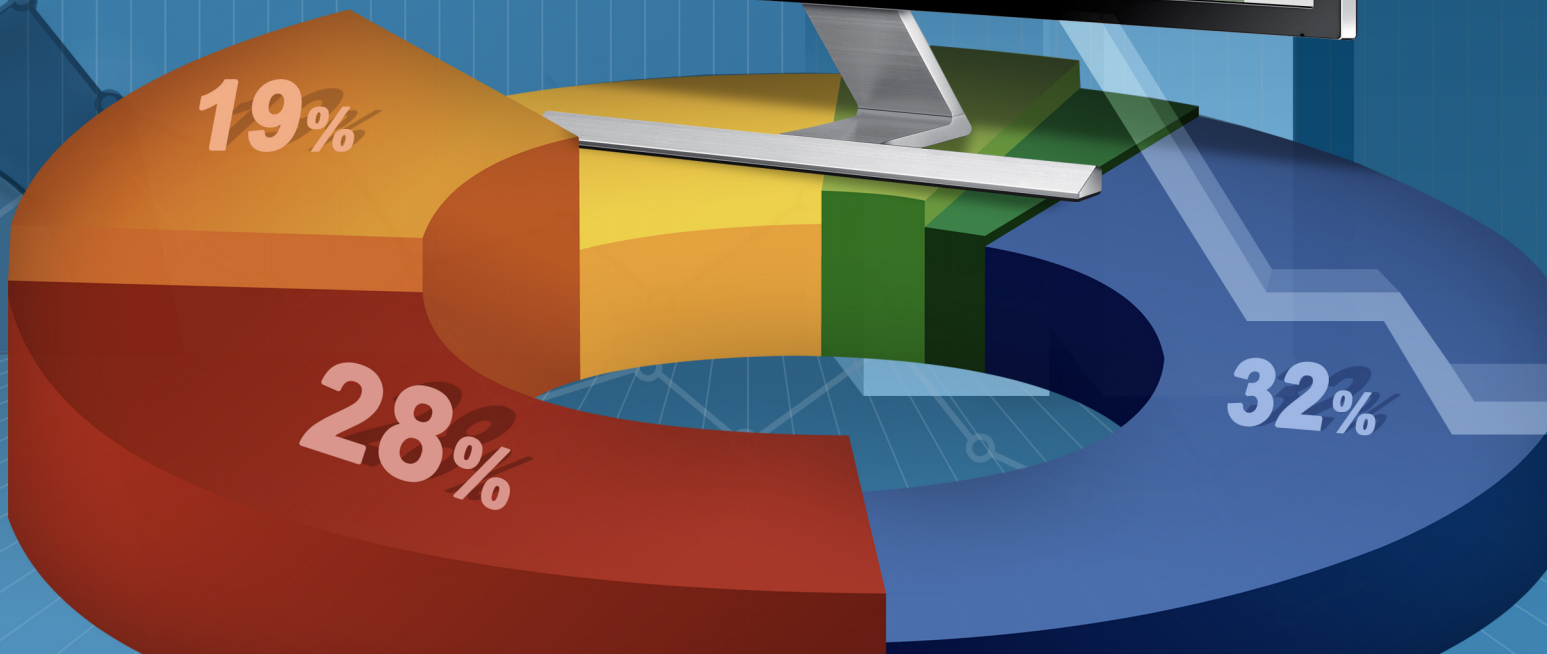
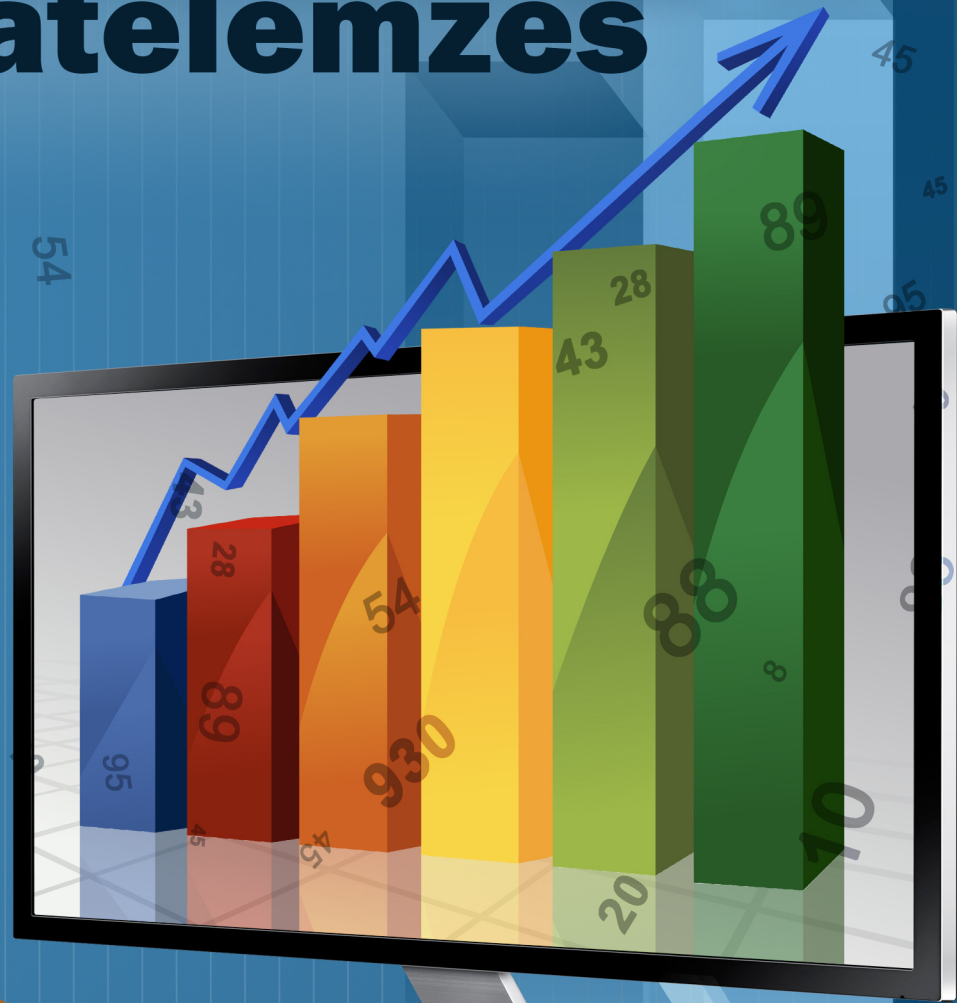




Ács Pongrác

GYAKORLATI adatelemzés



GYAKORLATI ADATELEMZÉS

Pécs, 2014.



PÉCSI TUDOMÁNYEGYETEM
UNIVERSITY OF PÉCS



GYAKORLATI ADATELEMZÉS

Alkotó szerkesztő: Dr. Ács Pongrác

Szerzők:

Dr. Ács Pongrác

Dr. Oláh András

Karamánné Dr. Pakai Annamária

Raposa László Bence

Lektorálta:

Dr. Balogh László

Felelős kiadó: Pécsi Tudományegyetem Egészségtudományi Kar

Műszaki szerkesztő: Varga Gábor

A tananyag a TÁMOP-4.1.2. E-13/1/KONV-2013-0012 pályázat támogatásával készült.

Copyright (szerzők és kiadó)

Pécs, 2014.

ISBN 978-963-642-682-8

A könyv a TÁMOP-4.1.2. E-13/1/KONV-2013-0012 című projekt keretében készült.

Nemzeti Fejlesztési Ügynökség
www.ujszechenyiterv.gov.hu
06 40 638 638



A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

TARTALOMJEGYZÉK

ELŐSZÓ	6
1. ELMÉLETI ÁTTEKINTÉS A TUDOMÁNYOS KUTATÁSOKRÓL (KARAMÁNNÉ PAKAI ANNAMÁRIA, OLÁH ANDRÁS)	9
1.1. BEVEZETÉS	9
1.2. AZ ISMERET FORRÁSAI A HÉTKÖZNAPI MEGISMERÉS ÉS TUDOMÁNYOS KUTATÓMUNKA SORÁN.....	9
1.3. A TUDOMÁNYOS KUTATÁS FOLYAMATA	11
1.4. A KUTATÁSI TÉMA KIVÁLASZTÁSA, A KUTATÁSI PROBLÉMA MEGHATÁROZÁSA	11
1.5. A KUTATÁS FŐBB CÉLJAI.....	15
1.6. A HIPOTÉZIS FELÁLLÍTÁS ALAPVETŐ KÖVETELMÉNYEI.....	16
1.7. A TUDOMÁNYOS KUTATÁS TERVEZÉSE ÉS AZ ELŐKÉSZÍTÉS FOLYAMATA.....	17
1.8. KONCEPTUALIZÁLÁS ÉS OPERACIONALIZÁLÁS.....	24
1.9. A KUTATÁS TÍPUSAI	27
1.10. A TUDOMÁNYOS KUTATÁS ETIKAI KÉRDÉSE	30
1.10.1. Embereken végzett kutatások.....	30
1.10.2. A hazai szabályozás.....	32
2. IRODALOMKUTATÁS A GYAKORLATBAN. IRODALOMKUTATÁS A LEGGYAKRABBAN HASZNÁLT ADATBÁZISOK FELHASZNÁLÁSÁVAL (KARAMÁNNÉ PAKAI ANNAMÁRIA, OLÁH ANDRÁS)	33
2.1. A RELEVÁNS SZAKIRODALOM GYŰJTÉSE ÉS FELTÁRÁSA	33
2.2. AZ IRODALOMKUTATÁS FORRÁSAI.....	34
2.3. AZ INTERNET HASZNÁLATA AZ IRODALOMKUTATÁS SORÁN	35
2.4. ELEKTRONIKUS KÖNYVTÁRAK.....	36
2.5. ELEKTRONIKUS ÚTON ELÉRHETŐ BIBLIOGRÁFIÁK.....	40
2.6. ON-LINE MÓDON ELÉRHETŐ SZAKMAI ADATBÁZISOK	44
2.6.1. EBSCOhost.....	45
2.6.2. MATARKA (Magyar Folyóiratok Tartalomjegyzékének Kereshető Adatbázisa).....	52
2.6.3. MOKKA (Magyar Országos Közös Katalógus).....	54
2.6.4. OVID	55
2.6.5. ScienceDirect.....	60
2.6.6. Scopus.....	63
2.6.7. SpringerLink.....	67
2.7. SZAKADATBÁZIS	69
2.7.1. MEDLINE.....	69
2.7.2. PubMed	70
2.7.3. SPORTDiscus	72
2.7.4. Semmelweis Egyetem Testnevelési és Sporttudományi Kar Digitális Könyvtár	72
2.8. ELEKTRONIKUS KÖNYVEK, ELEKTRONIKUS FOLYÓIRATOK	74
2.8.1. Open Access folyóiratok	74
2.8.2. Open Access Folyóiratok Könyvtára (Directory of Open Access Journals).....	75
2.9. INTERNETES KERESŐRENDSZEREK HASZNÁLATA A VILÁGHÁLÓN.....	77

2.10. IDÉZETEK ÉS SZÖVEGKÖZI HIVATKOZÁSOK	77
3. STATISZTIKAI ALAPFOGALMAK, AZ ADATOK FAJTÁI, VÁLTOZÓK ÉS ISMÉRVVÁLTOZÓK TÍPUSAI (ÁCS PONGRÁC).....	85
3.1. A STATISZTIKA MEGHATÁROZÁSA	85
3.2. STATISZTIKAI ADATOK.....	86
3.3. A VÁLTOZÓK CSOPORTOSÍTÁSA, SKÁLATÍPUSOK.....	87
4. ONLINE KÉRDŐÍVEK SZERKESZTÉSE A GYAKORLATBAN (RAPOSA LÁSZLÓ BENCE).....	91
4.1. BEVEZETÉS	91
4.2. A KÉRDŐÍV ELŐKÉSZÍTÉSE, ÁLTALÁNOS TUDNIVALÓK, A KÉSZÍTÉS ALAPJAI.....	92
4.3. KÉRDŐÍVSZERKESZTÉS A GYAKORLATBAN: A GOOGLE DOCS ALKALMAZÁS SEGÍTSÉGÉVEL	102
4.4. VÁLASZOK ÖSSZEGZÉSE, KIÉRTÉKELÉSE	109
4.5. KÉRDŐÍVSZERKESZTÉS EGY GYAKORLATI PÉLDÁN KERESZTÜL	112
5. AZ SPSS PROGRAM FELÜLETE, AZ ADATOK BEVITELE ÉS EXPORTÁLÁSA, MENÜPONTOK BEMUTATÁSA (ÁCS PONGRÁC).....	122
5.1. AZ SPSS PROGRAM FELÜLETE	122
5.2. ADATBEVITEL	127
5.3. MENÜPONTOK BEMUTATÁSA	131
6. ADATTISZTÍTÁS, EGYSZERŰBB MŰVELETEK A PRIMER ADATOKKAL, ADATMANIPULÁCIÓ (ÁCS PONGRÁC).....	140
7. LEÍRÓ STATISZTIKA, STATISZTIKAI TÁBLÁZATOK, STATISZTIKAI ÁBRÁK (ÁCS PONGRÁC, RAPOSA LÁSZLÓ BENCE).....	162
7.1. ELMÉLETI ALAPOK, LEÍRÓ STATISZTIKAI MUTATÓSZÁMOK	162
7.2. STATISZTIKAI TÁBLÁK, TÁBLÁZATOK	175
7.3. STATISZTIKAI ÁBRÁK, DIAGRAMOK.....	182
8. ÖSSZEFÜGGÉS-VIZSGÁLATOK (ÁCS PONGRÁC).....	203
8.1. STATISZTIKAI KAPCSOLATOK ÖSSZEFÜGGÉSEINEK ELMÉLETI ALAPJAI.....	203
8.2. ASSZOCIÁCIÓS KAPCSOLATOK VIZSGÁLATA, KERESZTTÁBLA ELEMZÉSEK.....	204
8.3. KORRESPONDENCIA- ANALÍZIS.....	223
8.4. VEGYES KAPCSOLAT VIZSGÁLATA	225
8.5. KORRELÁCIÓS KAPCSOLAT VIZSGÁLATA.....	228
9. KÖVETKEZTETÉSES VIZSGÁLATOK (ÁCS PONGRÁC).....	235
9.1. BEVEZETÉS, ELMÉLETI ALAPOK	235
9.2. A STATISZTIKAI BECSLÉSEK	238
9.3. KÜLÖNBÖZŐSÉG-VIZSGÁLATOK.....	244
9.4. HIPOTÉZIS-VIZSGÁLATOK (PARAMÉTERES ÉS NEM PARAMÉTERES PRÓBÁK A GYAKORLATBAN)	245
10. BEVEZETÉS A REGRESSZIÓ ANALÍZISBE (ÁCS PONGRÁC).....	270
10.1. KÉTVÁLTOZÓS LINEÁRIS REGRESSZIÓ	270

10.2. TÖBBVÁLTOZÓS LINEÁRIS REGRESSZIÓ.....	275
12. IRODALOMJEGYZÉK.....	280
13. MELLÉKLETEK	286
14. FÜGGELÉK (TÁBLÁZATOK)	288

ELŐSZÓ

Ennek az elektronikus tankönyvnek a megírásakor az alapvető célunk volt, hogy szilárd alapot nyújtsunk a szakdolgozatok írásakor leggyakrabban használt adatelemzési módszerek alkalmazásában. Tesszük ezt úgy, hogy tisztában vagyunk azzal, hogy az adatelemzések során alkalmazott és használt módszertanok ismeretanyaga olyan mértékben változott és hatalmasra duzzadt, hogy azoknak részletekbe menő összefoglalása, illetve a felhasználó oldaláról annak befogadása szinte lehetetlen lenne. Természetesen minket sem az a cél vezérelt, vagyis nem kívántuk és nem is tudnánk az összes ismeretet az olvasó elé tárni. A vezérelvünk sokkal inkább az volt, hogy a magyar felsőoktatásban minden képzési szinten kötelező szakdolgozat írásához gyakorlatban is használható módszertani segítséget adjunk. Tesszük ezt úgy, hogy számtalan BSc, MSc, valamint PhD képzésben vállaltunk és vállalunk témavezetést, illetve a kutatómódszertan kurzusok gondozása és oktatása során a hallgatókkal napi kapcsolatban vagyunk, így láthatóvá vált számunkra, hogy mely módszertanok preferáltak és alkalmasak a szakdolgozatok során a leggyakrabban felállított hipotézisek vizsgálatára. A tudományos kutatással kapcsolatos alapismeretekre a hallgatóknak a szakirodalom tanulmányozásához, a különböző beadandó dolgozataik és prezentációik, valamint a szak- illetve diplomadolgozatuk elkészítéséhez feltétlen szükségük van, ahogyan a TDK és OTDK munkák során sem megkerülhetők ezen ismeretek alkalmazása.

Fontos megjegyeznünk, hogy az elektronikus tankönyvünk, melynek fő témája az adatelemzés, nem tekinthető számítástechnikai tankönyvnek, viszont statisztikai tankönyvnek sem. Annak ellenére jelenthetjük ezt ki, hogy statisztikai módszertant alkalmazó elemzések bemutatása található meg benne túlsúlyban, viszont törekedtünk az alapvető elméleti részek mentén inkább a gyakorlati megközelítésre. Tettük ezt úgy, hogy a felvállalt és bemutatásra kerülő statisztikai részek is egyszerű és érhető formában kerüljenek az olvasó elé, viszont a szakmai hitelessége is végig megmaradjon. Véleményünk szerint a könyv önmagában is alkalmas a kitűzött módszertanok szemléltetésére, vagyis nem szükséges plusz statisztikai tankönyv a megértéshez. Azoknak az olvasóknak, akik mélyebben a témával ismerkedni szeretnének, ajánlásokat adunk, hogy azokat hol keresse.

A tankönyv szerkezeti felépítésében a tudatosság megfigyelhető, hiszen az első *(Elméleti áttekintés a tudományos kutatásokról)* fejezetben az olvasó alapvető információkat kap a tudományos kutatás fajtáiról, menetéről, kutatási célokról. Fontosnak tekintettük, hogy itt is a gyakorlati megközelítés megmaradjon a fő célunk, ezért egy részletesen kidolgozott szakdolgozati kutatási tervet is szemléltetünk az olvasókkal, mely nagy segítséget jelenthet azoknak, akikre ez a feladat a jövőben vár.

A második fejezetben a kutatások és szakdolgozatok alapjául szolgáló irodalomkutatás (*Irodalomkutatás a gyakorlatban. Irodalomkutatás a leggyakrabban használt adatbázisok felhasználásával*) kerül részletes bemutatásra. Törekedtünk arra, hogy a leggyakrabban használt elektronikus felületeket bemutassuk, mely által az olvasó saját maga is könnyebben tudhat a saját témájában fontos irodalmakhoz eljutni.

A harmadik részben (*Statisztikai alapfogalmak, az adatok fajtái, változók és ismérvváltozók típusai*) az adatelemzések során nélkülözhetetlen statisztikai alapfogalmak kerülnek röviden bemutatásra.

A negyedik fejezetben (*Online kérdőívek szerkesztése a gyakorlatban*) a primer adatbázisok előállítása során az egyik leggyakrabban használt módszer, az online kérdőív bemutatására tettünk kísérletet. Az alapvető információk és kérdéstípusok bemutatása után, egy preferált program segítségével mutatjuk be, miként kell egy ilyen konkrét kérdőívet összeállítani, és azokból kapott információkat adatbázisba rendezni. Érdekessége a bemutatott példának, hogy a tankönyv későbbi részeiben felhasznált adatbázis online kérdőív formában való megjelenítését demonstrálja.

Az ötödik fejezettől kezdődően sok más hasonló tankönyvhöz hasonlóan az SPSS program bemutatása következik, az abban legpreferáltabb adatelemzési módszerek szemléltetésén keresztül. Az elektronikus tankönyvben használt adatbázisok a www.etk.pte.hu weblapon az olvasó rendelkezésére állnak. Természetesen a program bemutatása során is az egymásra épülést céloztuk meg, melyet a fejezetek építkezése is jól mutat. Ennek megfelelően az ötödik részben (*Az SPSS program felülete, az adatok bevitele és exportálása, menüpontok bemutatása*) a program felépítése, alapvető funkciói kerülnek tárgyalásra.

A hatodik fejezetben (*Adattisztítás, egyszerűbb műveletek a primer adatokkal, adatmanipuláció*) az alapvető „adatmanipulációk” szemléltetése történik, gyakorlati példákon keresztül.

A hetedik részben (*Leíró statisztika, statisztikai táblázatok, statisztikai ábrák*) található a leíró statisztika elméleti és gyakorlati bemutatása, melyben külön kitértünk az adatprezentációk során leggyakrabban használt statisztikai táblázatok és ábrák előállításának bemutatására.

A nyolcadik fejezet (*Összefüggés-vizsgálatok*) a változók között alapvető kapcsolatok, összefüggések bemutatására tesz kísérletet.

A kilencedik részben (*Következtetési vizsgálatok*) mindazon vizsgálatok gyakorlati szemléltetése történik, melyet a leginkább fellelhetünk ebben a témakörben (statisztikai becslések, hipotézisvizsgálatok) a kutatási adatok feldolgozása során. Természetesen ennél a fejezetnél is szükség van alapvető elméleti tudásra, melyet a fejezet elején röviden leírunk.

Az utolsó tizedik fejezet (*Bevezetés a regresszió analízisbe*) rövid betekintés enged a regresszió analízis széles tárházába, rövid gyakorlati példafeldolgozáson keresztül.

Segítő szándékkal engedje meg az olvasó, hogy felhívjam a figyelmét arra a tényre, hogy a tankönyv többszöri olvasása sem jelenti azt, hogy egy kutatási problémát azonnal meg fog tudni oldani, ugyanis előtte a számítógépen történő gyakorlás elengedhetetlen. Kutatói tapasztalatunk alapján teljes bizonyossággal kijelenthetjük, hogy a számítógép előtt, gyakorlással töltött órák száma erős korrelációt mutat a sikeres probléma megoldásokra.

Kétségtelen tény, hogy az orvos és egészségügy, sporttudomány és szinte valamennyi tudományterület eredményei szinte „kimeríthetetlen kincsesbányái” a kutatóknak, érdeklődőknek, ami jelenti azt is, hogy az adatokban rejlő információkat kritikai szemmel vizsgálhatjuk és elemezhetjük, majd lehetőségünk van az eredményeket, következtetéseket kiadni, publikálni.

Fontos megemlítenem, hogy külön hálás vagyok szerzőtársaimnak, akik szem előtt tartva elképzeléseimet mindig pontos, precíz és egyben gyakorlatban használható részeket írtak, melyek nagyban emelik az elektronikus tankönyv minőségét. Szeretnék köszönetet mondani kollégáinknak, munkatársainknak, akik bátorítottak, és ötleteikkel inspiráltak a könyv megírására. Köszönet illeti azon tanáraink, mentoraink alapos munkáját is, akik elindítottak és a mai napig támogatnak bennünket. Külön köszönet illeti Makai Alexandra PhD hallgatót, aki elvégezte a fásasztó és hálátlan szövegkorrekciót és a javításokat.

Hálával tartozom továbbá neves lektorunknak is, aki kellő kritikussággal és szakmaisággal javította és segítette a végleges formába való megjelenést.

Szeretettel tisztelgek Dr. Pintér József Tanár Úr emlékének, aki felhívta figyelmemet a kutatómódszertan és azon belül is a sporttudományi kutatások mostoha helyzetére.

Meg szeretném kérni Kollégáimat, Hallgatóimat, hogy észrevételeikkel, javaslataikkal támogassák a könyv folyamatos karbantartását, aktualizálást, tegyék lehetővé, hogy az esetleges bennmaradó (és csak a szerzőt terhelő) hibákat kijavíthassam!

Pécs, 2014. szeptember 22.

Ács Pongrác
szerző

1. ELMÉLETI ÁTTEKINTÉS A TUDOMÁNYOS KUTATÁSOKRÓL

(Karamánné Pakai Annamária, Oláh András)

1.1. Bevezetés

Napjainkban a tudományos területen dolgozó szakemberekkel szemben alapvető elvárás, hogy bizonyítékokon alapuló, kutatómódszertani és biostatisztikai ismeretekre alapozott, releváns hazai és nemzetközi kutatások eredményeire támaszkodó vizsgálatokat végezzenek. Ezen ismeretek nemcsak hozzájárulnak az önálló kutatások sikeres kivitelezéséhez, de nagymértékben elősegítik a tudományos konferenciákon, szakmai rendezvényeken tartott előadásokon elhangzó és folyóiratokban megjelenő publikációk megértését és azok kritikus értékelését is (Lampek – Kívés 2012, Pakai – Kívés 2013).

1.2. Az ismeret forrásai a hétköznapi megismerés és tudományos kutatómunka során

A *tudományos kutatás* a probléma megoldására irányuló tervszerű, adott helyen és időben végzett, tudományos ismereteken és előre felállított hipotéziseken alapuló emberi tevékenység. A fogalom jól tükrözi, hogy a tudományos munka egy olyan szisztematikus tevékenység, ahol racionális gondolkodás, bizonyítékokon alapuló állítások megfogalmazása és logikus érvelés szükséges.

Az ember alapvető törekvése, hogy a körülötte lévő világot megismerje és megértse. A valóság megismerésének kétféle formája különíthető el. A *tapasztalati valóság* azok a dolgok, amit az ember az élete során saját közvetlen tapasztalatából származik. Példaként említhető az egészségügyben dolgozó szakember, aki a mindennapi tevékenységét a páciensek körében tölti, megismeri azok személyiségét, magatartását, a betegségekre való reagálását vagy a feladatainak ellátása során tanulja meg az akut, életveszélyes helyzetben hogyan kell gyorsan dönteni, helytállni. A személyes tapasztalatok és megfigyelések által tud hivatkozni és megbizonyosodik az információk valóságtartalmáról, megbízhatóságáról és objektivitásáról. Hasonlóan a sport területén az edzők szerepe, hiszen az edzéseken, versenyeken tapasztalt és „megélt” szituációkból tanul, és ezt használja fel a későbbi döntései során. A *konszenzuális valóság* azt jelenti, hogy az ember a saját kritériumrendszere alapján a mások által közvetített információt, tudást elfogadja és egyetért vele. Például a mozgásterápia során a páciens gyógyulása érdekében a gyógytornász tanácsait, javaslatait követi.

A mindennapi életben is lehet különböző megfigyeléseket, a kíváncsiság kielégítése céljából akár kisebb kísérletek is elvégezni, azonban e tevékenységek során számtalan hiba fordulhat elő. Általánosságban elmondható, hogy a hétköznapi megfigyelések véletlenszerűek, felületesek és csak félig meddig tudatosak. Erre jó példa, hogy nem biztos, hogy fel tudja pontosan idézni, hogy a legutolsó tanórán a tanárnő milyen szoknyát viselt, egyáltalán szoknyában volt-e. A hallgatók jelentős része erre a kérdésre valószínűleg nem tudna pontos leírást adni, ennek pedig alapvető oka a **pontatlan megfigyelés**, amelyből a következő hibaforrás is előrevetíthető. A **túláltalánosítás** olyan esetben fordul elő, amikor néhány hasonló eset után egy általános összefüggést bizonyítéknak tekintenek. Ilyen esetek gyakrabban időhiány miatt fordulnak elő, például, ha csak egy-két kollégistát kérdeznek meg a kollégiumok állapotáról és ebből vonnak le általános következtetéseket. A túláltalánosításból következhet az újabb hiba, **a szelektív észlelés**. Ha a mindennapi élet során megállapításra került egy összefüggés és arra egy elmélet is született, hajlamos lesz rá, hogy a közeljövőben csak olyan eseményekre és helyzetekre fókuszál, amelyek ezt az összefüggést igazolják, ezáltal figyelmen kívül marad sok más tapasztalat. Például a faji és etnikai előítéletek kialakulásában is nagy szerepet játszik ez a tényező. A hétköznapi megismerésben jellemző még a **hozzáköltés és az illogikus gondolkodás** is, amikor az ember számára érthetetlen dologra nem talál releváns magyarázatot. Példa, amikor valaki felkészül arra, hogy a sok jeles osztályzat után egy rosszabbnak, akár elégtelennek kell következnie. Gyakran előfordul az is, hogy a világról alkotott előzetes tudás, ismeret meghatározza az adott helyzet értelmezését, így **elfogulttá válik annak megértésében**. Ha egy hallgató elégtelen érdemjegyet kapott a kollokvium során sokkal hamarabb fogja azt a magyarázatot adni, hogy a tanár szándékosan tett fel olyan kérdést, ami nem is szerepelt a tételsorban. Számtalan olyan helyzet, esemény is szerepel az ember életébe, amit nem ért vagy nem tud megmagyarázni. Ennek következtében bizonyos dolgokat, érthetetlen jelenséget megismerhetetlennek tekint, vagy misztikus jellemzőt tulajdonít a számára. **Misztifikáció** a hallgatók körében a vizsgaidőszakban különösen jellemző, amikor egy-egy kabala nélkülözhetetlen kelléktára a vizsgáknak. A sportolók körében a kabalákon túl, gyakori a megszokott ruhadarabok használata, mint például Király Gábor sűrű hosszú szárú melegítője.

A fenti hibákat figyelembe véve a tudományos munkát végző szakemberrel szemben alapvető elvárásként fogalmazódik meg, hogy szelektáló képességgel rendelkezzen, a valóság egy részének megismerésére célirányos, megtervezett és adekvát módszereket alkalmazzon és állításainak logikailag és empirikusan alátámasztottnak kell lennie. Ennek következtében

ritkábban fordulnak elő a hétköznapi megismerést jellemző megismerési hiányosságok is. Egy tudományos kutatás során a megfigyelés tudatos tevékenység. Ha a megfigyelés szándékos, akkor a hibák előfordulási aránya is jóval kevesebb lesz. Tudatos megfigyelést különböző, mérésre szolgáló segédeszközök alkalmazásával lehet objektiválni. A túláltalánosítás elkerülése miatt a kutatást végző szakemberek ügyelnek arra, hogy a megfigyeléseiket nagyszámú mintán végezzék el, illetve törekednek a vizsgálat ismételt elvégzésére is. A tudományos munka során a szelektivitás úgy előzhető meg, ha a kutató egy átgondolt, mindenre kiterjedő kutatási tervet készít, melyben pontosan meghatározza a megfigyelésének célját és annak szempontjait, következtetéseit az eredmények alapján teszi. Ha a kutató ellentmondó tényeket lát, további tudományos kutatást kezdeményez. A tudományos munkát végzők nemcsak az adatgyűjtés során, de az okfejtésben is tudatosak és óvatosak. A tudományos nyilvánosság arra is lehetőséget ad, hogy ugyanazt a munkát más kutatói team is elvégezheti és leellenőrizheti. Tudományos alapelvként jelenik meg, hogy minden megismerhető, ezért nincs helye a misztifikációnak. Emellett soha nem szabad egy kutatást idő előtt lezárni, mert végső soron minden kérdést nyitottnak kell tekinteni. Napjainkban gyakran tapasztalható, hogy újabb és újabb tudományos eredmények cáfolják meg az addigi eredményeket.

Röviden összefoglalva megállapítható, hogy hétköznapi megismerés a mindennapi élettevékenységek folyamán történik, az így nyert adatok, ismeretek véletlenszerűek, ezzel szemben a tudomány tudatos, tervszerű, szisztematikusan felépített tevékenység, ahol a kutató tisztában vannak a tévedések kockázatával, de tudatosan védekezik ellene. (Babbie 2004, Bíróné et al 2011, Gőcze 2011).

1.3. A tudományos kutatás folyamata

A kutató a tudományos kutatás során a probléma meghatározás, célkitűzés és hipotézis megfogalmazása után tervszerűen, a kutatás szabályait szem előtt tartva adatokat gyűjt, majd ezeket az általa választott statisztikai program segítségével feldolgozza, elemzi, majd tudományos munka tartalmi és formai követelményeit szem előtt tartva mások számára is közzé teszi (Lampek – Kívés 2012). A tudományos kutatás lépéseit az 1/1. ábra szemlélteti.

1.4. A kutatási téma kiválasztása, a kutatási probléma meghatározása

A kutatási folyamat első lépése a kutatási téma kiválasztása és a probléma meghatározása. A témaválasztást egyaránt befolyásolja a kutató érdeklődési területe, motivációja, a

szakirodalomban való jártassága valamint az elvégezendő feladat célja (például szakdolgozat készítés). A témaválasztás során felmerülő probléma adódhat a mindennapi gyakorlatból (amikor egy megoldatlan probléma kerül az érdeklődés középpontjába), a választott diszciplína elméleti tételeiből vagy a releváns szakirodalomból. Felsőoktatási intézményben tanuló hallgatók esetében a témaválasztást elősegíti az előadásokon vagy a szakmai gyakorlatokon felmerülő kérdések, a hallgatók által korábban írt szakdolgozatok választott témái, de jelentős támaszt nyújt az adott intézet és tanszék által kiadott szakdolgozati és Tudományos Diákkör (TDK) témakör listája is.

A kutatás megkezdése előtt azonban néhány alapvető kérdést érdemes tisztázni. A kutatni kívánt problémát több oldalról is át kell tanulmányozni abból a célból, hogy egyáltalán szükséges-e az adott vizsgálat elvégzése. A kérdésre a választ a témában megjelenő hazai és nemzetközi publikációk eredményei adhatják. Lényeges szempont a megfelelő keresési technikával feltárt szakirodalom kellő mélységű és alapos áttanulmányozása. Az irodalomkutatásról bővebb ismereteket az *„Irodalomkutatás a gyakorlatban. Irodalomkutatás a leggyakrabban használt adatbázisok felhasználásával.”* című fejezetben kaphatunk.

A releváns szakirodalmak áttekintése több szempontból is hasznos lehet a kutató számára, nemcsak segít pontosítani a probléma megfogalmazását, de az eredmények ismerete új kutatásra inspirálhatja a szakembert.

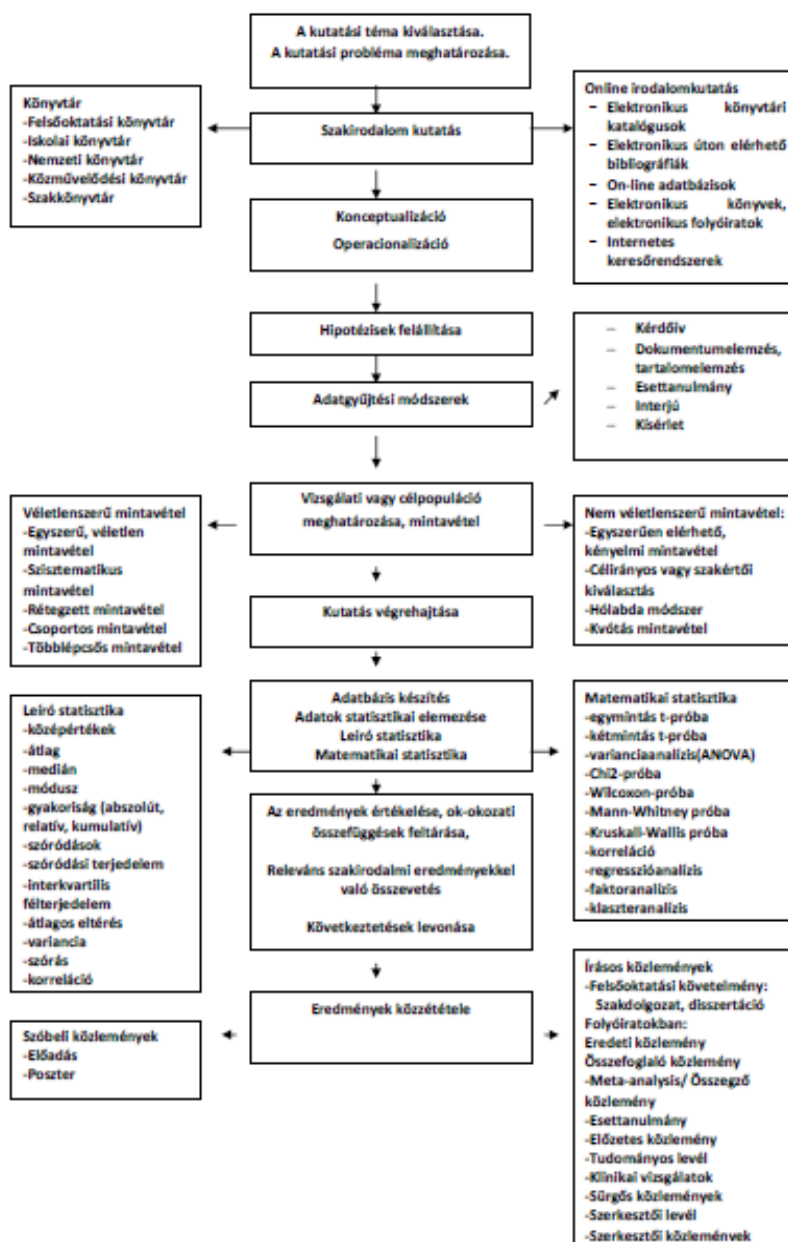
A választás során célszerű meggyőződni a választott téma aktualitásáról is. A szakirodalom áttekintése az olvasó számára további pontos képet nyújt arról, hogy a választott terület, érdeklődésre számot tartó. Amennyiben a témában nagyon sok vizsgálat született, célszerű lefedetlen területet találni vagy a szakirodalmak során felmerült kérdések mentén tovább haladni.

Meg kell vizsgálni azt is, hogy a kutató felkészültsége, szakmai ismerete elegendő-e a téma feltárására, rendelkezik-e azokkal a képességekkel, ami a kutatás tervezéséhez, a hozzá kapcsolódó szakirodalom feltáráshoz, az eredmények statisztikai feldolgozásához és az okozat feltárásához egyaránt szükséges. Ma már nélkülözhetetlen az idegen nyelv(ek) ismerete is, hiszen a releváns szakirodalmak többsége angol nyelven íródik, és magyar nyelven nem érhető el.

A kutatásban jártas szakembernek azonban már azt is szem előtt kell tartania, hogy a kutatás során kapott eredményeket mennyire lehet majd alkalmazni a gyakorlatban, elősegíti-e az adott gyakorlati munka hatékonyságának növelését.

A végleges választás, döntés előtt célszerű a határidőkkel is tisztában lenni. Alapos átgondolást igényel, hogy elegendő idő áll-e rendelkezésre a kutatás elvégzésére és pontosítani kell a benyújtás dátumát.

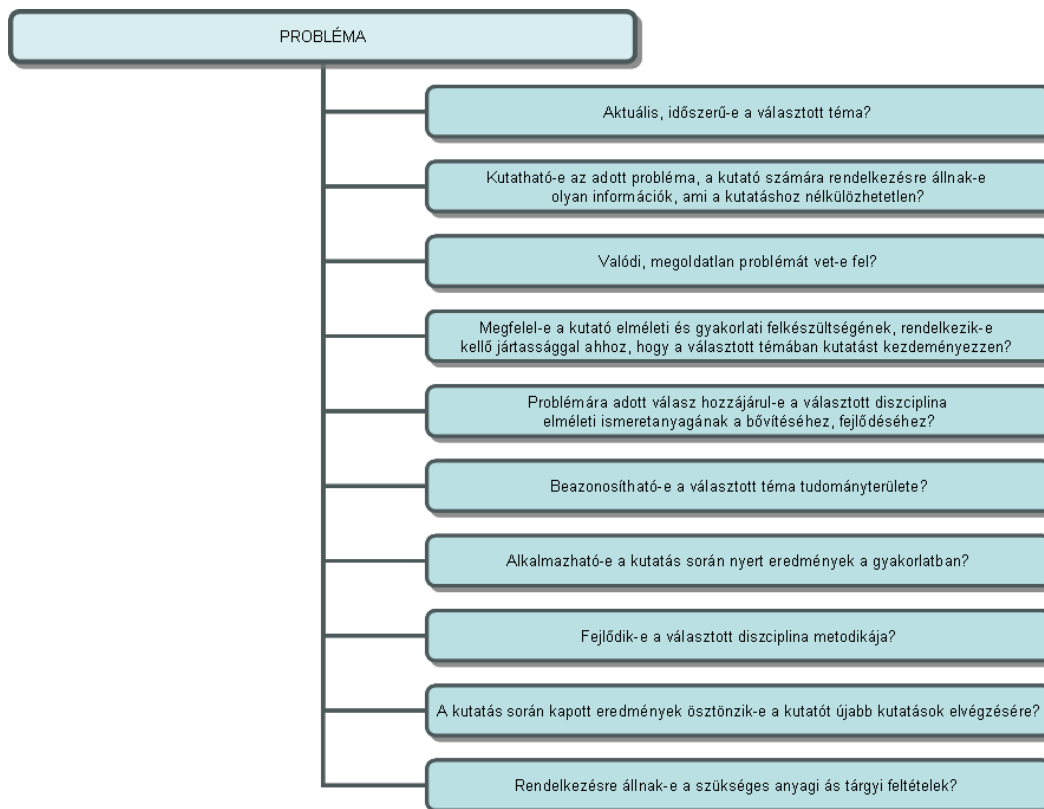
Amennyiben a 1/2. ábrán is látható kérdésekre választ adott a kutató és a szakirodalom keresés során meggyőződött arról, hogy a vizsgálat középpontjában álló kérdésben nem állnak rendelkezésre bizonyítékokon alapuló evidenciák, a hazai viszonylatban is megválaszolandó kérdések merülnek fel, a kutatás lebonyolítása mindenképpen ajánlott. A gyakorlott kutató a szakirodalmi találatokból már pontos képet kap arról, hogy a szakma mely területekre fókuszál, milyen témák élveznek prioritást, ami később a publikációs lehetőségeket is befolyásolhatja.



1/1. ábra A kutatási folyamat lépései

Forrás: Lampek – Kívés 2012, Betlehem 2010, saját szerkesztés

A felsorolt ajánlások ellenére előfordulhat, hogy a probléma megfogalmazása, annak pontos meghatározása mégis nehézségbe ütközik, ilyenkor ajánlott megfigyelést végezni a választott témában, további szakirodalmi kutatást folytatni vagy egy témában járatos, szakmai tapasztalattal rendelkező szakembert felkeresni, aki segít azonosítani, pontosítani a problémát (Papp 2013, Cserné Adermann 1999, Falus 2004, Ács 2009).



1/2. ábra A kutatási probléma aktualitásának vizsgálata

Forrás: Falus 2004 nyomán
saját szerkesztés

1.5. A kutatás főbb céljai

Egy kutatás különböző célból valósulhat meg úgy, mint a kutató valamilyen jelenség leírására vagy törvényszerűségek feltárása törekszik. Alapvetően három típust lehet elkülöníteni. A **felderítő kutatás** akkor végezhető, ha a kutató egy adott területről csak tájékozódni kíván, nem ismeri mélyrehatóan. Mivel nem rendelkezik elegendő információval, a vizsgálat alapvető célja annak megállapítása, hogy vajon érdemes-e egy alaposabb, mélyreható vizsgálatot elvégezni. Egy felderítő kutatás azonban abban az esetben is elvégezhető, ha egy új módszer hatékonyságát szeretnének kipróbálni, amit a későbbiekben a gyakorlatban is alkalmazni lehet. Ezek a kutatások többnyire nem nagy volumenűek, így az ezzel kapcsolatos költség sem számottevő. Eredményeikből messzemenő következtetés nem vonható le, azonban egy következő kutatás alapját képezheti, mint például segít eldönteni, hogy érdemes-e a témával tovább foglalkozni, alapot nyújt a kutatási probléma konkrét meghatározásába, a hipotézisek felállításában és a megfelelő adatgyűjtési módszer kiválasztásában is. A felderítő kutatások során a megfigyelések, az interjúk és esettanulmányok készítését részesítik előnyben. Erre jó példa a Kútvölgyi Klinikai Tömbben végzett vizsgálat, ahol a vizsgálat középpontjában az ápolói kompetenciába tartozó gyógyszerosztás folyamatát helyezték. A

szerzők direkt megfigyelési módszerrel kívánták a kórházi gyógyszeresztás folyamatához kapcsolódó gyógyszerelési hibák előfordulási gyakoriságát feltárni és a gyakorlatban előforduló hibatípusokat azonosítani (Lám et al 2011).

A **leíró kutatás** célja egy vizsgált jelenség, esemény, egy helyzet vagy egy kiválasztott csoport pontos és részletes bemutatása. A leíró kutatás jó alapot nyújthat egy későbbi magyarázó célból kezdeményezett kutatásnak. Példaként emelhető ki Boncz Imre et al (2013) publikációja, amelyben az Országos Egészségbiztosítási Pénztár finanszírozási adatbázisából, a szervezett országos emlőszűrési program 2008-2009. évi részvételi mutatóit elemzi. A 45-65 éves nők körében meghatározásra került azok aránya, akik a program 4. ciklusában (2008-2009) akár szűrési, akár diagnosztikai célú képalkotó emlővizsgálaton vettek részt.

Magyarázó kutatást ok-okozati kapcsolatok feltárása céljából végeznek. Ebben az esetben a vizsgált esemény, jelenség, csoport nem csak leíró statisztikával kerül bemutatásra, hanem bizonyos jellemzőket magyarázni is kíván, közöttük oksági kapcsolatot feltételez. Például Kopp Mária vezetése alatt a Hungarostudy 2002 országos, a magyar felnőtt lakosságra életkor, nem és településnagyság szerint reprezentatív egészségfelmérést végeztek. A monográfia egyes fejezeteiben a felmérésben résztvevő kutatók, az adott témában ismertetik legfontosabb, ok-okozati kapcsolatokat feltáró eredményeiket (Kopp – Kovács 2006).

A felsorolt három kutatási típus nem különül el élesen egymástól, a kutatás más-más fázisában eltérő hangsúlyt kaphatnak. A mindennapi gyakorlatban, egy nagyobb volumenű kutatás azonban mindhárom típust magába foglalhatja.

1.6. A hipotézis felállítás alapvető követelményei

A hipotézis a kutatás elvárt eredményeiről szóló előzetes feltevés, állítás, melyet a kutatás elvégzése során a megfelelő adatgyűjtési és statisztikai módszer segítségével igazolni vagy cáfolni lehet. A hipotézisalkotás induktív vagy deduktív módon történhet. Az **induktív hipotézisalkotás** során a gyakorlatban megfigyelt törvényszerűségek előfordulását feltételezik, míg a **deduktív módszer** esetén az elméleti tételből kerül levezetésre a hipotézis, vagyis az általános elméletből halad a speciális esetek felé. A hipotézis felállításával kapcsolatos alapvető követelményrendszer:

- A kutatás megkezdése előtt, mint vezérfonalként kell felállítani.
- A meglévő ismeretekre támaszkodva egyszerűen, tömören és állítás formájában (kijelentő mondat) kerüljön megfogalmazásra.
- Jelölje a függő és független változók közötti kapcsolatát.

- Egyértelmű, világosan megfogalmazott fogalmakat kell használni.
- A felállított hipotézisek igazolására megfelelő adatgyűjtési módszerek álljanak a rendelkezésre.
- A kiválasztott adatgyűjtési módszer feldolgoása során az alkalmazott statisztikai próbákkal egyértelműen igazolható, vagy elvethető legyen.
- Helytálló (igazolt) hipotézisek megtartásra kerüljenek.
- Összességében a kutatás során megfogalmazott problémára összességében választ kell adnia.

Egy kutatás során felállított hipotézisek számát a kutatás céljai és annak körülményei egyértelműen meghatározzák. Például egy szakdolgozat elkészítése során ideális, ha maximum 5 hipotézis kerül felállításra.

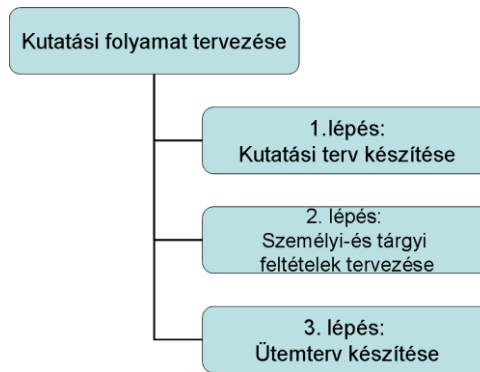
A következő minta a hipotézis felállításához szükséges alkotóelemeket szemlélteti. A hipotézis felállítása során alapvető feladat a függő és független változók közötti kapcsolat meghatározása. A közöttük lévő ok-okozati kapcsolatban a **függő változó** az okozat, míg a **független változó** az ok. A kutatások alapjait képezi, hogy a vizsgálat során a független változó hogyan határozza meg a függő változót. A következő példa is jól szemlélteti a függő - független változók kapcsolatát a felállított hipotézisben:

- A Triclosan varróanyag alkalmazása során kisebb eséllyel fordul elő sebgyógyulási zavar, mint a hagyományos varróanyagoknál (Pakai et al 2013).

A példa alapján a függő változónak tekintjük a sebgyógyulási zavart, amit a független változó, vagyis a varróanyag típusa (Triclosan, hagyományos) befolyásolja.

1.7. A tudományos kutatás tervezése és az előkészítés folyamata

A kutató munka első fázisában és a hipotézis felállítása során számos megválaszolandó kérdésre váró kérdés merül fel, mint például Mit kíván vizsgálni a kutató: az alapsokaságot vagy egy mintát?; milyen adatgyűjtési módszert célszerű alkalmazni (kérdőív, dokumentumelemzés, kísérlet stb.)?; milyen változókkal lehet a felállított hipotézist számszerűsíteni?; milyen statisztikai módszerekkel lehet a hipotéziseket alátámasztani?; milyen korlátai lehetnek az adott kutatásnak?. Mindezen kérdések a kutatót arra ösztönzik, hogy a kutatás folyamatát tervszerűen átgondolja, és egy átfogó, mindenre kiterjedő kutatási tervet állítson össze. A kutatási terv készítése három fő területre fókuszál (1/3. ábra).



1/3. ábra. A kutatás folyamat megtervezése

Forrás: Hornyacsek 2013 nyomán
saját szerkesztés

A *tervezés első területe* a kutatás legfontosabb alapelemeinek meghatározása, melynek egy tervezetben mindenképpen szerepelnie kell (1/1. táblázat):

- Probléma és a vizsgálat céljának megfogalmazása. Röviden, 5-10 mondatban érdemes leírni releváns szakirodalommal alátámasztva: Mit kíván vizsgálni? Miért érdemes a választott témában kutatást kezdeményeznie? Milyen gyakorlati jelentőséggel bír?
- Hipotézis(ek) felállítása. A hipotézisek száma függ a választott kutatás céljától és a körülményeitől. Egy szakdolgozat készítés során minőségi kutatás érdekében ötnél több hipotézis felállítása nem célszerű.
- A kutatás típusa. A kutatási tervben csak felsorolás jelleggel kell leírni, hogy kvantitatív vagy kvalitatív, prospektív vagy retrospektív, longitudinális vagy keresztmetszeti stb. vizsgálatot kíván folytatni.
- Célcsoport rövid, lényegre törő leírása.
- A beválasztási és kizárási kritériumok. A tervezés során már célszerű átgondolni azokat a jellemzőket, tulajdonságokat, ami biztosítja a minta homogenitását.
- Mintaválasztás során el kell dönteni, hogy a célcsoporton belül véletlenszerű-, vagy nem véletlenszerű mintaválasztás kerül alkalmazásra.
- Az esetcsoport, és ha van, a kontrollcsoport elemszáma.
- A vizsgálat pontos helyszínének meghatározása.
- A kutatás tervezett ideje (időintervallumban meghatározva).
- Az adatgyűjtési módszerek és eszközök kiválasztása (kérdőív, dokumentumelemzés, tartomelemzés, esettanulmány, interjú, kísérlet).
- A vizsgált (függő- és független) változók bemutatása.
- A statisztikai elemzés módjának ismertetése.

- A várható eredmények rövid összefoglalása.

A kutatási terv készítése során érdemes az irodalomkutatás alkalmával talált releváns szakirodalmak feldolgozását is elkészíteni. A publikáció bibliográfiai adatai mellett az alábbi fő területekre érdemes fókuszálni a releváns cikk feldolgozása során: a cikk háttérproblémájának rövid összefoglalása, a kutatás típusának, a vizsgálat helyének és idejének meghatározása, a mintaválasztás módja és a minta rövid ismertetése, az adatgyűjtési módszer felsorolása és a statisztikai próbák és legfontosabb, a kutató számára értékes eredmények rövid leírása (1/1. táblázat). A **második tervezési terület** a személyi és tárgyi feltételek átgondolása. Személyi feltételként alapvető elvárás, hogy a kutató szakmai felkészültsége megfelelő legyen, a választott témában kellő jártassággal rendelkezzen, mind hazai, mind a nemzetközi szakirodalmat széles körben ismerje, emellett rendelkezzen az általánosítás, szelektálás és a rendszerezés képességével, és munkáját hatékonyság, kitartás jellemezze. A kutatás vezetőjének emellett azt is át kell gondolni, hogy ki vegyen részt a kutató team munkájában. Nagyobb volumenű vizsgálat esetén szükséges lehet olyan kollégák munkájára, akik segítenek az adatgyűjtésben, a statisztikai adatok bevitelében. Érdemes a vizsgálatba olyan szakembereket is invitálni, akik a statisztikai adatok elemzésében és az eredmények publikálásában kellő jártassággal rendelkeznek. A kutatási téma igényelheti a társszakmák képviselőinek bevonását is. A kutatás végrehajtásához rendelkezésre kell állnia a szükséges tárgyi feltételeknek, eszközöknek és berendezéseknek egyaránt. A kutatás költségeinek fedezetére is elő kell teremteni a szükséges anyagi forrásokat. Ezek a források származhatnak az adott intézmény kutatásra fordítható összegéből, pályázati pénzekből, alapítványok vagy más szervek támogatásából. A **tervezés harmadik területe** egy ütemterv elkészítése, ahol a kutatás menetének egyes lépései mellé az általa szükséges időpontok és határidők is meghatározásra kerülnek. A 1/3. számú táblázat röviden összefoglalja egy szakdolgozat/diplomamunka készítésének ütemtervét. A megfelelő ütemterv elősegíti a sikeres dolgozat elkészítését, segíti a dolgozat készítőjét a tervezett dátumokat betartására, rámutat a párhuzamos feladatok elvégzésére, feltárja a megvalósíthatóság nehézségeit és korrekció elvégzésére is ad lehetőséget (Hornycsek 2013).

1/1. táblázat Egy szakdolgozat kutatási terve (minta)

A kutatás témája/címe: Egy multinacionális cég egészségfelmérése				
<p>Probléma meghatározás: A XX. század végére nőtt a daganatos, a szív- és keringési rendszeri megbetegedések száma; a munkahelyi stressz fokozódott, az ezekkel kapcsolatos megbetegedések aránya is magasabb lett. Tudatosítani kell az emberekben, hogy munkakörülményeik jelentős hatással bírnak az általános egészségügyi kondícióikra, ezért az elsődleges szempont az érdeklődés felkeltése az egészséges életmód, egészségfejlesztő programok iránt.</p> <p>A jelentős kockázattal és veszéllyel járó munkafolyamatok is magasabb stressz-szintet eredményeznek. Ezek a munkahelyeken a szervezetnek állandó készenlétben kell lennie a gyors reagálásra. Az állandó feszültség, a növekvő adrenalin szint, a légzés szabálytalansága, az izmok állandó ingerlése is vezethet hosszútávon egészségkárosodáshoz. Ezen faktorok figyelmen kívül hagyása később mind a munkaadónak, mind a munkavállalónak, de közvetve az egész társadalomnak is sokba kerülhet. A rendszeres testmozgás hozzájárul az egészségesebb életvitelhez, a jó közérzethez, a munkateljesítmény fokozódásához. Segít megelőzni a szívpanaszokat, kontrollálni a testsúlyt, normális szinten tartani a vérnyomást, segít megszabadulni a feszültségtől, a stressztől, általános jó közérzetet biztosít, s hosszútávon a nyugdíjas években is az egészségesebb, aktívabb élethez járul hozzá.</p>				
<p>Célkitűzés: Jelen vizsgálat célja egy multinacionális cég termelő tevékenységében és adminisztratív munkakörében alkalmazott dolgozóinak egészségi állapotfelmérése belgyógyászati és mozgásszervrendszeri szempontból, melyre a főbb problémakörök kiszűrése után felvilágosító, egészségnevelő, egészségfejlesztő előadások, bemutatók építhetők.</p>				
<p>Hipotézisek:</p> <p>Feltételezem, hogy a termelésben dolgozók több mint 25%-ánál a normál érték feletti pulzusszám, vérnyomásérték; 20%-ánál magasabb koleszterin- szint, 5%-ánál magasabb vércukorszint mérhető.</p> <p>Feltételezem, hogy az egyoldalú terhelés miatt az elmúlt egy évben a termelésben dolgozók körében a nyak-, hát-, derékfájás és az izomerő-, nyújthatóság pozitivitása a vizsgáltak több mint 25 %-ánál áll fenn.</p> <p>Feltételezem, hogy a szalag mellett dolgozók körében a diagnosztizált krónikus betegségek, illetve az elmúlt egy évben felhasznált táppénzes napok száma szignifikánsan magasabb, mint az adminisztrációs munkakörben, illetve a termelésben egyéb területen dolgozók körében.</p>				
<p>Kutatás típusa: keresztmetszeti, kvantitatív</p>				
<p>A vizsgálni kívánt célcsoport: Egy multinacionális cég termelésben dolgozó, illetve adminisztrációval foglalkozó alkalmazottak. (N=120 fő)</p>				
<p>Beválasztási kritériumok: cégnél dolgozó alkalmazott beleegyeznek a vizsgálatba</p>		<p>Kizárási kritériumok: nem alkalmazottak nem egyeznek bele a vizsgálatba</p>		
<p>amennyiben véletlenszerű mintaválasztás: -----</p>		<p>amennyiben nem véletlenszerű mintaválasztás kvótás mintavétel a nem, a kor, a munkában eltöltött évek száma alapján</p>		
<p>Az esetscsoport- és ha van a kontrollcsoport – elemszáma: 100 fő termelésben dolgozó alkalmazott 20 fő adminisztrációt végző alkalmazott</p>				
<p>A vizsgálat helye: Egy Nyugat – Dunántúli régióban működő, profitorientált multinacionális cég</p>				
<p>A vizsgálat ideje: 2012.07.-2013.01.</p>				
<p>Adatgyűjtési módszer, eszközök (kérdéscsoport ismertetése, forrás megnevezése):</p>				
<p>Standardizált kérdőív: Egészségi állapotra vonatkozó</p>	<p>Saját szerkesztésű kérdőív: Demográfiai adatok</p>	<p>Strukturált interjú: -</p>	<p>Dokumentum-elemzés: üzemorvosi adatbázis, krónikus betegségek, táppénzes napok száma</p>	<p>Egyéb: Belgyógyászati: vérnyomás, vércukor, pulzus, oxigén-szaturáció, koleszterin szint, haskörfogat mérése, Fizikális: nagyzületek mozgásterjedelem vizsgálata, izomerő,- nyújthatóság mérése:</p>

Forrás megnevezése: Országos Lakossági Egészségfelmérés 2003 önkitöltős kérdőíve	
Függő változó: pulzus-szám, vérnyomás, vércukor, koleszterin, nyak és hát fájdalom, izomerő,- nyújthatóság derék és hátproblémák, mozgásszegény életmód, káros szenvedély, stressz szint, egészségi állapot	Független változó: termelésben való munka munkakör bejárás időtartama
Konkrét statisztikai módszer/próba megnevezése: leíró statisztika: abszolút- és relatív gyakoriság, átlag, szórás, medián, módusz, matematikai statisztika: korreláció, egy-mintás és két-mintás t-próba, egyutas variancia-analízis (ANOVA), Chi ² -próba Az eredményeket p<0,05 mellett tekintem szignifikánsnak.	
Az alkalmazott statisztikai szoftver megnevezése: PASW Statistics 18.0 programcsomag	
Várható eredmények: Várhatóan a vizsgálatnak köszönhetően az eddig nem diagnosztizált belgyógyászati és mozgás-szervrendszeri problémák feltárásra kerülnek, így a kezelésük is megindítható. Feltárásra kerül a munkahelyi stressz és a vizsgálatra épített felvilágosító és egészségnevelő programok hosszú távon csökkentik ezek előfordulását és növeli a dolgozók egészségtudatosságát. A dolgozók körében az általános ízületvédelem használata nőni fog.	

Forrás: Bajsz 2013

1/2. táblázat Egy releváns szakirodalom feldolgozása (minta)

Publikáció adatai /hivatkozás (szerző, megjelenés éve, cím, folyóirat neve, évfolyam, kötetszám, a cikk oldalszáma (től-ig): Móczár CS., Borda F., Faragó K., Borgulya G., Brauniczer F., Vörös V. (2007) Egészséges életmód hatása túlsúlyos és elhízott betegeken, Orvosi Hetilap 148(2):65–69.
Probléma és célkitűzés A túlsúlyos és elhízott betegek kezelésében alapvető szempont az egészséges életmód elsajátítása és beépülése a beteg mindennapjaiba. A kecskeméti Irányított Betegellátási Modellkísérletben indított prevenció program célja a túlsúlyos és elhízott betegek cardiovascularis kockázati tényezőinek csökkentése, illetve a szív- és érrendszeri megbetegedések megelőzése volt.
Kutatás típusa: kvantitatív, prospektív
Célcsoport, beválasztási- és kizárási kritérium, mintavétel: A program keretében 2489 túlsúlyos és elhízott 18 éven felüli egyént szűrtek ki és vettek gondozásba. A páciensek bevonása a BMI alapján történt. A háziorvosi praxisok rendelési idejében megjelent 25 kg/m ² feletti BMI-értékű betegeket vonták be. A betegek bevonása és a kizárási kritériumok felmérése (a testsúlynövekedés másodlagos okainak kizárása) rendelői vizsgálat és a rendelkezésre álló betegdokumentáció alapján történt. A bevonás a prevenció program ismertetése és a beleegyező nyilatkozat aláírását követően történt.
Vizsgálat helye és ideje: Kecskemét és vonzásköre. 2001. április 1. és 2002. március 31. A nyomon követés két éven át tartott, a vizsgálat 2004. március 31-én zárult.
Adatgyűjtési módszer: A bevonást követően egy adatlap kitöltésére került sor. Az adatlap az alábbi adatokat tartalmazta: betegazonosító, a beteg életkora, a beteg neme, foglalkozásának jellege, rendszeres testmozgás, a dohányzási szokások (igen/nem kategóriák szerint), alkoholfogyasztás, táplálkozás jellemzői. A vizsgálat forrásdokumentumaként a betegkarton szolgált, amelyből rögzítették a BMI-t, derékkörfogatot, vérnyomást és pulzust (nyugalmi és 15 guggolás utáni terheléses), a laboratóriumi vizsgálatok eredményeit: éhomi szérumkoleszterin, HDL-koleszterin, triglicerid, vércukor.
Statisztikai elemzés: A betegek BMI-idő grafikonjaira egyénenként lineáris regresszióval egyeneseket illesztettek. A BMI változását két többváltozós statisztikai módszerrel is vizsgálták: a) döntési fa módszerével és b) kovarianciaanalízis segítségével. A BMI általános trendjének vizsgálatát Wilcoxon próbával elemezték.
Eredmények, következtetés, javaslat: Kismértékű, szignifikáns csökkenést észleltek a BMI-ben (átlagos csökkenés: 0,5605; $p < 0,001$), a koleszterinszintben (átlagos csökkenés: 0,23; $p < 0,001$). A többi anyagcsere-paraméterben is statisztikailag jelentős csökkenést észleltek: éhomi vércukor 0,15 mmol/l ($p < 0,001$), ill. 0,19 mmol/l ($p < 0,03$), triglicerid: 0,18 ($p < 0,001$), ill. 0,08 mmol/l (nem szignifikáns), a HDL-koleszterin szintekben lényeges változást nem észleltek. Az első év végére a nyugalmi vérnyomás 5,9 Hgmm-rel ($p < 0,001$), majd a második év végére további 0,11 Hgmm-rel (nem szignifikáns) csökkent. Ez alapján a prevenció programot sikeresnek, hatásosnak tekintjük. Összességében a betegek cardiovascularis veszélyeztetettsége csökkent a két év alatt.

Forrás: Móczár et al 2007 alapján,

saját szerkesztés

1/3. táblázat Szakdolgozat készítés ütemterve

A kutatás folyamata	Tartam	Időpont
Témaválasztás, konzulens felkérése, konzultáció a konzulenssel, irodalomkutatás, kutatási modell elkészítése, kutatási terv készítés.	3 és fél hónap	2014. szeptember 01- december 15.
Kutatási terv leadása.	1 hét	2014. december 10.-15. között
Tanszék nyilatkozata választott téma elfogadásáról.	1 hónap	december 16 - 2015. január 30.
Irodalomkutatás folytatása.	2 hónap	2015. február-március
Az adatgyűjtési módszer pontosítása, véglegesítése.	1 hónap	2015. április
Adatgyűjtés. Adatbevitel.	4 hónap	2015. május-augusztus
Az adatok statisztikai feldolgozása.	2 hónap	2015. szeptember-október
A szakdolgozat szakirodalmi részének a megírása.	2 hét	2016. január
Az eredmények és a következtetés fejezet megírása.	1 hónap	2016. február-március
Konzulensi bírálat és javítás	1 hét	2016. március közepe
Dokumentációs papírok kitöltése	2 nap	2016. március vége
A szakdolgozat köttetése, CD-re való kiírás	3 nap	2016. április eleje
A szakdolgozat leadásának határideje		2016. április eleje
Szakdolgozat védésre való felkészülés, prezentáció elkészítése	2 hét	2016. június
Szakdolgozat védelem		2016. június második hete

Forrás: Majoros 1997 alapján,
saját szerkesztés

1.8. Konceptualizálás és operacionalizálás

A kutató a probléma megfogalmazása, a kutatás kérdések feltevése és a hipotézisek felállítása során különböző fogalmakat használ. Ezek a fogalmak származhatnak szakirodalmakból, kutatási eredményekből egyaránt. A fogalmak közül vannak olyanok, ami mindenki számára egyértelmű, nem indokolt annak magyarázata, ilyen például a műtét során használt varróanyag. Vannak olyan összetett, nem mindenki számára ugyanazt a jelentést hordozó elvont fogalmak, mint például stressz, egészség, életminőség, depresszió, kiégés, szorongás, amelyek értelmezése elengedhetetlen.

A **konceptualizálás** az a folyamat, amikor a kutató vizsgálat során használt elvont fogalmakat és a felhasznált változókat pontosan és egyértelműen definiálja. Emellett lényeges a dimenziók meghatározása és az egyes dimenziókhöz tartozó különféle indikátorok kidolgozása. A dimenzió nem más, mint valamely fogalom meghatározott aspektusa, nézőpontja, melyet további, ún. fő- majd aldimenziókra bonthatóak. Példaként Salavecz Gyöngyvér et al munkáját emelnénk ki, ahol a rövidített Erőfeszítés-Jutalom Egyensúlytalanság Kérdőív magyar verziójának validálását végezte el. A vizsgálat középpontjában a munkahelyi stressz mérése áll. Siegrist munkahelyi stressz modellje szerint, ha a munkahelyen a kifejtett erőfeszítések és az ezért kapott jutalmak aránya nem egyenes arányos, az ebből adódó feszültség hatására egészségromlás következik be. A társadalmi szabályoknak megfelelő kölcsönösséget elvárva a dolgozó a munkában megnyilvánuló erőfeszítéseit, munkájáért elismerésben, jutalomban részesül. Amennyiben ez nem így történik, és a munkavállalónak magas erőfeszítés mellett alacsony jutalma, ez az egyensúlytalanság testi (pl. szív- és érrendszeri) és lelki (pl. kiégés, depresszió) betegségek kialakulásához vezethet (Siegrist 1997). Ez a modell nem csak a munkahelyi környezethez köthető stresszforrásokat veszi figyelembe, hanem az egyéni jellemzőket is (Salavecz et al 2006). A munkahelyi stressz három fő dimenziója: a munkahelyi erőfeszítések (1), a munkahelyi jutalmak (2) és a túlvállalás (3). E három dimenzió meghatározása önmagában is összetett, ezért további aldimenziók meghatározására kerül sor. A *munkahelyi erőfeszítés* magába foglalja a munka során jelentkező időzavart, a megszakításokat, a zavaró tényezőket és a növekvő erőfeszítést. A *jutalom* dimenzió további alegységei az anyagi juttatások, a vezetői és közvetlen munkatársak elismerései, biztonság és karrierlehetőségek. A *túlvállalás* dimenzió tovább bontható a munkából kivonódás képessége vagy a túlterheltség aldimenzióra. Végezetül eljutunk az egyértelműen mérhető indikátorokhoz, melyekkel közvetlenül mérhető példa: a munkából kivonódás képessége az alábbi kérdések, felmérő skálák, indexek segítségével: „amint hazaér, könnyen ellazul, és félreteszi a munkahelyi

gondokat”. Az indikátorok olyan mutatók, amelyek az adott fogalom meglétét vagy hiányát jelzi a kutató számára. Az indikátorokra jellemző a felcserélhetőség elve, vagyis az indikátorok helyettesíthetik egymást és összefüggésüket tekintve különböző indikátorok segítségével is mérhető ugyanaz a fogalom. Miután az elvont fogalmak pontos definiálása megtörtént a következő lépésben a konkrét mérési eljárásokat, lépéseket kell meghatározni. Az **operacionalizálás** nem más, mint azon konkrét lépések meghatározása, melynek segítségével a kutatás során megfogalmazott elvont fogalmak empirikusan vizsgálhatóvá válnak, vagyis megfogalmazásra kerülnek mindazok a kérdések, amit a kutatás során feltesznek a különböző adatgyűjtési eszközök (kérdőív, interjú, dokumentumelemzés) segítségével.

Az operacionalizálás során az alkalmazott *adatgyűjtési módszer kiválasztása* mellett ügyelni kell a mérési terjedelem és a mérési szint meghatározására is. A *mérési terjedelem* jelentése, hogy a vizsgálni kívánt érték melyik tartománya érdekes a kutatás szemszögéből. Például egy elszegényedett társadalmi réteg jövedelmi helyzetét vizsgálva felesleges a havi élelmiszerre fordított összeghatárt 200.000.-Ft-ba meghatározni. A **mérési szintek** pontos meghatározása elengedhetetlen, hiszen mind az adatgyűjtés, mind a statisztikai módszerek meghatározása skálafüggő. A mérési szintek a 4. fejezetben kerülnek részletes bemutatásra. A konceptualizálás és operacionalizálás folyamatában már jól látható, hogy meghatározásra kerülnek a tárgyak, jelenségek, tulajdonságokra vonatkozó attribútumok és változók (1/4. táblázat). Például az iskolai végzettség, mint változó attribútumai lehetnek: általános iskola, középiskolai érettségi, diploma; a hajszín, mint változó attribútumai a fekete, barna, szőke és vörös; vagy a vércukorszint változó egyénenként eltérő értéket mutat (Pakai – Kívés 2013, Falus 2004, Héra – Ligeti 2006).

1/4. táblázat Egészségtudományi karon egy BSc hallgatóra jellemző változók és azok attribútumai

változó	attribútum
életkor	19 éves 20 éves 21 éves 22 éves
nem	férfi nő
képzési központ	Pécs Kaposvár Szombathely Zalaegerszeg
szakirány	ápoló védőnő gyógytornász mentőtiszt dietetikus szülésznő népegészségügyi ellenőr egészségfejlesztés rekreációs szervező diagnosztikai képalkotó analitika orvosdiagnosztikai laboratóriumi analitika
évfolyam	I. évfolyam II. évfolyam III. évfolyam IV. évfolyam
tagozat	nappali levelező

Forrás: saját szerkesztés

A méréssel szemben három alapvető követelményt állítanak fel: érvényesség (validity), megbízhatóság (reliability) és objektivitás. **Érvényesség** magába foglalja, hogy a kutatás során kiválasztott adatgyűjtési eszköz valóban azt a dolgot mérje, amire a kutató vállalkozott.

Különböző fajtái lehetnek:

- Tartalmi érvényesség (content validity): azt jelenti, hogy a fogalom minden egyes eleme lefedésre kerül.
- Konstruktív vagy fogalmi érvényesség (construct validity) azt mutatja, hogy a felhasznált mérőeszköz megfelel az elvárásnak.
- Egyeztetésen alapuló érvényesség (current validity) azt szemlélteti, hogy a kutatás során alkalmazott új mérőeszköz eredménye mennyire egyezik a korábbi kutatások során igazolt mérőeszköz eredményével.

- Előrejelző, prognosztikus érvényesség (predictív validity) azt fejezi ki, hogy az aktuális mérés során kapott eredmények mennyire felelnek meg egy későbbi mérés eredményének.

A **megbízhatóság** abban fejeződik ki, hogy a megismételt eljárások során ugyanazzal az eszközzel, az eredetivel egyező vagy kevésbé eltérő eredmény születik. Az **objektivitás** azt jelenti, hogy a vizsgálatot végző személy mennyire tudja tárgyilagosan, függetlenül végrehajtani az adott vizsgálatot, vagyis a mérés eredménye nem függhet mástól, csak a kérdéses dologtól (pl. hallgató teljesítménye) (Falus 2004).

1.9. A kutatás típusai

Minden tudományos kutatás megkezdése előtt érdemes tisztázni azt a kérdést is, hogy milyen kutatási eljárás alkalmas a kitűzött cél elérése érdekében. A döntést - az egyes kutatások között tapasztalható különbség miatt a személyi, tárgyi és anyagi feltételek is befolyásolhatják. A kutatási típusok besorolása különböző megközelítési módok szerint történhet, melyek közül a leggyakrabban használt fogalmak kerülnek ismertetésre.

A kutatások szintje alapján alap-, alkalmazott és fejlesztő kutatás különíthető el. Az **alapkutatás** empirikus megfigyelésen alapuló kísérleti, tapasztalati, rendszerező vagy elméleti munka, amely a világ jelenségeinek megismerésével és a törvényszerűségeinek feltárásával foglalkozik. Elsődleges céljai közé tartozik az új információk/ismeretek szerzése az elmélet finomításának céljából. A tiszta alapkutatás nem tartalmaz konkrét gyakorlati alkalmazásra és felhasználásra vonatkozó célkitűzéseket, míg a célzott alapkutatás lényege, hogy alapul szolgál a jövőben felmerülő problémák megoldására. Alapkutatást inkább felsőoktatási intézmények, egyetemek és kutató intézetek végeznek. A teljesség igénye nélkül egy példa kerül kiemelésre a gyógyszerésztudományok területéről. Az Észak-Umbriai Egyetem kutatói bizonyos halolajtartalmú étrendkiegészítő hatását vizsgálták. A kettős-vak, placebo kontrollós vizsgálatban 22 fiatal felnőtt (átlag életkor=21,96 év) vett részt. Az első csoport dokozahexaénsavban gazdag, a második csoport eikozapentaénsavban (EPA) gazdag halolajat, a harmadik (placebo) csoport olivaolajat kapott. Az eredményekből kiderült, hogy a különböző omega3 tartalmú halolaj rendszeres szedése nem befolyásolta jelentősen a mentális funkciókat, de csökkentette a szellemi fáradtságot és javította a reakcióidőt (Jackson et al 2012). Az **alkalmazott kutatás** az alapkutatások elméleti eredményeire támaszkodva a gyakorlat során felmerülő problémák megoldására koncentrál. A teljesség igénye nélkül említhetőek a piackutatások, ahol a szolgáltató az eredmények alapján tervezi meg arculatát vagy változtat az üzleti stratégiáján, vagy a gyógyszeriparban a hatóanyagokkal,

gyógyszerkészítményekkel végzett különböző kísérletek. A **fejlesztő kutatás** olyan kutatótevékenység, melynek célja a különböző elméletek, módszerek, eszközök, termékek továbbfejlesztése, korszerűsítése. Példaként emelhető ki, hogy a GVOP által konzorciumi pályázat keretében a Budapesti Műszaki és Gazdaságtudományi Egyetem három tanszéke egy egészségi állapotot, szakképzettség nélkül otthon monitorozó készülék kifejlesztésére vállalkozott, a PTE ETK Zalaegerszegi Képzési Központ munkacsapata a készülék fejlesztésével kapcsolatos tájékoztatást és igényfelmérést vállalta magára (Jobbágy et al 2008, Karamánné et al 2006).

A kutatás során az idő is meghatározott szerepet játszik. Egy vizsgálat történhet egy adott időpontban vagy hosszabb időintervallumban. A **keresztmetszeti vizsgálatra** jellemző, hogy az információgyűjtés a vizsgálni kívánt mintán egyszeri alkalommal, konkrét időpontban történik, és nem kerül sor az ismétlésre. Ezzel a vizsgálati típus inkább az egyén aktuális állapotára, véleményére vagy adott helyzetére fókuszál, bár előfordulhat, hogy olyan kérdés is felmerül, ami a múltra (néhány hónappal, évvel ezelőtti állapotra) vonatkozik, melyeknek a megbízhatósága azonban megkérdőjelezhető. Vajda Réka et al keresztmetszeti kutatásában a méhnyakrák-szűréssel kapcsolatos ismereteket vizsgálta 9-14 éves leánygyermeket nevelő nők körében (Vajda et al 2014). A **longitudinális (követéses)** vizsgálat azonban hosszabb időn keresztül megy végbe. A vizsgálatot a kiválasztott mintán a kutatást végző szakemberek meghatározott időközönként megismétlik. Példaként emelhető ki az a vizsgálat, ahol a kutatók 1986-2006 között, négyévenként az egészségügyi dolgozók életmódjában bekövetkező változásokat vizsgálták. A résztvevők a kutatás megkezdése előtt normál testsúlyúak voltak és krónikus megbetegedésben nem szenvedtek. Az eredmények alapján a testsúly negyedévenként átlagosan 1,5 kg-mal emelkedett, húsz év alatt a 7,6 kg volt (Mozaffarian et al 2011). A longitudinális vizsgálat további három típusa különíthető el. A kutatás során kiválasztott, ugyanannak az alapsokaságban idővel bekövetkező változásaival foglalkozik a **trendvizsgálat**. Kisebb, populációt (kohorsz) vizsgál a **kohorszvizsgálat**, ahol a mintán bekövetkező változások időbeli követése történik. A **panelvizsgálat során** ugyanaz a minta, ugyanazok az emberek megkérdezése, vizsgálata történik több alkalommal. A longitudinális vizsgálatok a keresztmetszeti vizsgálatokkal szemben több és megbízhatóbb információt nyújtanak a vizsgált mintáról a kutatást végző szakemberek számára, ugyanakkor a kutatás költségei is jóval magasabbak. Az időbeliség szempontjából a prospektív és retrospektív vizsgálatról kell szót ejteni. **Prospektív** vizsgálatnak tekinthető, ha a mérési és a megfigyelési feladatok elvégzése a vizsgálat során történik, tehát az adatok a vizsgálat során keletkeznek. **Retrospektív vizsgálat** során az analízis a vizsgálat időpontjához képest korábbi

eseményekkel, illetve korábban keletkezett adatokkal történik. Érdeemes szem előtt tartani, hogy a kísérlet mindig prospektív típusú, míg a megfigyeléses vizsgálatoknál mindkét típus előfordulhat. Téves, ha a kohorsz vizsgálatokat a prospektív vizsgálatokkal azonosítják, mert a kohorsz lehet egyúttal retrospektív is (levéltári, irattári dokumentumok feltárása pl. egy településen élő lakosság vizsgálata során). Az eset-kontroll (amennyiben pl. egy adott betegségben szenvedő és nem szenvedők csoportját tekintjük) minden esetben retrospektív vizsgálatnak kell tekinteni (Reiczigel 2005). Egy kutatás tervezése során alapvetően - a kutatási probléma jellege alapján - három kutatási stratégia áll a kutatást végző szakember rendelkezésére. A döntés számottevő, mert jelentősen meghatározza a kutatás egészének további menetét. A minőségi jellemzésre szolgáló, nem számszerű eredményeket szolgáltató **kvalitatív kutatás** célja a mélyebb, árnyaltabb ismeretek megszerzése, a nem nyilvánvaló ok-okozati összefüggések feltárására és a probléma jobb megértése. A feltáró jellegű, tipikusan a probléma meghatározás során megfogalmazott „miért” és a „hogyan” kérdésre keresi a választ. Az adatgyűjtés egy kisméretű, behatárolt területtel rendelkező, nem reprezentatív mintán történik. A kutatás alkalmával a kutató bizalmi kapcsolat kialakítására is képes a kutatásba bevont személlyel, így olyan lényeges információkat gyűjthet, ami egy kvantitatív kutatás nem biztos, hogy lehetővé tenné. Hátránya, hogy objektív, számszerűsíthető adatok gyűjtésére nem használható, megbízhatósága alacsony és magas a szubjektivitás esélye. A leggyakrabban alkalmazott kvalitatív kutatási módszerek: a megfigyelés, a kísérlet, a mélyinterjú, a szakértői interjú, a fókuszcsoport és az esettanulmány. A kutatások másik csoportját képviselő **kvantitatív kutatás** a mennyiségileg mérhető információk gyűjtésére alkalmazzák. Ennél a típusnál a kutatásba bevont alapsokaságról strukturált kérdőív, strukturált interjú vagy strukturált megfigyelés során kapnak - a kutatási probléma megválaszolására szolgáló - választ. Az adatgyűjtés során kapott nagy mennyiségű, megbízható, számszerűsített adatokat különböző leíró és matematikai statisztikai elemzéssel lehet feldolgozni és magyarázni. Egy komplex, átfogó kutatás során azonban - a lényeges különbségek ellenére is - egymást kiegészítő módszerekként is gyakran alkalmazzák a **kvalitatív és kvantitatív vizsgálatot**, mint Szabó et al (2009) Zala Megyei Kórház Pszichiátriai Osztályának Addiktológiai részlegén alkoholfüggő betegeket vontak be a vizsgálatba. A vizsgálat során arra kérték a betegeket, hogy saját kézzel írjanak magukról egy önéletrajzot. Elemzésre az ATLASZ.ti 5.00 tartalomelemző programot alkalmazták, melynek segítségével a szociális szavak gyakoriságára fókuszáltak. A kutató munkát végző szakemberek először a kiválasztott program segítségével kigyűjtötték az önéletrajzban található összes szót, majd kiválogatták a társas kapcsolatokra utaló szavakat, és a tartalmuk alapján kategóriákat

állítottak fel, ezt követően az azokhoz tartozó szociális szavak megjelenési gyakoriságát vizsgálták.

1.10. A tudományos kutatás etikai kérdése

A kutatást végző szakembereket komoly felelősség terheli a kutató munka folyamán. A kutatónak nemcsak szakmai tudással, kellő jártassággal kell rendelkeznie az adott témában, de általános etikai elvárásoknak is meg kell felelnie.

A vezető kutató felelős a kutatás etikai vonatkozásáért. A vizsgálatba bevont egyének személyiségi jogait tiszteletben kell tartani, biztosítani kell anonimitásukat. A kutató köteles őszinte, részletes tájékoztatást nyújtani a vizsgálat céljáról, a várható időtartamáról, a vizsgálat eredményeiből származó előnyökről és a fellépő kockázatoktól, ami egyben a bizalmas kapcsolat kialakítását is elősegíti. Egy vizsgálat csak a résztvevő beleegyezése után kezdődhet el, amit bármikor visszavonhat. A kutatásban résztvevők kockázatát minimálisra kell csökkenteni és a biztonságos környezet megteremtése is alapvető követelmény. Nem utolsósorban a résztvevők idejét ugyancsak tiszteletben kell tartani.

A kutatóval szembeni etikai elvárások során megfogalmazódik, hogy a munkáját szakmai presztízisének megfelelő, magas színvonalúan, precízen, pontosan és megbízhatóan kell végeznie. Az etikai vétségek elkerülése céljából ügyelni kell az adatvédelem, a forrásfelhasználás, a kutatási eredmények etikai szabályainak betartására is. Etikátlan, ha az eredményeket torzítják, meghamisítják vagy a terhelő adatokat elhallgatják. Mivel egy kutató munka során más kutató munkáját is felhasználhatja, a tudományos tisztaság megóvása és a szerzői jogok megsértésének elkerülése céljából nagy hangsúlyt kell fektetni a hivatkozások korrekt alkalmazására (Falus 2004, Cserné 1999, Majoros 2004, Dempsey 1999).

1.10.1. Embereken végzett kutatások

Az embereken valószínűleg ősidők óta végeztek különböző kísérleti vizsgálatokat, ami nem mindig szolgálta a kísérleti alanyok jóllétét, azok beleegyezés nélkül és az akaratuk ellenére történt. Ezt bizonyítja Hérodotosz feljegyzése is, amit a Héphaisztosz papjainak elbeszélése alapján írt le i.e. VII. században. Pszammetikosz fáraó rendelkezésére két újszülöttet bíztak egy pástorra, hogy a nyája között, emberi kapcsolatoktól elkülönítve nevelje fel őket, úgy hogy emberi szót nem szabad előttük kiejteni. Ezzel azt szerette volna megtudni, hogy a gyermekek milyen nyelven szólalnak meg először, mert a feltételezése szerint az a nyelv egyben az emberiség ősnyelve is.

A történelmi eseményeket áttekintve embereken végzett kísérleteket leginkább rabszolgákon, a háború során foglyul ejtett alanyokon vagy olyan embereken végeztek, akiket nem tekintettek olyan embernek, mint saját magukat. A II. Világháborúban sem tartották egyesek tiszteletben a kutatásban bevont emberek alapvető jogait. Ilyen kísérletek voltak a például Auschwitz koncentrációs táborában felállított kísérleti laboratóriumában Mengele ikerkutatásai. Napjainkban azonban az emberi alanyokon végzett kísérleteket mind nemzetközi, mind nemzeti szinten nagyon szigorú szabályozás alatt áll. Például a Nürnbergi Kódex 10 pontja rögzíti az embereken végzett kísérletek alapvető feltételeit, melyben biztosítják a vizsgálatban résztvevő alanyok méltóságát és biztonságát (1/5. táblázat). Kiemeli, hogy a kutatásban bevont személyeket előzetesen fel kell világosítani és beleegyezésüket kell adni a kutatások elvégzésére (Kerpel-Fronius 2008, Kovács 2007).

Orvostudományi Világtalálkozón (World Medical Association) 1964-ben fogadták el a Helsinki Deklarációt, amely irányt mutat az orvosi kutatásokban résztvevő szakemberek részére. A jelenleg 35 pontból álló orvosi kutatások etikai alapját biztosító kódexet 6-8 évente vizsgálják felül és módosítják. A deklaráció 3 fő részből áll: Bevezetés (1), Valamennyi orvosi kutatásra vonatkozó alapelvek (2) és kiegészítő elvek olyan orvosi-biológiai kutatásokhoz, amelyek az orvosi beteggondozás keretében történik (3).

Európában jelentős szerepet tölt be az orvosi és biológiai kutatások etikai szabályozásában az Európa Tanács testülete és a Steering Committee on Bioethics (CDBI) munkája. 1997-ben született meg az Oviedói Emberjogi és Orvosbiológiai Konvenció, amelyet Magyarország is aláírt. A CDBI a kötelezően alkalmazandó dokumentumai a következők: az emberi klónozás tiltása, a szervkereskedelem tiltása, az orvosbiológiai kutatások szabályozása, a genetika vizsgálatok szabályozása egészségügyi célokra. Emellett más ajánlások sorát is közzé tett, mint például a xenotranszplantáció, az elmebetegek jogainak védelme, a humán biológiai anyagokkal történő kísérletek.

Az Európai Unió direktívái sem hagyhatóak figyelmen kívül, ami a humán vizsgálatokra, gyógyszerkísérletekre, állatkísérletekre egyaránt vonatkoznak (Fésűs 2014).

1/5. táblázat Nürnbergi Kódex

- A kutatásba bevont emberi alany önkéntes beleegyezése feltétlenül szükséges.
- A kísérlet eredményeinek a társadalom javát kell szolgálnia, nem lehet ötletszerű és szükségtelen.
- Az állatkísérletek eredményei szolgáljanak alapul a kísérlet tervezésénél.
- A kísérleti alanyok számára a vizsgálat nem okozhat fizikai és lelki szenvedést és károsodásokat.
- Olyan kísérlet nem végezhető, amiben feltételezhető a halálos kimenetel vagy a maradandó károsodás.
- A viselendő kockázat mértéke soha nem haladhatja meg a megoldandó probléma humanitárius fontosságát.
- A kísérletbe bevont személyek védelme céljából a vizsgálatokat kellő körültekintéssel kell előkészíteni és biztosítani kell a szükséges eszközöket.
- A kísérletet kizárólag olyan szakemberek vezethetik, akik tudományos fokozattal rendelkeznek.
- A kísérletnek alávetett személyek a beleegyezésüket bármikor visszavonhatja.
- A kísérletet a vizsgálat bármely szakaszában be kell fejezni, ha a kísérleti alanyok épségét veszélyezteti (sérüléséhez, rokkantságához vagy halálához vezet).

Forrás: <http://www.hhs.gov/ohrp/archive/nurcode.html> nyomán saját szerkesztés

1.10.2. A hazai szabályozás

A kutatás megkezdése előtt kutatási engedély kérése szükséges. Hazánkban a megfelelő kutatásetikai bizottsághoz kell a kérelmet benyújtani. Szakmai etikai engedélyt embereken végzett orvostudományi kutatás alkalmával a 23/2002 (V.9.) EüM rendelet az emberen végzett orvostudományi kutatásokról /módosította: 31/2009. (X.20.) EüM rendelet/ szerint

- Az Egészségügyi Tudományos Tanács Tudományos és Kutatásetikai Bizottság (ETT TUKEB) vagy
- a Regionális Kutatásetikai Bizottság (REKEB) vagy
- az Intézményi Kutatásetikai Bizottság (IKEB) adhat.

A kutatási engedély beadása előtt érdemes a hazánkban hatályos törvényi szabályozásokat áttanulmányozni (1. számú melléklet), ami biztosítja az embereken végzett kísérletek során az alapvető emberi jogok érvényesülését. Nem szabad elfeledkezni arról sem, hogy tájékozódni kell a benyújtott dokumentumokkal kapcsolatos elvárásokról is. Az eljárási rendekről az Egészségügyi Tudományos Tanács honlapján a www.ett.hu lehet bővebb információhoz jutni. Felhívjuk a figyelmet arra is, hogy a beavatkozással járó és a beavatkozással nem járó kutatási tervekben foglalt kutatásról a szakmai vezetőnek beszámoló jelentést is kell írni (Alexin – Lelovics 2009).

2. IRODALOMKUTATÁS A GYAKORLATBAN. IRODALOMKUTATÁS A LEGGYAKRABBAN HASZNÁLT ADATBÁZISOK FELHASZNÁLÁSÁVAL (Karamánné Pakai Annamária, Oláh András)

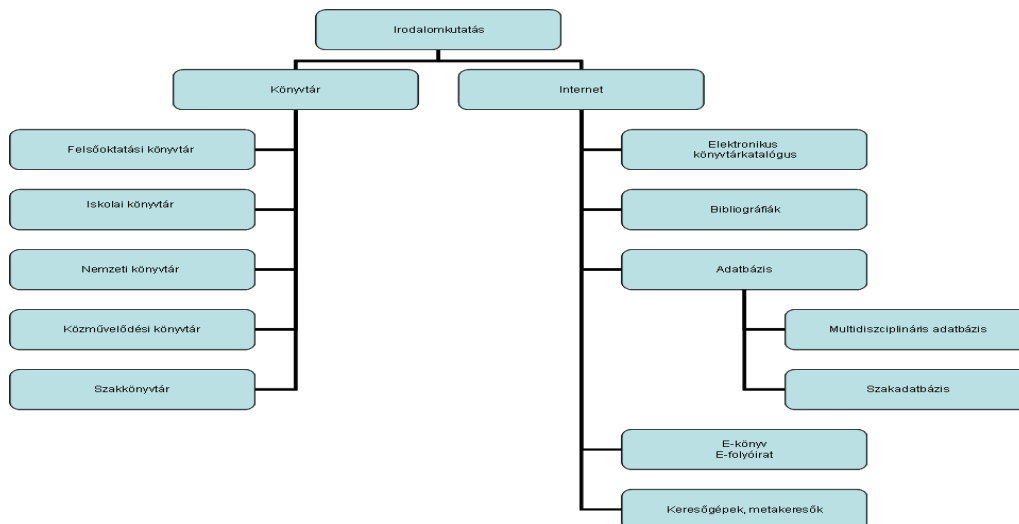
2.1. A releváns szakirodalom gyűjtése és feltárása

A kutatási téma kiválasztása és a probléma felvázolását követően elengedhetetlen a témával kapcsolatos nemzetközi és hazai releváns (releváns: a tárgyhoz tartozó, fontos, lényeges, meghatározó) szakirodalom feltárása. Napjainkban a szakirodalom feltárásához a kutatók számára széles körben állnak rendelkezésre a könyvtárak, szakkönyvtárak mellett elektronikus adatbázisok, melyeken a publikációk jelentős része teljes terjedelemben is elérhetőek. (2/1. ábra).

Naprakész, bizonyítékokon (evidence-based) alapú, tudományos értékű munka akkor készíthető, ha a kutató a választott témában kellő jártassággal rendelkezik. A kutatás előzményeinek feltárása során a kutatónak tájékozódni kell arról, hogy a kijelölt témában milyen nemzetközi és hazai publikációk jelentek meg. Kellő jártasság akkor alakítható ki egy témában, ha a kutató ismeri, hogy kik, milyen típusú kutatást végeznek, milyen problémára fókuszáltak, hol és milyen célcsoportban és mintán milyen adatgyűjtési módszert alkalmaztak (kérdőív, dokumentumelemzés, kísérlet stb.), milyen változókat mértek és milyen statisztikai módszerekkel történt a hipotézisek alátámasztása, az eredmények milyen új problémafelvetéshez vezettek, valamint milyen korlátai voltak az adott kutatásnak.

A releváns szakirodalom áttekintése segítséget nyújt a tudományos kérdések pontos megfogalmazásában, a tudományos fogalomkészlet meghatározásához/megismeréséhez, az adatgyűjtési módszer, eszközök megválasztásához, az eredmények értékeléséhez és az önálló munka tartalmi és szöveges összeállításához is.

Az irodalomkutatás nemcsak időigényes folyamat, de nem nélkülözheti a kutatómódszertan és statisztikai ismereteket sem, mivel a kutatónak képesnek kell lennie arra, hogy az adott publikáció minőségét - ha nem is minden szempontból - és a benne lévő információk, eredmények megbízhatóságát is értékelni tudja.



2/1. ábra: Irodalomkutatás kiinduló pontjai
 Forrás: saját szerkesztés

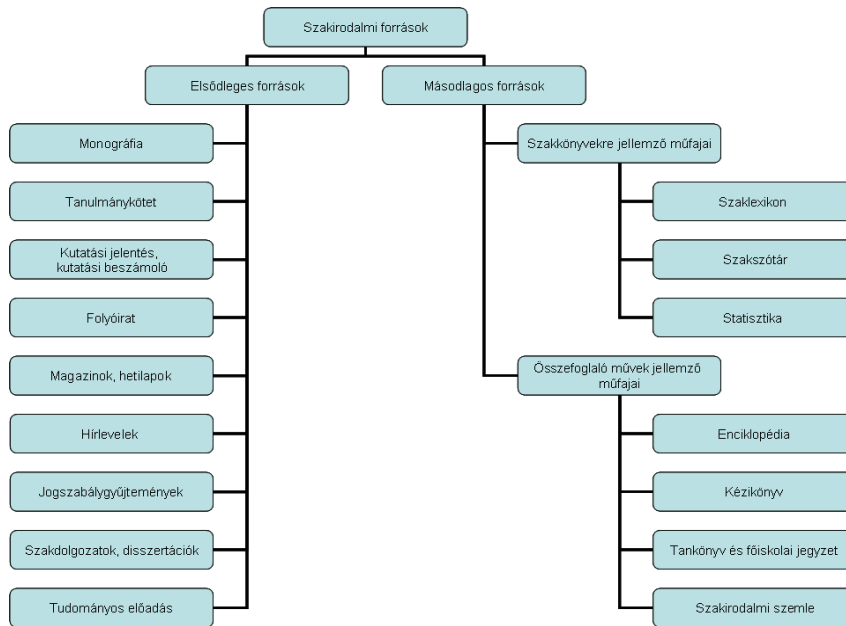
2.2. Az irodalomkutatás forrásai

Egy tudományos mű elkészítéséhez szükséges releváns források egy része intézményi könyvtárakban vagy szakkönyvtárakban lelhető fel. A feltáró munka folyamán talált források között elsődleges és másodlagos forrásokat lehet felhasználni (2/2. ábra), de nem szabad a szürke irodalom jelentőségéről sem megfeledkezni.

Az **elsődleges forrás** olyan eredeti mű, amely egy bizonyos témával vagy tudományterülettel foglalkozik és a szerző a kutatás eredményeit közvetlenül mutatja be például egy folyóirat cikkben, disszertációban vagy egy konferencia előadás keretében.

A **másodlagos forrásnak** az adott tudományterület elsődleges forrásai alapján készített összefoglalókat nevezik, amelyek átfogó képet nyújtanak az adott kutatási területről. A másodlagos forrás nemcsak bemutatja, de elemezheti, cáfolhatja, kiegészítheti vagy akár bizonyíthatja az elsődleges forrást.

A „**szürke szakirodalom**”, olyan műveket foglal magába, melyek viszonylag szűk körben kerülnek felhasználásra és megismerésre, ugyanakkor gyakran nyújt hasznos segítséget a kutató számára. Szürke szakirodalomnak tekintik a felsőoktatási intézetek, kutatóközpontok, kormányzati és egyéb szervek által készített korlátozott számban megjelent kiadványokat, amelyek nehezen hozzáférhetőek és a kereskedelmi forgalomba sem kaphatóak.



2/2. ábra Szakirodalmi források

Forrás: Oláh (szerk.) 2008 nyomán saját szerkesztés

2.3. Az internet használata az irodalomkutatás során

Az interneten való irodalomkutatás többféle forrásból indulhat ki, melyek a továbbiakban részletesen bemutatásra kerülnek.

- Elektronikus könyvtári katalógusok
- Elektronikus úton elérhető bibliográfiák
- On-line adatbázisok
- Elektronikus könyvek, elektronikus folyóiratok
- Internetes keresőrendszerek

Kiegészítő szolgáltatások lehetnek:

- hírlevelek, levelezőlisták és a szakmai szervezetek, társaságok honlapjai, mint például:
 - a Magyar Egyetemi - Főiskolai Sportszövetség hírlevele
(elérhetősége: www.mefs.hu/hirlevel),
- a Magyar Sporttudományi Társaság honlapja (elérhetősége: www.sporttudomany.hu),
- a Magyar Asztalitenisz Szövetség (elérhetősége: www.moatsz.hu)

Mielőtt megkezdődne az irodalomkutatás, fontos lépés a témával kapcsolatos magyar és idegen nyelvű kulcsszavak, tárgyszavak összegyűjtése, a témával kapcsolatos fogalmak és szinonimák átgondolása és a keresési szempontok (korcsoport, nem, nyelv, idő)

meghatározása. Nem szabad elfelejteni, hogy az adatbázisban való sikeres keresés kulcsa a kereső kifejezések pontos megfogalmazása.

2.4. Elektronikus könyvtárak

A könyvtárak állományában megtalálható könyvtári katalógusok az adott könyvtár dokumentumainak több szempont alapján történő keresésére szolgálnak. Napjainkban a könyvtárak jelentős részében a cédulaalapú katalógusok szerepét átvette a technikailag modernebb elektronikus katalógusok. Előnye, hogy nemcsak az adott könyvtárban, de otthoni számítógépekről is elérhetőek a könyvtár állományában megtalálható irodalmak adatai. Eszerint választ ad arra, hogy a kutató által keresett szakirodalom vagy a kutatás témájában elismert szakember művei megtalálhatóak-e a könyvtárban vagy a kiválasztott dokumentum milyen státusszal rendelkezik (kölcsonözhető, vagy csak olvasótermi használatban olvasható).

Online Public Access Catalogue (OPAC)

Az OPAC interneten keresztül tértől és időtől függetlenül mindenki számára hozzáférhető, ingyenes szolgáltatást nyújtó, nyilvános könyvtári katalógus. A kutatók számára széleskörű keresési lehetőséget biztosít az adatbázisban, így a keresés történhet: szerző (szerkesztő, egyéb alkotó), cím, vagy annak szavai, tárgyszó (a dokumentum tartalmára utaló, téma szerinti keresés), kiadó, ISBN, ISSN szám (könyv ill. folyóirat egyedi azonosító száma), ETO (Egyetemes Tizedes Osztályozás, a PTE könyvtár által nem használt), lelőhely (könyvtár, intézetek, klinikák kódjaival a könyvtáros segítségével), raktári jelzet-helyrajzi szám (könyvtáros segítségével), kiadás éve (adott évnél mindkét mezőbe ugyanazt az évszámot kell írni, szűrőfeltétel), dokumentumtípus (könyv, periodika, számítógépfájl, vizuális, stb. szűrőfeltétel), és nyelv (szűrőfeltétel) alapján.

Az illusztrált példákat a Pécsi Tudományegyetem Egyetemi Könyvtárának elektronikus katalógusának felhasználásával történik. Elérhetősége: www.lib.pte.hu

Keresési mód

Az online katalógusok a legtöbb esetben három keresési módot ajánlanak fel a kutató számára: 1. egyszerű keresés (search), 2. összetett keresés (advanced search). 3. böngészés (browse). Az **egyszerű keresés** történhet szerző, cím, tárgyszó, kiadó, könyv illetve folyóirat azonosító száma (ISBN, ISSN) alapján. Az egy oldalon megjelenő találatok száma 5-100-ig állítható be. (2/3. ábra). **Csonkolt keresés** alkalmazható, amennyiben nem ismert a keresett kifejezés pontos tartalma vagy a kereső kifejezés csak egy részletét kívánjuk beírni. A

csonkolás a karaktersorozat elejének vagy végének a levágásával történhet, de a szó közben is alkalmazható helyettesítésre. A leggyakrabban alkalmazott csonkolási jelek pl.: dollár (\$), hashmark (#), a csillag (*), a százalékjel (%), a kérdőjel (?), azonban a betűk helyettesítése alsó kötőjellel (_) alkalmazásával is megoldható.

Példa: az egészség% csonkolt formában történő keresése során az egészséggel kapcsolat valamennyi találatot, mint pl.: egészséges, egészségügy, egészségkárosodás, egészségpszichológia stb. megkapjuk.

Összetett keresés lehetőséget kínál több egyszerű keresés kombinálására. Ebben az esetben ún. Boole-operátorokat (vagy logikai operátorokat) használnak abból a célból, hogy az adatbázisban való keresés során a különböző kereséseket összekapcsolják. A kapcsolatok létesítésére három operátort alkalmaznak **and**; **or** és **not**, melyek eltérő kapcsolatot létesítenek az egyes keresések során. (2/4. ábra)

AND (logikai ÉS kapcsolat) segítségével olyan találatokat kapunk, amelyek a kereső kifejezések mindegyikét magába foglalják.

Példa **sport and egészség**

Ha egyik keresés a **sport**, a másik az **egészség**, akkor a **sport and egészség** azokat a találatokat válogatja ki, amelyben a *sport* és az *egészség* szó **egyaránt** előfordul. Ebből az is következik, hogy a sorrend nem számít, tehát *sport and egészség* = *egészség and sport*

OR (logikai VAGY kapcsolat) olyan találati eredményekhez vezet, amelyek a legalább az egyik kereső kifejezést tartalmazzák.

Példa **sport or egészség**

A megadott kifejezések keresésekor mind a *sport*, mind az *egészség* szavak is megtalálhatóak az adatbázisban, de azok is, ahol mindkettő előfordul. (Ez a közös rész határozható meg a fenti *and* operátorral.) A sorrend ebben az esetben sem számít, tehát *sport or egészség* = *egészség or sport*.

NOT (logikai NEM, vagy kizárás) használatával a kapott találatok az elsőnek megjelölt kifejezést tartalmazzák, a másodikat azonban nem.

Példa **sport not egészség**

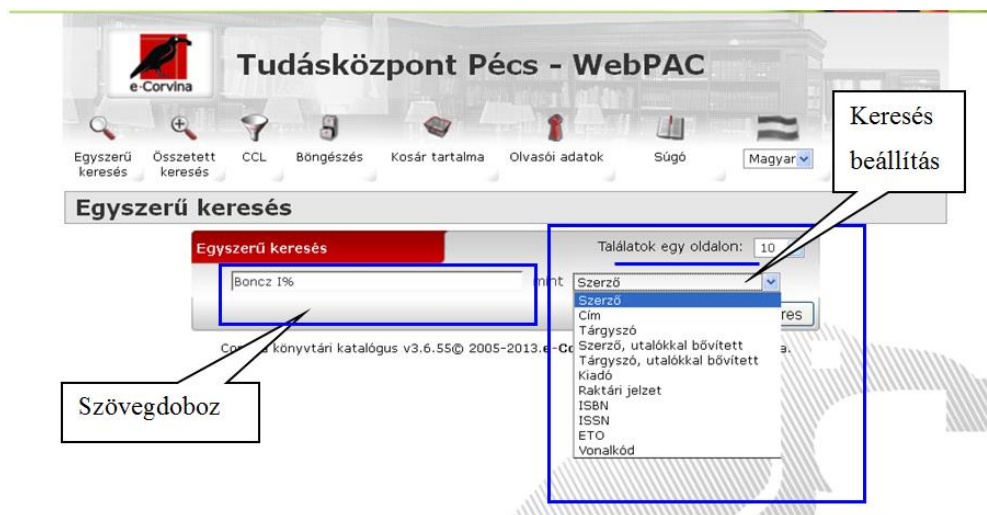
A kifejezések beírása alkalmával a sport feltételt teljesítő rekordok kereshetőek, és ezekből eltörli az egészség szót is tartalmazókat. A sorrendre azonban itt nagyon vigyázni kell: a sport not egészség mást jelentéssel bír, mint az egészség not sport.

A keresés során a kiválasztott mezőbe be kell írni a kereső kifejezést, majd a legördülő listából hozzá kell rendelni a megfelelő típust, de emellett beállítható a leőhely, a kiadás dátuma, a keresett dokumentum nyelve (magyar, angol), a dokumentum típusa (bármilyen, könyv, folyóirat, cikk, film stb.). A keresést a *KERESÉS* felirattal ellátott gombbal lehet indítani.

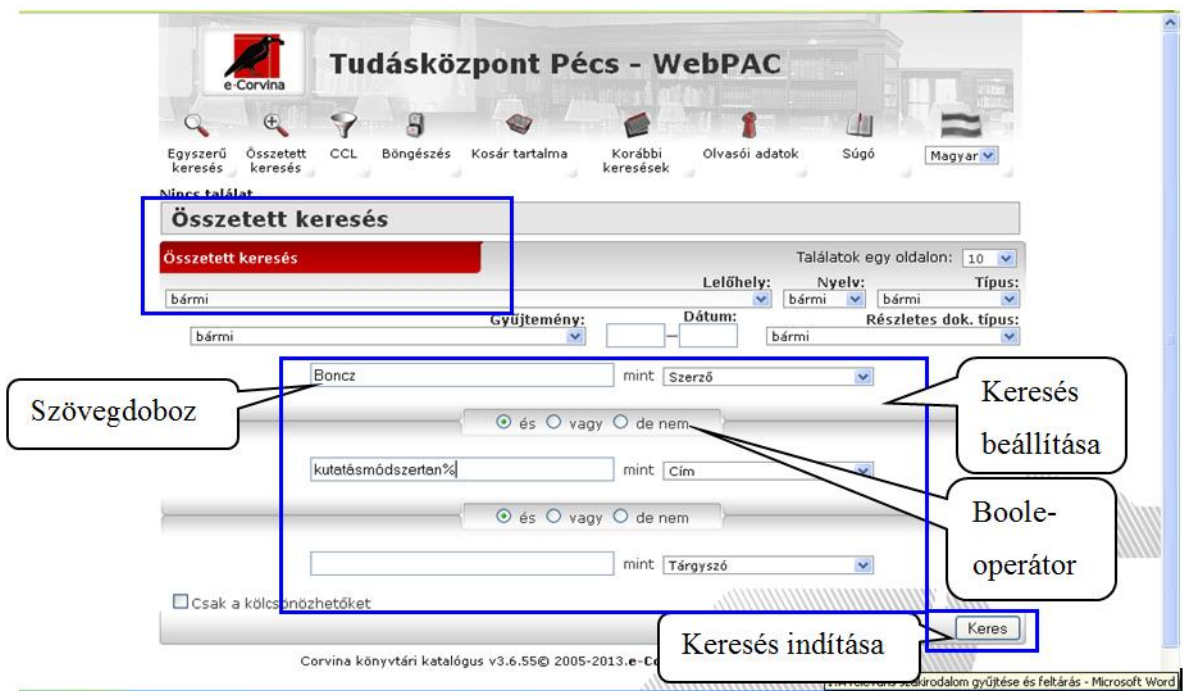
A **böngészés** lehetőséget kínálja az adatbázis egységesített besorolási adatai közötti keresésre (nevek, címek, tárgyszavak) (2/5. ábra). Olyan esetben érdemes ezt a lehetőséget választani, ha bizonytalan a keresett kifejezés írásmódjában, a szerző pontos nevében, valamint a kifejezés katalógusban alkalmazott formáját tekintve. Itt csak a szövegben lehet alkalmazni a csonkolási jeleket. Böngészés segítségével nyújt az adatbázistól a megfelelő kereső kifejezés kiválasztásához. Ezáltal nemcsak a szerzőre lehet keresni (pl.: Zala György), hanem mindenre, amiben a „Zala” kifejezés megtalálható.

A találatok megjelenítése

A megadott kifejezések alapján egy találati lista *rövid* formában jelenik meg, amely tartalmazza a dokumentum azonosítására szolgáló legfontosabb információkat, mint a szerző nevét, a mű címét, a megjelenés évét és a dokumentum típusát. A talált rekordok megjelenítéséhez több lehetőség is rendelkezésre áll. Így a találatok rendezhetőek a találatok mennyisége, részletessége (rövid/hosszú), a kiadás dátuma, a szerző vagy mű címe alapján betűrendi sorrend szerint, de történhet a rendezés csökkenő vagy növekvő sorrendben is. A találati oldalon a *RÉSZLETEK* feliratú fülre kattintva továbbá információt ad a kiválasztott mű kölcsönzés helyéről és aktuális kölcsönzési státuszáról is (2/6. és 2/7. ábra).



2/3. ábra. Egyszerű keresés a könyvtári katalógusban



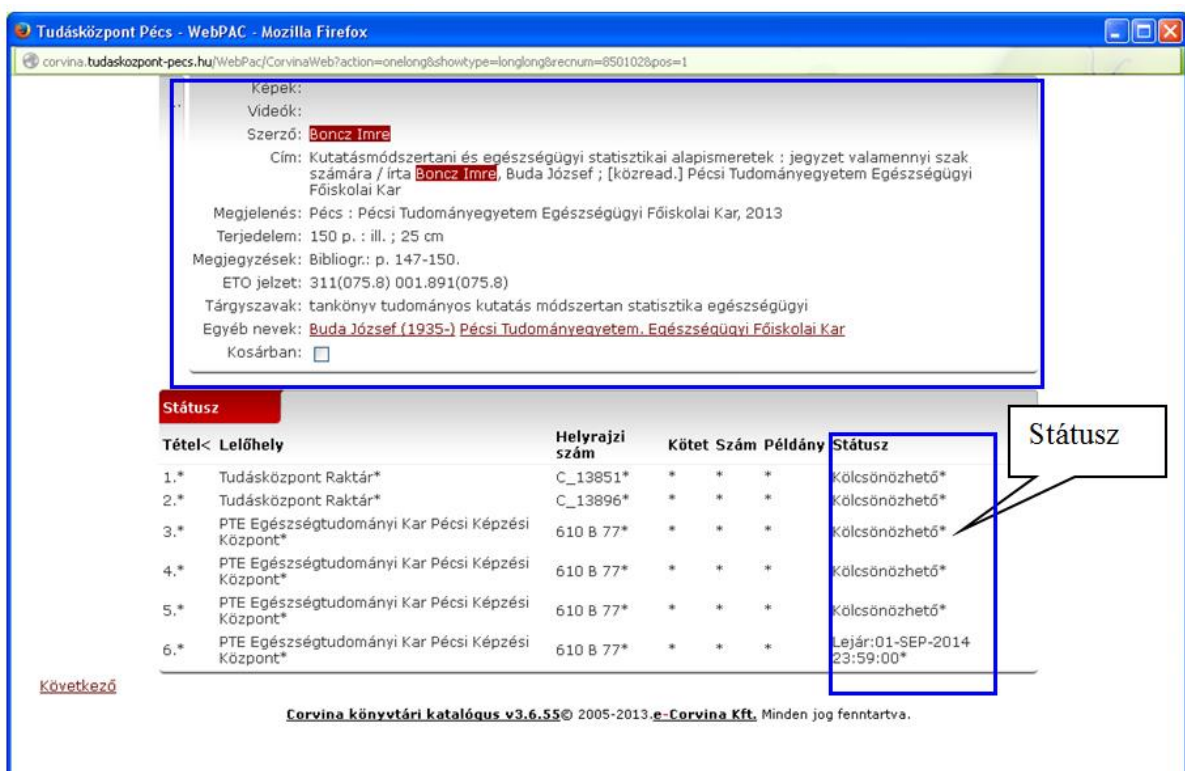
2/4. ábra. Összetett keresés a könyvtári katalógusban



2/5. ábra. Böngészés a könyvtári katalógusban



2/6. ábra. Találati lista



2/7. ábra Találati lista részletei

2.5. Elektronikus úton elérhető bibliográfiák

A könyvtári katalógusok mellett napjainkban számos elektronikus úton is elérhető rendszeresen frissített, naprakész bibliográfia, szakbibliográfia segíti a mélyreható releváns szakirodalom további feltárását. A **bibliográfia** egy olyan produktum, melyben a

dokumentumokról meghatározott szempontrendszer alapján jegyzék, lista készül. A csoportosítás lehet téma, megjelenés hely, idő, dokumentum típusa alapján, a rendezési elv pedig leggyakrabban betűrend vagy időrend szerint történik. A **szakbibliográfia** szerepe adott szakterületen megjelent szakirodalmak összegyűjtése és rendszerezése, mint például a Magyar Orvosi Bibliográfia.

Magyar Orvosi Bibliográfia (MOB)

A Gyógyszerészeti és Egészségügyi Minőség- és Szervezetfejlesztési Intézet Informatikai és Rendszerelemzési Főigazgatóság (GYEMSZI IRF) Országos Egészségtudományi Szakkönyvtár révén kerül feldolgozásra a hazánkban megjelenő orvostudományi és annak határterületeit átfogó szakfolyóiratainak cikkei és módszertani levelei. Online módon mindenki számára hozzáférhető adatbázis, mely korábban nyomtatott formában is megjelent, azonban 2006-tól évente 4 alkalommal CD-ROM-formájában kerül kiadásra. Egy-egy CD mintegy 1000 új tétel foglal magában. Az adatbázisban egyaránt kereshető könyvek, folyóiratcikkek és más dokumentumok is, mint például időszaki kiadványok, tudományos értekezések, kongresszusok és konferencia kiadványok. A csonkolás százalékjellel (%) történik.

Keresési módok

A MOB online felületén - hasonlóan a könyvtári katalógusban való keresésnél - lehetőség van egyszerű és összetett keresésre valamint böngészésre. Az *egyszerű keresés* során cím, MOB, szerzők, tárgyszavak, folyóirat és kulcsszó alapján történhet a keresés, a kutatók számára leginkább a cím és kulcsszó alapján történő keresést ajánlják, mert a találati eredmények a bibliográfiai tételek címe szerinti alfabetikus sorrendben jelenik meg. Amennyiben a cím alapján történik a keresés, nem szükséges a cím elején megjelenő névelőket beírni, mert az A/Az névelővel kezdődő összes dokumentum az A betűnél jelenne meg. (2/8. ábra)

Az összetett keresés egyszerre több kereső kifejezés megadásával történik. A kutató által megadott kereső kifejezések „and” (azaz logikai ÉS) kapcsolatban állnak egymással. A találatok megjelenése hasonló, mint az egyszerű keresésnél.

A böngészés történhet MOB füzetszám, szerzők, tárgyszavak, kulcsszavak és folyóiratcímek alapján.

A találatok megjelenítése

A keresés alkalmával kapott találatok egyszerre 20-as listában jelennek meg a felületen maximum 20 oldalon keresztül. Ez mindegy 400 tételt is jelenthet. A találati listából kiválasztott címre kattintva az adott tétel részletes leírását és annak összefoglalója tekinthető meg. 2011 év vége óta bizonyos folyóiratcímek alá weblinkeket is elhelyeznek, melyek segítenek a folyóiratok saját honlapjaihoz való eljutását. Emellett a részletes leírásban megtalálható "A cikk állandó MOB linkje" is, melynek célja, ha más adatbázisokba vagy internetes alkalmazásokba kívánja a kutató beemelni, mindig erre a cikkre fog rámutatni. Az érdeklődést felkeltő tételek mellett megjelenő négyzetre kattintással, lehetőség van arra is, hogy az összes (maximum 20) találat részletes leírása megjelenjen ISBD formátumban, de annak összefoglalója nélkül. A tételek kiválasztása azonban egyszerre csak egy találati oldalról történhet.

A kutatás szempontjából releváns kijelölt tételek megtekintésére később is van mód, de adathordozóra történő lementés is választható. (2/10. ábra, 2/11. ábra)

Hozzáférés:

A MOB az alábbi címen érhető el a <http://www.eski.hu>



2/8. ábra Egyszerű keresés a MOB felületén

Magyar Orvosi Bibliográfia
BIBLIOGRAPHIA MEDICA HUNGARICA

Egyszerű keresés | Összetett keresés | Bongészés | Kosár | Sütő

Összetett keresés

Szerzők
 Szerzők
 Szerzők

Összetett keresés a szerzői nevek alapján

- Kistérségi egyenlőtlenségek az otthoni szakápolás vonatkozásában a dél-dunántúli térségben
- A kivonuló mentődolgozók egészségi állapotát befolyásoló főbb tényezők hazánkban
- A munka hatása a kórházi ápolók jólétére Magyarországon az EU csatlakozáskor
- Az OEP otthoni szakápolási kassza igénybevételének területi egyenlőtlenségei
- Az otthoni szakápolás egészségbiztosítási vonatkozásainak elemzése Magyarországon
- Az otthoni szakápolás igénybevételének területi egyenlőtlenségei a dél-dunántúli térségben
- Tudományos közlések az egészségügyben
- Önkéntes ápolói nyilvántartás Németországban

Mind kijelölve: Kiválasztottak megjelenítése

2/9. ábra Összetett keresés és találati lista

Magyar Orvosi Bibliográfia
BIBLIOGRAPHIA MEDICA HUNGARICA

Egyszerű keresés | Összetett keresés | Bongészés | Kosár | Sütő

Kérjük, hogy használat előtt olvassa át a Sütőt!

Egyszerű keresés

Szerzők

<< 1 2 3 4 >>

Találati lista

- 10 éves a HBCS!: a Homogén Betegcsoportok (HBCS) rendszerének tapasztalatai finanszírozói oldalról
- A 2002. évi szervezett lakossági emlőszűrés monitorozásának eredményei
- A 2007. áprilisi 1-i reform hatása a dél-dunántúli egészségügyi intézmények piaci részesedésére
- A 2007. áprilisi 1-jel egészségügyi reformintézkedések hatása az összes kórházi ágyszámra
- 60 év alatti combnyaktrótták csavaros osteosynthesisit követő további ellátások és rizikótényezők kapcsolata
- Az aktív fekvőbeteg szakellátás finanszírozásának visszavezetése a depresszív TVK irányába 2010-2012. között
- Az akut myocardialis infarctus betegségterhe Magyarországon, 2003-2005
- Az akut stroke előfordulása és betegségterhe hazánkban, OEP-adatok alapján
- Battlány-Strattmann László: a ferences-herceg-orvos
- A carpal box radiográfia költséghatékonysági vizsgálata

2/10. ábra Találati lista

Magyar Orvosi Bibliográfia
BIBLIOGRAPHIA MEDICA HUNGARICA

Egyszerű keresés | Összetett keresés | Böngészés | Kosár | Súgó

Részletek

A cikk állandó MOB linkje:
<http://mob.gyemsz.hu/detailsperm.jsp?PERMID=105913> MOB link

MOB: 2013/1
Szerzők: Boncz Imre; Donkáné Verebes Éva; Oberfrank Ferenc
Tárgyszavak: EGÉSZSÉGÜGYI GAZDASÁGTAN; GYÓGYSZERTÁRAK; STATISZTIKA; TÖRVÉNYEK, GYÓGYSZERTÁRI
Folyóirat: Acta Pharmaceutica Hungarica - 2013. 83. évf. 1. sz. [http://www.mgyt.hu] A folyóirat weblinkje

A magyar gyógyszergazdaságossági törvény hatása a patikák számára 2007-2010. között / Boncz Imre, Donkáné Verebes Éva, Oberfrank Ferenc
Bibliogr.: p. 32. - Abstr. hun., eng.
In: Acta Pharmaceutica Hungarica. - ISSN 0001-6659. - 2013. 83. évf. 1. sz., p. 28-32. : ill.

Cékitűzés: Elemzésünk célja a biztonságos és gazdaságos gyógyszer- és gyógyászatisegédeszköz-ellátás, valamint a gyógyszer-forgalmazás általános szabályairól szóló 2006. évi XCVIII. törvény, az ún. "gyógyszer-gazdaságossági" törvény patikák számára gyakorolt hatásának áttekintése. Adatok és módszer. Elemzésünket az Országos Egészségbiztosítási Pénztár (OEP) gyógyszerforgalmi adatbázisa alapján végeztük el. Az elemzés tehát az összes OEP által finanszírozott gyógyszerre kiterjed.

2014. augusztus 14.

2/11. ábra. A releváns találat részletei

2.6. On-line módon elérhető szakmai adatbázisok

Az on-line módon elérhető általában angol nyelvű adatbázisokon tudományos folyóiratokat, releváns szakkönyveket, konferencia és kongresszusi kiadványok érhetőek el a különböző tudományterületeket felölelve, mint az orvostudomány, egészségtudomány, társadalomtudomány, természettudomány és műszaki tudományok.

Az adatbázisok egy része csak kivonatokat és bibliográfiai adatokat tartalmaz, de jelentős számban már elérhetővé teszik a publikációk teljes szövegét és a hozzátartozó illusztrációkat, grafikonokat és táblázatokat is. A kapott találatokat a kutató lementheti, kinyomtathatja, vagy akár elektronikus levélben elküldheti. Az adatbázisok használata (pl. teljes szöveg elérése) a könyvtárak számítógépes hálózat gépeiről lehetségesek, azonban például a PTE Egyetemi Könyvtára is biztosítja hallgatóinak, oktatóinak, hogy - bizonyos beállítások után - az otthoni gépeikről is elérjék a szolgáltatást. Az adatbázisokra való egyéni előfizetések száma a szolgáltatások magas ára miatt alacsony, de gyakran biztosítanak meghatározott intervallumon belül ingyenes kipróbálási lehetőségeket. Az on-line adatbázisok két nagy csoportra oszthatók: a multidiszciplináris adatbázisok, melyek több tudományterület publikációit foglalják összes valamint az ún. szakadatbázisok, melyek egy adott tudományterület bibliográfiáját teszik elérhetővé.

MULTIDISZCIPLINÁRIS ADATBÁZISOK

2.6.1. EBSCOhost

Az egészségügyi adatbázisok közül egyik leggyakrabban használt online szolgáltatás az EBSCOhost. A bostoni székhelyű EBSCO Publishing az *EBSCO Information Services* egyik nagy ágazata működteti. A cég bibliográfiai és elektronikus folyóirat adatbázis forgalmazásával foglalkozik. A forgalmazott adatbázisokhoz egy egységes keresőfelületet biztosít. A könyvtárak adatbázisokhoz való hozzáférését az EBSCO Publishinggel kötött licenc-szerződés határozza meg. A szakirodalom keresés során érdemes arról is tudni, hogy bizonyos folyóiratoknál előfordul az ún. *embargók*, ami azt jelenti, hogy a legfrissebb cikkek teljes szövege csak hónapok (pl.:3, 6, 9,12 hónap) múlva lesznek elérhetőek.

Keresési módok

Mielőtt az adatbázisban megkezdődik a kutatás (2/12. ábra), első lépésében érdemes a *Languages* menüpont alatt a kutató által használt ***nyelvet kiválasztani***, így magyar nyelven is használható az adatbázis keresőfelülete. Az EBSCOhost adatbázisban alapvetően angol nyelvű szakirodalmak kerülnek feldolgozásra, de kis számban más idegen nyelven megjelent publikációkat is tartalmaz, amelyek azonban rendelkeznek angol nyelvű címmel és bibliográfiai leírással is. Az adatbázis keresési nyelve az angol, amerikai és angol helyesírás használatával.

Második lépésben az ***adatbázis kiválasztása*** történik. A Pécsi Tudományegyetem jelenleg a felsorolt adatbázisokkal rendelkezik előfizetéssel. Ezekről az adatbázisokról rövid magyar összefoglaló is olvasható, amely megkönnyíti a kutató számára a választást.

- *Academic Search Complete* - elsősorban felsőoktatási és kutatási intézmények részére készült multidiszciplináris adatbázis.
- *Business Source Premier* - üzleti iparág által leggyakrabban használt kutatói adatbázis.
- *EconLit* - közgazdaságtani adatbázis.
- *Eric* - oktatásügyi- és neveléstudományi információs rendszer.
- *MasterFILE Premier* - általános tárgyú adatbázis.
- *Health Source-Consumer Edition* - fogyasztói egészségügyi tájékoztatási gyűjteményeket tartalmaz az orvostudományok, a táplálkozástudomány, a gyermekgondozás, a sportgyógyászat és az általános közegészségügy területeiről.

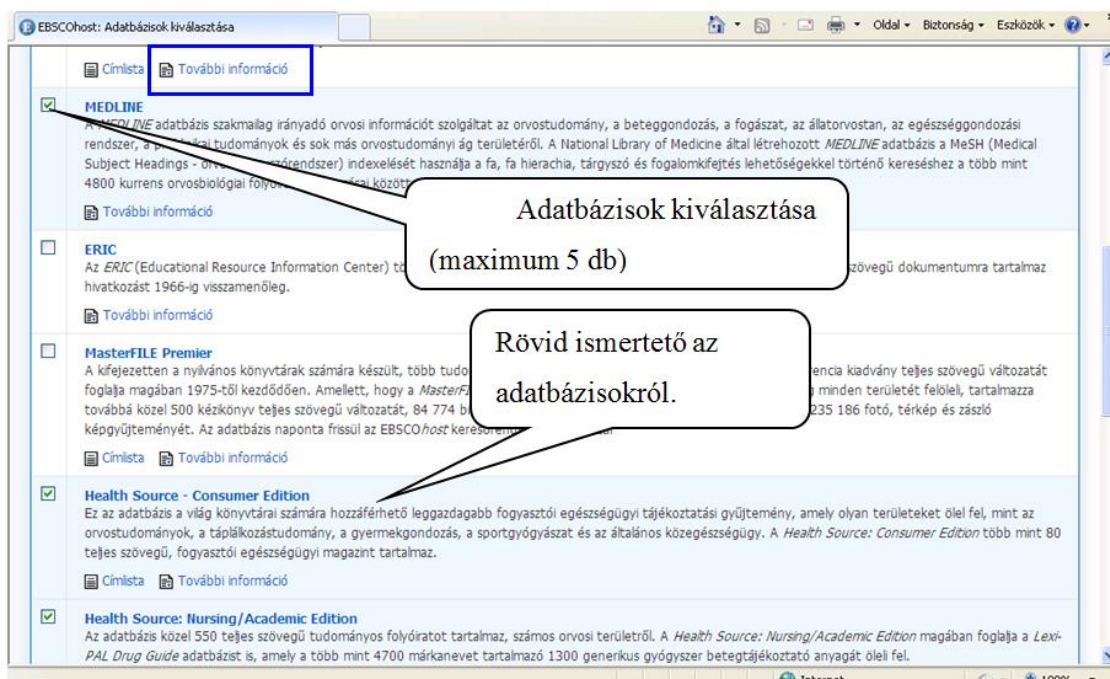
- *Health Source Nursing/Academic Edition* - orvostudományi adatbázis, de itt található a különböző generikus gyógyszerek betegájékoztató anyagát magába foglaló *Lexi-PAL Drug Guide* adatbázis is.
- *MEDLINE* - orvostudományi adatbázis.
- *Regional Business News* - regionális üzleti kiadványok teljes szövegét tartalmazó adatbázis.
- *Library, Information Science & Technology Abstracts* - könyvtár- és információtudomány területének adatbázisa.
- *GreenFILE* - környezetvédelem, globális felmelegedés, alternatív üzemanyagok stb témaköröket feldolgozó adatbázis.
- *Newspaper Source Plus* - több mint 700 újságban megjelenő cikkeket rendszerező adatbázis.

A címlistára klikkelve részletes leírást ad az egyes adatbázisok folyóiratainak listájáról és a kiadványok adatairól (2/13. ábra).

A **keresés** történhet egy vagy több adatbázisból is. Amennyiben egyetlen adatbázisra esik a választása, a kiválasztott adatbázis nevére kell közvetlenül rákattintani vagy kijelölni és a *Folytatás* gombbal tovább lépni. Több adatbázis használata esetén a kiválasztott adatbázisokat kell kijelölni, majd a *Folytatás* gombra kell kattintani.



2/12. ábra Belépés az EBSCOhost adatbázisba



2/13. ábra Az adatbázisok közötti tallózás

A kívánt felületre jutva a keresőmezőbe be kell írni a kereső-kifejezést, majd az egyszerű - (Basic Search) vagy az összetett keresés (Advanced Search) közül választani. Az összetett keresés során több keresőmező áll rendelkezésre. A keresés történhet teljes szöveg, szerző, cím, kulcsszavak, folyóirat neve stb. alapján.

A keresés során a szolgáltató lehetőséget kínál a találati halmaz szűkítésére és bővítésére, ami a mező alatt található. A **bővítési lehetőségek (expanders)** az alábbiak (2/14. ábra):

- *Boole-operátor/kifejezés (Boolean/Phrase)*: A keresés a Boole-operátor támogatásával történik.
- *Keresés az összes kifejezésre (Find all my search terms)*: A keresés során automatikusan az AND logikai operátor kerül alkalmazásra, így csak azok a találatok jelennek majd meg, melyekben a kutató által megadott valamennyi keresőszó egyszerre fordul elő.
- *Keresés a kifejezések bármelyikére (Find any of my search terms)*: A keresés során automatikusan az OR logikai operátor kerül alkalmazásra, vagyis azok a találatok jelennek meg a találati listában, melyekben legalább - a kutató által megnevezett - egyik keresőszó előfordul.
- *Smart text keresés (Smart text searching)*: Alkalmazása révén hosszabb szövegdarabok (legfeljebb 5000 karakter) illeszthetők be a keresőablakba. Segítségével olyan kulcsszavas tömörítvényt hoz létre, amely alapján az adatbázisban lévő cikkek tartalmi kivonataiban

történik a keresés. A szűkítőkkal és bővítőkkal egyaránt használható. Felhívjuk a figyelmet arra is, hogy érdemes lehet a cikkek teljes szövegében való kereséssel együtt alkalmazni.

- *Kapcsolódó szavak alkalmazása (Apply related terms)*: Segítségével a beírt keresőszó szinonimáira és többes számú alakjaira is egyaránt rákeres.
- *Keresés a cikkek teljes szövegében is (Also search within the full text of the articles)*: Itt a keresés a publikáció teljes szövegére vonatkozik. Alkalmazása hasznos, de meg kell jegyezni, ha túl tág keresőszavak kerülnek beírásra, sok lehet az irreleváns találatok száma.

A saját **találatok szűkítési lehetőségei** az alábbiak, melyek együtt, de külön-külön is használhatóak:

- *Teljes szöveg (Full Text)*: A találati oldalon csak a teljes szöveggel rendelkező publikációk listája olvasható.
- *Tudományos (lektorált) folyóiratok (Scholarly (Peer Reviewed) Journals)*: A találati rekordok listáján csak a lektorált, tudományos cikkek jelennek meg.
- *Kiadvány (Publication)*: A kutató által megadott folyóirat cikkeire szűkíti a keresést.
- *Referenciák rendelkezésre állnak (References Available)*: A kereséssel egyidejűleg a keresett cikk szerzője által hivatkozott további cikkeket /Cited References/ is ki lehet gyűjteni, amennyiben a szövegük vagy bibliográfiai adataik rendelkezésre állnak valamely az EBSCOhost-on keresztül elérhető adatbázisban. *Kiadás dátuma...-tól ...-ig (Published Date from ... to ...)*: A keresés során lehetőség van a megjelenés időintervallumának meghatározására is.
- *Kiadvány típusa (Publication Type)*: Egy bizonyos dokumentumtípusra is történhet a szűkítés.
- *Oldalak száma (Number Of Pages)*: Csak a kutató által meghatározott oldalszámú cikkekre fókuszál a keresés során.

A találatok megjelenítése

A találati oldal tetején továbbra is látható a keresőmező és a benne beírt kereső-kifejezés. Ez alatt a *keresési eredményekben (Search Results)* valamennyi, a keresőkérdésnek megfelelő eredmény megjelenik. A rendszer a keresett kifejezéseket kiemeli (félkövér és dőlt betűvel) a megjelenő találati listában.

A keresési eredmények fölött az *Opciók elrejtése (Page Options)* gomb található, melyre kattintva a találatok formátuma állítható be, így a kapott lista olvasható *szabványos*

(Standard), rövid (Brief), csak cím (Title Only) vagy részletes (Detailed) megjelenítésben. Itt állítható be a találatok oldalanként megjelenő száma is (Results per page).

A publikáció címe mellett látható ikonra kattintva a cikk legfontosabb adatai jelennek meg egy ablakban: szerzői nevek (Authors), a forrás (Source), a megjelenés dátuma (Date), a kiadvány típusa (Publication Type), a tárgykörök (Subject Terms), a referátum (Abstract) és az adatbázis (Database).

A találati oldalon lehetőség van további finomításra, így a jobb és bal oldali oszlopok segítségével további szűkítés végezhető el. A baloldali függőleges menüpont további keresési lehetőséget kínál, mint például: teljes szöveg (Full Text), referenciák rendelkezésre állnak (References Available), tudományos (lektorált) folyóiratok (Scholarly (Peer Reviewed) Journals), kiadás dátum (Publication Date), tárgykör (Subject), kiadvány (Publication), vállalat (Company), életkor (Age), földrajz (Geography), Adatbázis (Database). (2/16. ábra)

A találatoknál a PDF teljes szövege (PDF Full Text) ikon jelzi a kutató számára, ha a cikk teljes szövege is olvasható. Erre kattintva azonnal megnyílik a cikk teljes tartalma és látható a publikációban megjelenő ábrák, képek, grafikonok ikonja is.

Új lehetőségként jelenik meg, hogy a Web of Science adatbázishoz is kapcsolódni lehet (Find authors articles in Web of Science), amely a szerző/szerzők további publikációiról nyújthat hasznos információt (2/15. ábra).

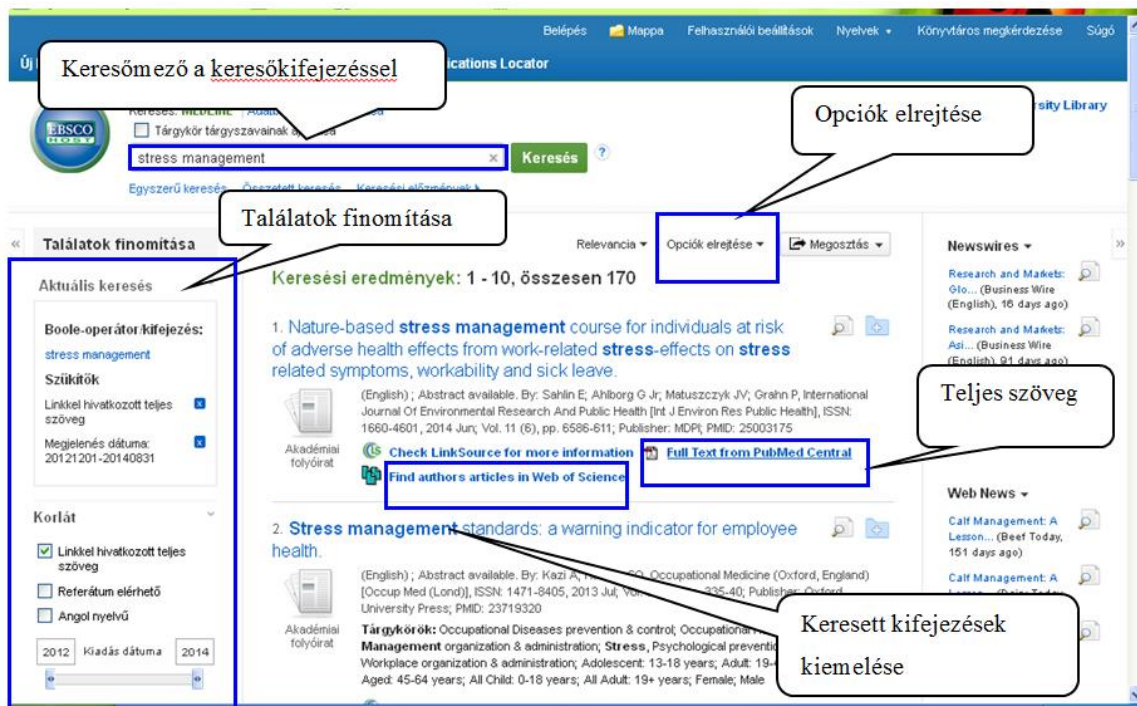
A találatok mappában rendezésére is lehetőség van: Megosztás (Share) → Hozzáadás a mappához (Add to folder). Az így iktatott publikációk az EBSCOhost-ról való kilépés után is megmaradnak. A szükséges dokumentumokat a mentés mellett kinyomtathatók, és az exportálási lehetőségek is adottak. (2/17. ábra)

Hozzáférés

Az EBSCOhost szolgáltatás a <http://search.epnet.com> URL linken keresztül érhető el.



2/14. ábra. Egyszerű keresés



2/15. ábra Találati lista

Összes keresési eredmény
 Akadémiai folyóiratok (54,944)
 Hírek (11,247)
 Magazinok (3,359)
 Szakmai kiadványok (219)
 CEU-k (108)
 Több megjelenítése
 Tárgykör: tezaurszkifejezés
 Tárgykör: fő címsor
 Tárgykör
 Kiadvány
 Vállalat
 Életkor
 Nem
 Földrajz
 NAICS/parág
 Adatbázis
 Összes adatbázis
 MEDLINE (31,641)
 Academic Search Complete (16,007)
 Newspaper Source Plus (10,672)
 CINAHL with Full Text (3,967)

A találatok további szűkítése

4. Second primary **cancers** in survivors of **cervical cancer** in the Netherlands: Implications for prevention and surveillance.

By: Arnold, Melina; Lifang Liu; Kenter, Gemma G.; Creutzberg, Carien L.; Coebergh, Jan Willem; Soerjomataram, Isabelle. *Radiotherapy & Oncology*. Jun2014, Vol. 111 Issue 3, p374-381. 8p. DOI: 10.1016/j.radonc.2014.04.011. Adatbázis: Academic Search Complete

Tárgykörök: CERVICAL cancer -- Patients; CERVICAL cancer -- Treatment; CERVICAL cancer -- Diagnosis; CANCER -- Radiotherapy; CANCER of unknown primary origin; NETHERLANDS

Full Text from ScienceDirect Find authors articles from Web of Science

5. The association between **cervical cancer** screening and mortality from **cervical cancer**: a population based case-control study.

(English) ; Abstract available. By: Vicus D; Sutradhar R; Lu Y; Elit L; Kupets R; Parsat L; Investigators of the Ontario Cancer Screening Research Network. *Gynecologic Oncology [Gynecol Oncol]*, ISSN: 1095-6850, 2014 May; Vol. 133 (2), pp. 167-71; Publisher: Academic Press; PMID: 24569414, Adatbázis: MEDLINE

Tárgykörök: Ontario; Early Detection of Cancer statistics & numerical data; Papanicolaou Test statistics & numerical data; Uterine Cervical Neoplasms diagnosis; Vaginal Smears statistics & numerical data; Adult: 19-44 years; Aged: 65+ years; Middle Aged: 45-64 years; Young Adult: 19-24 years; All Adult: 19+ years; Female

Full Text from ScienceDirect Find authors articles from Web of Science

6. Younger age distribution of **cervical cancer** incidence among survivors of pediatric and young adult **cancers**.

By: Djha, Rohit P.; Jackson, Bradford E.; Tota, Joseph E.; Offutt-Powell, Tabatha N.; Hudson, Melissa M.; Gurney, James G. *Gynecologic Oncology*. Aug2014, Vol. 134 Issue 2, p309-313. 5p. DOI: 10.1016/j.ygyno.2014.05.011. Adatbázis: Academic Search Complete

Tárgykörök: AGE factors in disease; CERVICAL cancer -- Treatment; CANCER in children; CERVICAL cancer; PROGNOSIS; COMPARATIVE studies; CANCER -- Epidemiology

2/16. ábra Szűkítési lehetőségek

←

Részletes rekord

Full Text from PubMed Central

Find authors articles from Web of Science

Hasonló találatok SmartText keresés használatával.

← Találati lista | Keresés finomítása | 1 / 4,689 →

Nature-based stress management course for individuals at risk of adverse health effects from work-related stress-effects on stress related symptoms, workability and sick leave.

Szerzők: Sahlin E
Ahlborg G Jr
Matuszczyk JV
Grahm P

Forrás: *International Journal Of Environmental Research And Public Health* [Int J Environ Res Public Health] 2014 Jun; Vol. 11 (6), pp. 6586-611.

Kiadvány típusa: Journal Article; Research Support, Non-U.S. Gov't

Nyelv: English

Folyóirat adatai: Publisher: MDPI Country of Publication: Switzerland NLM ID: 101238455 Publication Model: Print Cited Medium: Internet ISSN: 1660-4601 (Electronic) Linking ISSN: 16604601 NLM ISO Abbreviation: Int J Environ Res Public Health Subsets: In Process; MEDLINE

Impresszumnevek: Original Publication: Basel : MDPI, c2004-

Referátum: Sick leave due to **stress**-related disorders is increasing in Sweden after a period of decrease. To avoid that individuals living under heavy **stress** develop more severe **stress**-related disorders, different **stress management** interventions are offered. Self-assessed health, burnout-scores and well-being are commonly used as outcome measures. Few studies have used sick-leave to compare effects of **stress** interventions. A new approach is to use nature and garden in a multimodal **stress management** context. This study aimed to explore effects on burnout, work ability, **stress**-related health symptoms, and sick leave for 33 women participating in a 12-weeks nature

Eszközök

Hozzáadás mappához

Nyomtatás

E-mail

Mentés

Hivatkozás

Exportálás

Megjegyzés hozzáadása

Állandó link

Megosztás

A Firefox automatikusan adatokat küld a Mozillának, hogy javíthassuk a felhasználói élményt.

Megosztandó adatok kiválasztása

2/17. ábra A találati eredmény bemutatása

2.6.2. MATARKA (Magyar Folyóiratok Tartalomjegyzékének Kereshető Adatbázisa)

Magyar Folyóiratok Tartalomjegyzékének Kereshető Adatbázisa (MATARKA) a Miskolci Egyetem Könyvtár, Levéltár, Múzeum vezetésével 2002 tavaszától kezdte meg működését. Az interneten szabadon hozzáférhető szolgáltatás jelenleg több, mint 1500 magyar nyelven megjelenő szakfolyóirat tartalomjegyzékét dolgozza fel könyvtári konzorcium keretében, melyben a Pécsi Tudományegyetem Könyvtára is aktívan közreműködik. Az egyes folyóiratok feldolgozása változó, nem minden folyóirat összes lapszáma tekinthető meg.

Keresési módok

A szolgáltatás lehetővé teszi a folyóiratszámok tartalomjegyzékének böngészése mellett a szerző és/vagy a címben levő szavak egyszerű és összetett keresését is.

Az *egyszerű keresés*en belül a kutató két keresési lehetőség közül választhat. A *gyors keresés (Tájékozódó keresés)* során a keresődobozba maximum öt szó írható be. A keresésben a szavak között ÉS kapcsolat van, és csak abban az esetben kap eredményt, ha a kutató által megadott minden egyes kifejezésre van találat. Természetesen a találatok alapján további szűkítésre van lehetőség. Az egyszerű keresés másik lehetősége, amikor a kutató négy egymás alatt megtalálható keresőmezőt tölthet ki: (1) szerző, (2) további szerző, (3) cikk címe-kulcsszavak, (4) cikk címe-részlet.

A felsorolt keresőmezőkben egyszerre több szó megadására is lehetőséget ad, melyeket automatikusan ÉS kapcsolattal köti össze. A kutató munka során azonban nem kötelező mindegyik beviteli mezőt kitölteni.

A keresés egyaránt szűkíthető - értelemszerűen akár együtt is alkalmazható - a kiválasztott évszám, szakterület, folyóirat, csak teljes szöveggel elérhető cikkek és csak az Országos Széchényi Könyvtár Elektronikus Periodika Archívumában (EPA) archivált cikkek alapján (2/18. ábra).

Böngészni az adatbázisban megtalálható szerzőnevek és cím-kulcsszavak alapján lehet.

Emellett a keresés történhet folyóiratok között is. A folyóiratok alfabetikus sorrendben állnak, és a kiválasztás kezdőbetű vagy szakterület szerint történhet. A kiválasztott kezdőbetűre kattintva a folyóiratok összefoglaló táblázata jelenik meg. A kiválasztott folyóíratra kattintva a feldolgozott évfolyamok, évszámok és az interneten elérhető cím jelennek meg, majd a szám kiválasztása után a tartalomjegyzék kerül bemutatásra. A keresett folyóirat adatairól pedig az alábbi információkat nyújtják a felhasználó számára: címadatok, közreadó, a folyóirat indulásának (és megszűnésének) éve, ISSN szám, szakterület, teljes szövege elérése, honlap és nyelv.

A találatok megjelenítése

A találati rekordok időrendi sorrendet követve jelennek meg a találati halmazban, elől mindig legfrissebb tanulmány. Megadásra kerül a szerző(k) neve(i), a publikáció címe és a megjelenéssel kapcsolatos adatok (folyóirat címe, megjelenés ideje, évfolyam, évszám, oldalszám) (2/19.ábra).

A közlemények körülbelül 15 százaléka érhető el közvetlenül teljes szöveggel, a többről térítés ellenében cikkmásolat kérhető.

A honlapon egyéb kiegészítő információk is olvashatóak. A *súgó* menüpont alatt az eddig készült dokumentumok és publikációk listája, a készítők (közreműködő könyvtárak) névsora jelenik meg. A *vendégkönyvben* a szolgáltatással kapcsolatos vélemények, javaslatok fogalmazhatóak meg, míg a *kapcsolattartók* segítséget nyújtanak a felmerült kérdésekben és cikkmásolat ügyében, a *statisztikák* fülre kattintva betekintést enged az adatbázissal kapcsolatos statisztikai mutatók megismerésére, mint például látogatók száma, és végezetül megismerhető a MATARKA Egyesület működéséhez szükséges alapvető dokumentumok is.

Hozzáférés:

A MATARKA az alábbi címről érhető el: <http://www.matarka.hu>

The screenshot shows the MATARKA search interface. At the top, there is a navigation bar with the MATARKA logo and the text "MATARKA - Magyar folyóiratok tartalomjegyzékeinek kereshető adatbázisa". Below this, there are several menu items: "Keresés", "Készítők", "Dokumentumok", "Vendégkönyv", "Kapcsolat", and "Egyesület". A "Gyors keresés" (Quick search) box is prominently displayed. Below the navigation bar, there are three search options: "Egyszerű keresés" (Simple search), "Haladó keresés" (Advanced search), and "Bongészés" (Browse). The "Egyszerű keresés" section is highlighted with a blue box and contains a search input field with the text "Tájékoztató keresés (párhuzamosan kereshet címben, szerzőben)" and a "Keresés" button. Below this, the "Keresés" section is highlighted with a blue box and contains several input fields: "Szerző:" with the value "Ács Pongrác", "További szerző:", "Szerző - kulcsszavak:", "Cikk címe - kulcsszavak:", and "Cikk címe - részlet:". A "Keresés" button is also present in this section. At the bottom, the "Találatok szűkítése" (Refine results) section is highlighted with a blue box and contains several filters: "1800 évtől" and "2014 évig", a "Válasszon szakterületet!" dropdown menu, and a "Válasszon folyóiratot!" dropdown menu. There are also two checkboxes: "Csak teljes szöveggel elérhető cikkek" and "Csak az EPA-ban archivált cikkek".

2/18. ábra Egyszerű keresés több szerzővel



2/19. ábra Találati lista

2.6.3. MOKKA (Magyar Országos Közös Katalógus)

A Magyar Országos Közös Katalógus egy olyan bibliográfiai és lelőhely-adatbázis, mely lehetőséget kínál a Magyarországon működő nemzeti, tudományos és felsőoktatási valamint megyei és városi könyvtárainak állományában történő keresésre.

Keresési lehetőségek

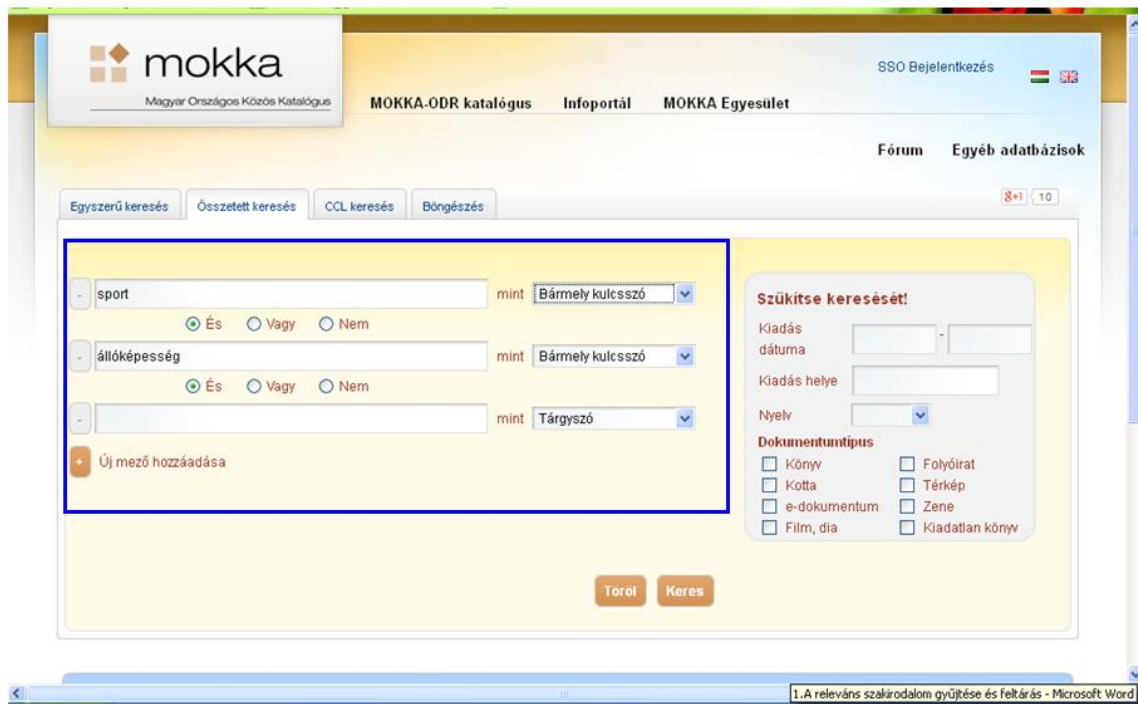
Az egyszerű és összetett keresés történhet szerző, cím, tárgyszó és bármilyen kulcsszó alapján. A keresődobozba beírt kereső-kifejezés után a keresés a *Search* gombra kattintva a keresés indítható (2/20. ábra).

A találatok megjelenítése

A kapott találatok megmutatják a felhasználói számára, hogy a keresett dokumentum mely könyvtárakban lelhető fel és ennek megfelelően a kiválasztott könyvtár adatbázisában folytathatja a keresést, ahol pontos információt kap annak állapotáról (kölcönözhető vagy nem) (2/21. ábra).

Hozzáférés:

A MOKKA a <http://www.mokka.hu> címen érhető el.



2/20. ábra. Összetett keresés



2/21. ábra Találati lista

2.6.4. OVID

Különböző tudományterületek (orvostudomány, társadalomtudomány, műszaki tudományok, humán tudományok) szakemberei számára az elektronikus információkhoz való hozzáférést biztosít az Ovid klasszikus tartalomszolgáltatója. Érdekességként kell megemlíteni, hogy a

MEDLINE adatbázist is ők adták ki először CD-ROM változatban. Az információszolgáltatásuknak 3 alappillére van: adatbázisok, elektronikus folyóiratok és elektronikus könyvek. A MEDLINE adatbázis mellett csak a Wolters Kluwer kiadó folyóiratai érhetőek el.

Keresési lehetőségek

Az online módon történő keresés történhet adatbázisból, könyvből és folyóiratokból. Ha a választás az *adatbázis szerinti keresésre* esik, első lépésben a listán látható adatbázisok közül kell választani. Egy alkalommal maximum öt adatbázist lehet kipipálni. Ezt követően az oldal bal alsó szélén található *Select Resource(s)* gombra klikkelve elindítható a keresés.

Lehetőség kínálkozik más adatbázisba való keresésre is, ehhez elegendő a *Change* gombra kattintani, ami visszaviszi a kutatót a választható adatbázisokhoz. Az OVID szolgáltatás hat keresési lehetőséget ajánl fel a felhasználó számára: *Basic Search, Find Citation, Search Tools, Search Fields, Advanced Ovid Search, Multi-Field Search* (2/1. táblázat)

2/1. táblázat. A keresés lefuttatásának lehetőségei

A keresés típusa	Összefoglalója
Basic Search,	Természetes nyelven való keresést jelent, ahol a keresés szintaxisát és a tudományterület szakkifejezéseit nem szükséges ismerni. A megadott keresőkifejezésekkel a találatok rangsoroltak, de a releváns találat nem számottevő. Segítő funkcióként jelenik meg: a helyesírás ellenőrző és a keresés kiterjesztése a kapcsolódó fogalmakra.
Find Citation,	Ezt a keresési módot akkor célszerű választani, ha egy <i>konkrét közleményt</i> keres néhány bibliográfiai adat ismeretében, pl. a publikáció megjelenése, címe vagy szerző megadása szerint
Search Tools,	A kutató számára segít megtalálni a <i>kulcsszavainkhoz tartozó megfelelő tárgyszavakat</i> illetve a keresést is felajánlja. Ezek a lehetőségek már az „Advance Ovid Search” részét is képezik.
Search Fields,	Különböző szempontok alapján nyújt segítséget a kereséshez, mint például intézményi név, DOI szám, ISSN szám, kiadó, ország stb.
Advanced Ovid Search,	A hagyományos összetett keresés során több lehetőség kínálkozik a keresésre, mint például kulcsszó, szerző, közlemény címe és folyóiratcím alapján. Amennyiben a kutató egy adatbázist választ ki, az Ovid felajánlja a tárgyszavas keresést is.
Multi-Field Search.	Több keresőmező kombinálásával nyújt gyorsabb keresési lehetőséget, így egyszerre több feltételnek megfelelő publikáció kerülhet a találati halmazba.

Forrás: Csajbók 2009

Az alapkeresés során a keresődobozba kell a kutató által meghatározott kereső-kifejezéseket beírni (2/22. ábra). Az összetett keresés történhet tárgyszavak (*Keywords*), szerző (*Author*), címe (*Title*) és folyóirat (*Journal*) szerint. A keresés során van lehetőség szűkítésre a *Limits*

gombra kattintva, ahol dátum vagy a dokumentum típusa alapján összefoglaló (*Abstracts*), teljes szöveg (*Full Text vagy Review* stb.) szűkíthető a találati halmaz.

The screenshot shows the OvidSP search interface. At the top, there are navigation tabs: Search, Journals, Books, My Workspace, Primal Pictures, and Multimedia. Below this is a 'Search History' section with a table showing search results. The main search area includes a search bar, a 'Search' button, and a 'Limits' section. The 'Limits' section contains several checkboxes for filtering results, such as 'Daily Update', 'Local Holdings', 'Review Articles', 'Ovid Full Text Available', 'PsycARTICLES', 'Articles with Abstracts', and 'Original Articles'. There is also a 'Publication Year' dropdown menu. Callouts point to the search bar area as 'Keresési lehetőségek' and the 'Limits' section as 'Szűkítési lehetőségek'. At the bottom, there are language selection options: English, Français, Deutsch, 日本語, 繁體中文, Español, 简体中文, and 한국어.

2/22. ábra Egyszerű keresés

The screenshot shows the OvidSP search interface. At the top, there are navigation tabs: Search, Journals, Books, My Workspace, Primal Pictures, and Multimedia. Below this is a 'Search History' section with a table showing search results. The main search area includes a search bar, a 'Search' button, and a 'Limits' section. The 'Limits' section contains several checkboxes for filtering results, such as 'Daily Update', 'Local Holdings', 'Review Articles', 'Ovid Full Text Available', 'PsycARTICLES', 'Articles with Abstracts', and 'Original Articles'. There is also a 'Publication Year' dropdown menu. Callouts point to the search bar area as 'Keresési lehetőségek' and the 'Limits' section as 'Szűkítési lehetőségek'. At the bottom, there are language selection options: English, Français, Deutsch, 日本語, 繁體中文, Español, 简体中文, and 한국어.

2/23. ábra Összetett keresés

A találatok megjelenítése

A találati eredmények az alábbiakról nyújtanak információt: cikk címe, szerző(k) nevei, folyóirat címe, publikáció bibliográfiai adatai (kiadási év, évfolyam szám, lapszám, oldalszám), és az absztrakt olvasható. Ha a szolgáltató engedélyezi, PDF formátumban a felhasználni kívánt cikk letölthető, vagy adathordozóra elmenthető (2/24. ábra).

A találati oldal jobb oldali oszlopában látható a *Complete Reference*, ami az adott rekordról nyújt bővebb információt, ezáltal megtekinthető a bibliográfiai adatok mellett a szerzők munkahelye, a cikk MeSH tárgyszavai, elolvasható a rövid kivonat, és - ha a hozzáférés szabad - a cikk elérhetőségét (2/25. ábra). A kutatás szempontjából releváns találatokat a felhasználó kinyomtathatja, lementheti vagy exportálhatja word vagy pdf formátumban.

Érdekes, hogy a szolgáltató minden hónapban elérhetővé tesz egy adatbázist és egy folyóiratot, amihez külön-külön kell regisztrálni. Az adatbázis regisztráció a *Resource of the Month*, a folyóirat regisztráció a *Journal of the Month* menüpont alatt található.

Primal Pictures

A keresések mellett a Primal Pictures az emberi test anatómiájáról ad átfogó szolgáltatást a dinamikus interaktív multimédia. 3D animációk segítik az tantárgy elsajátítását, kiegészítve biomechanikai és sebészi folyamatok prezentálásával (2/26. ábra).

Hozzáférés

<http://www.ovid.com/site/index.jsp>

Results Tools Options

All Print Email Export Add to My Projects Keep Selected

Search Information Clear Selected View: Title Citation Abstract 10 Per Page 1 GO Next

You searched:
 stress management.mp.
 [mp=title, abstract, full text, caption text]
 - Search terms used:
 management
 stress

Search Returned:
 9869 text results

Sort By:

Customize Display

Filter By

Add to Search History

+ Selected Only (0)

- Years
 All Years
 Current year
 Past 3 years
 Past 5 years
 ▶ Specific Year Range

1. **The Relationship of Chronic and Momentary Work Stress to Cardiac Reactivity in Female Managers: Feasibility of a Smart Phone-Assisted Assessment System.**
 Lumley, Mark A. PhD; Shi, Weisong PhD; Wiholm, Clairy PhD; Slatcher, Richard B. PhD; Sandmark, Helene PhD; Wang, Shinan MS; Hytter, Anders PhD; Arnetz, Bengt B. MD, PhD
Psychosomatic Medicine.
 [Original Article: PDF Only]
 AN: 00006842-900000000-99173.
 B 13 Raktari jelzet: P 59 (Psychosom Med); USA, New York, ISSN 0033-3174; 1960:22 -
 Status
 Publish Ahead of Print, POST AUTHOR CORRECTIONS, 30 July 2014
 ▶ View Abstract
 PDF (341KB) + My Projects

- Ovid Full Text
- Table of Contents
- Abstract Reference
- Complete Reference
- Find Similar
- Find Citing Articles
- Library Holdings
- Request Permissions
- Internet Resources

2. **Web-Based Parenting Skills Program for Pediatric Traumatic Brain Injury Reduces Psychological Distress Among Lower-Income Parents.**
 Raj, Stacey P. MA; Antonini, Tanya N. MA; Oberjohn, Karen S. MA; Cassidy, Amy PhD; Makoroff, Kathi L. MD; Wade, Shari L. PhD
Journal of Head Trauma Rehabilitation.
 [Original Article: PDF Only]
 AN: 00001199-900000000-99766.
 Status

- Table of Contents
- Abstract Reference
- Complete Reference
- Find Similar
- Find Citing Articles
- Library Holdings

2/24. ábra Találati lista

1 GO Search Results | Next

1.

Accession Number 00006842-900000000-99173.

Author Lumley, Mark A. PhD; Shi, Weisong PhD; Wiholm, Clairy PhD; Slatcher, Richard B. PhD; Sandmark, Helene PhD; Wang, Shinan MS; Hytter, Anders PhD; Arnetz, Bengt B. MD, PhD

Institution From the Departments of Psychology (M.A.L., R.B.S.), Computer Science (W.S., S.W.), and Family Medicine and Public Health Sciences, Cardiovascular Research Institute and Institute of Environmental Health Sciences (B.B.A.), Wayne State University, Detroit, Michigan; Department of Public Health and Caring Sciences (C.W., B.B.A.), Uppsala University, Uppsala, Sweden; Department of Public Health Sciences (H.S.), Malardalen University, Vasteras, Sweden; and School of Business and Economics (A.H.), Linnaeus University, Vaxjo, Sweden

Title **The Relationship of Chronic and Momentary Work Stress to Cardiac Reactivity in Female Managers: Feasibility of a Smart Phone-Assisted Assessment System.**[Article]

Source Psychosomatic Medicine.

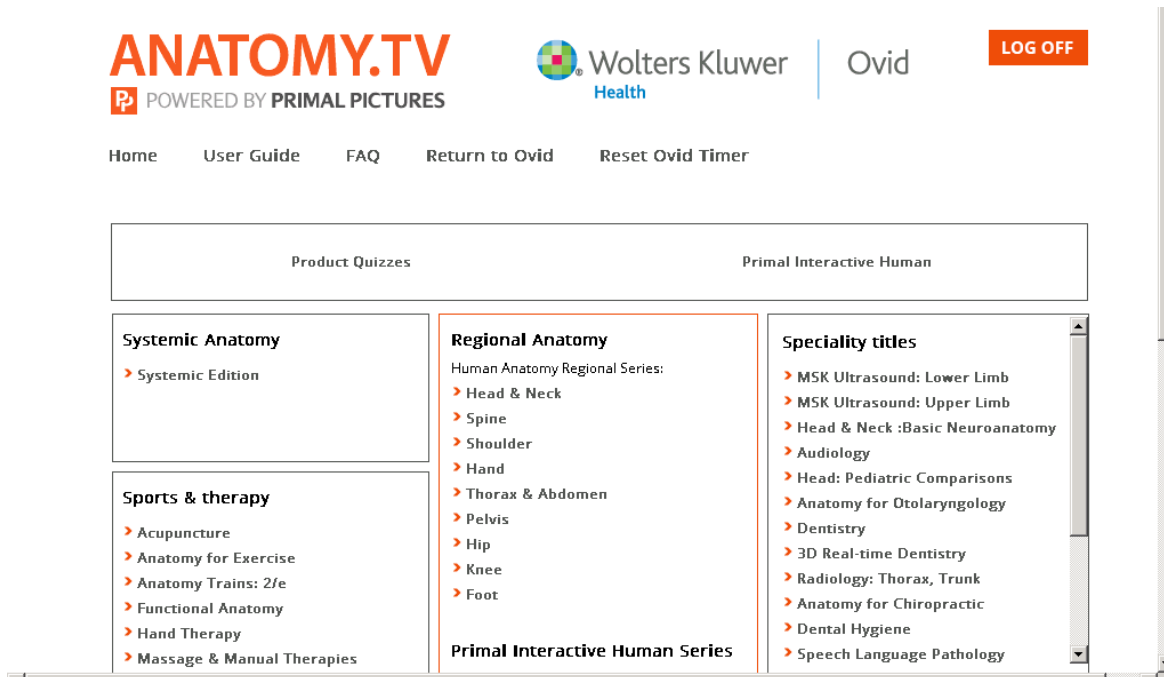
Status Publish Ahead of Print, POST AUTHOR CORRECTIONS, 30 July 2014

Local Message B 13 Raktari jelzet: P 59 (Psychosom Med); USA, New York, ISSN 0033-3174; 1960:22 -

Abstract
 Objectives: To evaluate a wireless smart phone-assisted (SPA) system that assesses ongoing heart rate (HR) and HR-triggered participant reports of momentary stress when HR is elevated during daily life. This SPA system was used to determine the independent and interactive roles of chronic and momentary work stress on HR reactivity among female managers.
 Methods: A sample of 40 female managers reported their chronic work stress and wore the SPA system during a regular workday. They provided multiple reports of their momentary stress, both when triggered by increased HR and at random times. Relationships among chronic stress, momentary stress, and HR were analyzed with hierarchical linear modeling.
 Results: Both chronic work stress (b = 0.08, standard error [SE] = 0.03, p = .003) and momentary work stress (b = 1.25, SE = 0.62, p = .052) independently predicted greater HR reactivity, adjusting for baseline

- Ovid Full Text
- Table of Contents
- Abstract Reference
- Find Similar
- Find Citing Articles
- Library Holdings
- Request Permissions
- Internet Resources

2/25. ábra A kiválasztott rekord információ tartalma



2/26. ábra Primal Pictures

2.6.5. ScienceDirect

Hazánkban is elérhető az Elsevier tudományos kiadó gondozásában 1997 óta működő Science Direct adatbázis, amely az orvostudomány, természettudomány, társadalomtudomány és műszaki tudományok területén mintegy 2500 saját kiadású papíralapú folyóiratának elektronikus változatához teljes szövegű hozzáférést biztosít, de kiegészül más kiadók e-folyóirataihoz is, és mintegy 26 000 könyvet indexel. Az 1995 előtt megjelent Elsevier folyóiratcikkek csak bibliográfiai adatokkal érhető el. Lehetővé teszi a bibliográfiai adatok közötti keresést, és együttesen oldja meg az adatbázis és a teljes szöveges hozzáférést. Az adatbázis használatának előnye, hogy tartalmaz olyan folyóiratokat is, amelyekben a megjelenő publikációk már a nyomtatott változat megjelenése előtt olvashatók.

Keresési módok

A ScienceDirect szolgáltatása széleskörű. A keresés lehet egyszerű, haladó (*Advanced search*), szakértői (*Expert search*) de emellett böngészésre (*Browse*) is lehetőséget kínál. Az egyszerű keresés alkalmával a kezdőfelületen található keresőmezőbe bármilyen mező tartalmára történhet keresés, azonban ezeket tovább lehet pontosítani a szerző (*Author name*), a folyóirat vagy könyv címének (*Journal or book title*), és a kiadással kapcsolatos adatok,

mint a pontos évfolyam (*Volume*), szám (*Issue*), oldal (*Page*) megadásával, ezt követően a „*Submit Quick Search*” gombra klikkelve a keresés indítható (2/27. ábra).

Az adatbázisban az Összetett keresés (*Advanced search*) a dokumentum típusok – folyóirat (*Journal*), könyv (*Books*), enciklopédia szócikkei (*Reference Works*), képek (*Images*)- szerint történik. Minden dokumentumtípus külön űrlappal rendelkezik, különböző adatmezőkkel és keresőfelülettel ellátva. Az adatbázisban való keresés alkalmával használhatóak a Boole-operátorok és a csonkolás karakterei, az idézőjelbe tett szavakra pontos kifejezésként keres.

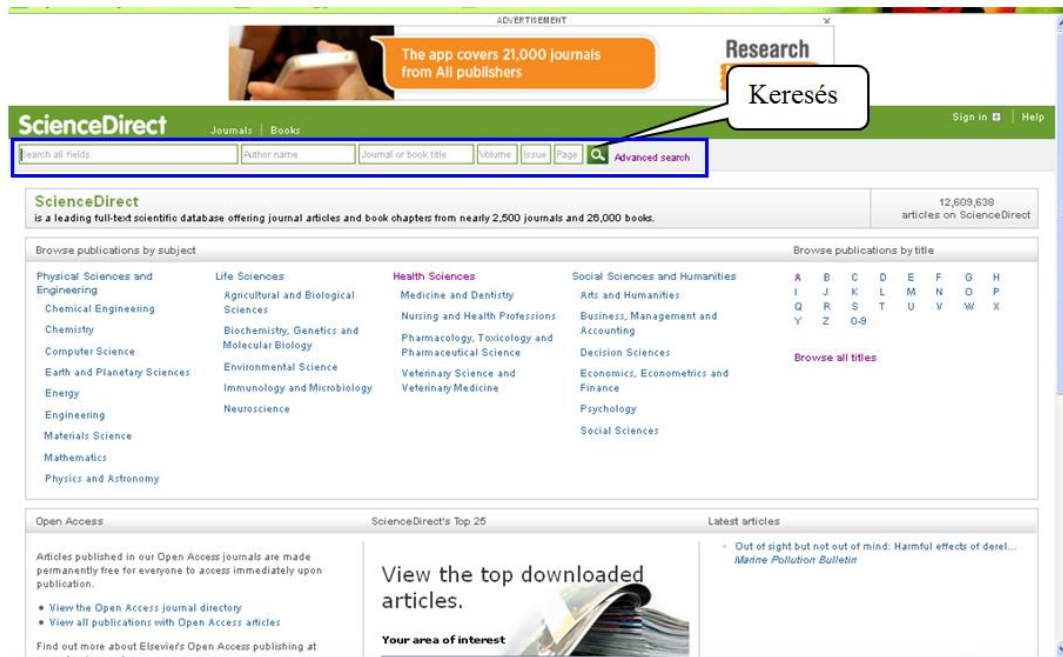
Az adatbázis nyitófelületén található böngésző (*Browse publications by title*) segítségével a ScienceDirect adatbázisában elérhető az összes folyóirat és könyv listája betűrendi sorrendben. A megfelelő kezdőbetű kiválasztása után, arra rákattintva a kezdő címeket kapja meg felhasználó. A kapott lista több szempont szerint tovább szűkíthető. Kiválasztható a bal oldali mezőből a tudományterület, a megjelenő lista felett az összes publikáció (*All publications*) gomb alatt pedig azt lehet eldönteni, hogy a folyóiratok (*All journals*), könyv (*Books*), könyvsorozatok (*Book Series*), kézikönyv (*Handbook*), enciklopédiák szócikkei (*Reference Works*) között kíván-e keresni. A hozzáférést tekintve előfizetett (*Subscription*), nyílt hozzáférésű (*Open Access Journal*) részleges nyílt hozzáférésű (*Contains Open Access*) tanulmányok közül lehet választani.

A találatok megjelenítése

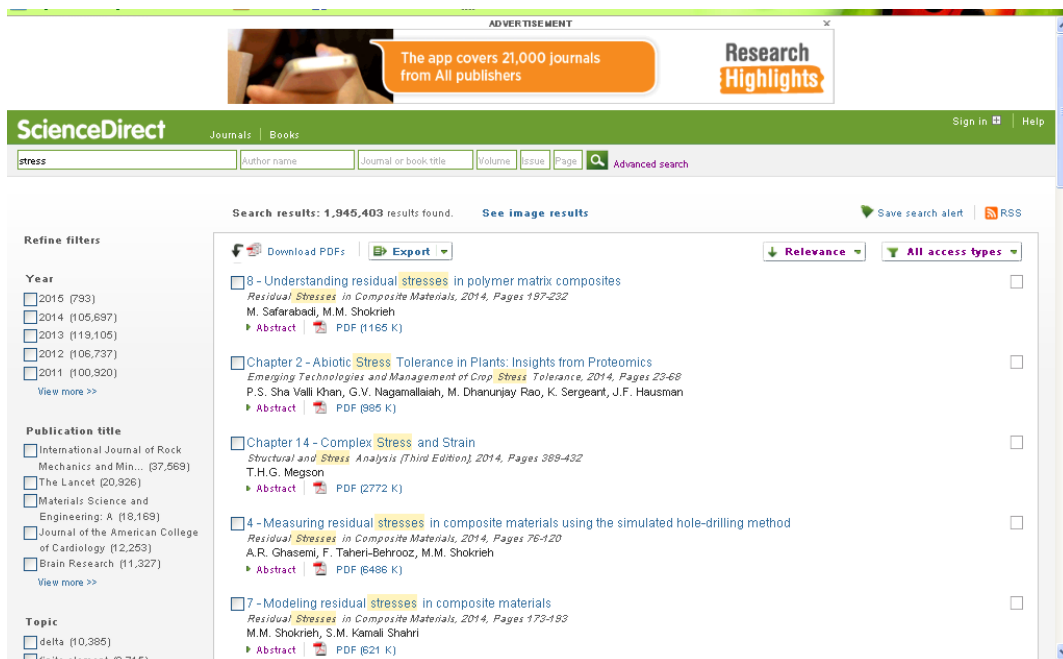
A találat megjelenítésénél a dokumentum bibliográfiai adatai találhatóak: a publikáció címe és típusa, a folyóirat neve, a megjelenésre vonatkozó adatok és a szerző(k) neve(i). A találatok mellett megjelenő zöld vonalas négyzet jelöli, hogy az adott publikáció teljes szövege elérhető, a fehér négyzet pedig azt jelzi, hogy nem rendelkezik előfizetéssel, emiatt csak a rövid összefoglaló tanulmányozható. A találatok további szűkítésére is van lehetőség: megjelenés éve (*Year*), folyóirat címei (*Publication Title*), témák (*Topic*), dokumentum típusa (*Content type*) alapján. A teljes szöveg PDF formátumban megnyitható, lementhető és kinyomtatható (2/28. ábra).

Hozzáférés:

A ScienceDirect a <http://www.sciencedirect.com> oldalon található.



2/27. ábra. ScienceDirect keresőfelülete



2/28. ábra Találati lista

2.6.6. Scopus

Az Elsevier kiadó által működtetett Scopus multidiszciplináris bibliográfiai adatbázis 2004 óta működik, melyet az egyik legnagyobb absztrakt és idézet kereső adatbázisként tartják nyilván.

Tematikai megoszlását tekintve a következő tudományterületeket foglalja magába: élet- és orvostudományok, kémia-, fizika-, matematika- és mérnöki tudományok, társadalomtudomány, pszichológia, közgazdaságtudomány, biológia, agrártudományok, környezetvédelem és általános tudományok. Az idézetek visszakeresésére 1996-ig van lehetőség. A naponta frissített adatbázis nemcsak az angolul megjelenő tudományos munkákat dolgozza fel, elérhetők francia, német, spanyol, kínai nyelvű publikációk is.

Keresési módok

A Scopus adatbázis keresőkínálata sokrétű. A nyitóoldalon végezhető egyszerű keresés dokumentum (*Document Search*), szerző (*Author Search*), intézmény név (*Affiliation Search*) keresésével, valamint összetett keresés (*Advanced Search*) melyek adatmezőkkel és keresőfelülettel ellátva külön űrlappal rendelkeznek (2/29. ábra).

A *Document Search* űrlapon a kereső mezőbe kell a kutatás szempontjából lényeges kulcsszavakat beírni. A kereső-kifejezéseket Boole-operátorokkal kell összekötni. A keresés finomítható az időintervallum, dokumentumtípus vagy tudományterületre egyaránt.

A szerzők keresése során a vezetéknevet és a keresztnév kezdőbetűit kell a kereső mezőbe beilleszteni, majd a *Search* gombbal a keresés indítható.

A találatok megjelenítése

A *Documentum Search* keresési mód találati oldalán jól látható a találati halmaz (*Document Results*), amely megmutatja, hogy hány szakirodalom foglalkozik az adott témával. Száz vagy annál több találat esetén mindenképpen ajánlott a találati eredményeket tovább szűkíteni a megadott szempontok alapján (*Refine results*): évszám (*Year*), szerzői nevek (*Author Name*), tárgyszó (*Subject Area*), dokumentum típusa (*Document Type*), forrás címe (*Source Title*), kulcsszavak (*Keyword*), munkahely (*Affiliation*), ország (*Country*), forrás típus (*Source Type*), nyelv (*Language*).

A találati listában a publikáció címe, a szerző(k) neve(i), a megjelenés éve, a megjelenéssel kapcsolatos adatok, a témakör és a citáció száma olvasható. A lista rendezhető (*Short on*): dátum (*Date*), citáció száma (*Cited by*) és relevancia (*Relevance*) vagyis szerző és publikációk címének alfabetikus sorrendje alapján is (2/30. ábra).

A citációra vonatkozó adatokat az idézettség áttekintése (*citation overview*) fülre való kattintás után teszi látható. A kapott oldalon egy táblázat 1996-tól szemlélteti, hogy az adott cikket évenkénti bontásban hányszor idézték. Az idézettséget évszám, idézetek számának növekvő vagy csökkenő sorrendjében is rendezhető. Lehetőség van arra is, hogy a szerző az önhivatkozásait kizárja a „kizárás az idézettségek áttekintéséből” (*Exclude from citation overview*) ablak kijelölésével.

A szerzők keresése során a találati halmazban a szerző neve (ékezet nélkül) és a szerző profiljához csatolt névváltozatok jelennek meg, emellett feltüntetésre kerül a szerző által publikált tanulmányok száma is. A kiválasztott névre való klikkelés után a szerző publikációs teljesítményére vonatkozó összefoglaló látható, ami tájékoztatást ad a megjelent publikációk számáról, a citáció-számáról, a Hirsch-indexről, a szerzőtársakról, a megjelent publikációk témaköreiről és dokumentumokban a felhasznált irodalmak számáról (*References*). Emellett grafikonok, összefoglaló táblázatok segítik az adott szerző megjelenő publikáció számának szemléltetését (*View Author Evaluator, View citation overview, View h-Graph*).

Az intézmények szerinti keresés (*Affiliation Search*) alkalmával az adott intézmény publikációs dokumentumait lehet megtekinteni, a kapott találatok dátum (*Date*), citáció (*Cited by*) és relevancia (*Relevance*) szerint is rendezhetőek (2/31. és 2/32. ábra).

Megjegyzés

Felhívjuk a figyelmet arra, hogy a Scopus és a Web of Science adatbázisában található adott szerzőre vonatkozó hivatkozás és Hirsch-index eltérő lehet, mert az adatbázisok a saját tartalmuknak megfelelően készíti el az adott szerző tudományos munkásságának összefoglaló listáját.

Hozzáférés

A Scopus <http://www.scopus.com> honlapon érhető el.

Document search | Author search | Affiliation search | Advanced search

Search for... *Eg., "heart attack" AND stress* **Article Title, Abstract, Keywords**

Limit to:

Date Range (inclusive)
 Published **All years** to **Present**
 Added to Scopus in the last **7** days

Document Type
ALL

Subject Areas
 Life Sciences (> 4,300 titles.)
 Health Sciences (> 6,800 titles. 100% Medline coverage)
 Physical Sciences (> 7,200 titles.)
 Social Sciences & Humanities (> 5,300 titles.)

Resources

Follow @Scopus on Twitter for updates, news and more

Access training videos

Learn about alerts and registration


About Scopus
What is Scopus
Content coverage

Language
日本語に切り替える
切换到简体中文

Customer Service
Help and Contact
Live Chat

About Elsevier
Terms and Conditions
Privacy Policy

Copyright © 2014 Elsevier B.V. All rights reserved. Scopus® is a registered trademark of Elsevier B.V. Cookies are set by this site. To decline them or learn more, visit our Cookies page.



2/29. ábra. Keresőfelület

Scopus | Register | Login

Search | Alerts | My list | Settings

Szűkítés | **Találati lista**

Save | Set alert | Set feed

View 6 patent results | Analyze results

Sort on: Date | Cited by | Relevance

Search results... | Export | Download | View citation overview | View Cited by | More... | Show all abstracts

Refine	Search Results	Authors	Year	Journal	Citations
Year	<input type="checkbox"/> Economic burden of long-term care of rheumatoid arthritis patients in Hungary	Horváth, Z., Sebestyén, A., Ósterle, A., (...), Bagosi, G., Boncz, I.	2014	The European Journal of Health Economics	0
<input type="checkbox"/> 2014 (2)	<input type="checkbox"/> Az otthoni szakápolás egészségbiztosítási vonatkozásainak elemzése Magyarországon [Health insurance data analysis on home nursing care in Hungary]	Cs. Horváth, Z., Sebestyén, A., Molics, B., (...), Bagosi, G., Boncz, I.	2014	Orvosi Hetilap	0
<input type="checkbox"/> 2012 (1)	<input type="checkbox"/> Is 'meaningfulness' a general mediating factor? The salutogenic revolution of question-setting in health science and occupational psychology	Varga, K., Tóth, Á., Rozniar, J., (...), Betlehem, J., Jeges, S.	2012	European Journal of Mental Health	0
<input type="checkbox"/> 2010 (2)	<input type="checkbox"/> A kivonuló mentő dolgozók egészségi állapotát befolyásoló főbb tényezők hazánkban [Major contributing factors of self perceived health in Hungarian ambulance personnel]	Betlehem, J., Horváth, A., Gondócs, Z., (...), Boncz, I., Oláh, A.	2010	Orvosi Hetilap	2
<input type="checkbox"/> 2008 (3)	<input type="checkbox"/> Prehospital emergency care in Hungary: What can we learn from the past?	Gondócs, Z., Oláh, A., Marton-Simora, J., (...), Schaefer, J., Betlehem, J.	2010	Journal of Emergency Medicine	1
<input type="checkbox"/> 2007 (6)	<input type="checkbox"/> Stress, geomagnetic disturbance, infodrain and circadian	Oláh, A., Jozsa, B.	2008	Neurotoxic Research	4

30. ábra Találati lista

Scopus Register | Login

Search | Alerts | My list | Settings Live Chat | Help and Contact | Tutorials

Author last name "Oláh", Author first name "Andras" Edit

3 author results About Scopus Author Identifier Sort on: Document Count | Author (A-Z) ...

Show exact matches only Show documents | View citation overview | Request to merge authors

Refine

Source Title

- 3rd International Symposium on Wireless Pervasive Computing ISWPC 2008 Proceedings (1)
- American Journal of Public Health (1)
- Annales Des Telecommunications Annals of Telecommunications (1)
- Annals of the New York Academy of Sciences (1)
- Applied Ecology and Environmental Research (1)

Affiliation

- Budapesti Corvinus Egyetem (1)
- Budapesti Muszaki es

<input type="checkbox"/> Oláh, András 1 Oláh, Andras Oláh, András Oláh, A.	24 Medicine ; Pharmacology, Pecs Tudomanyegyetem Pecs Hungary Toxicology and Pharmaceutics ; Nursing; ...
<input type="checkbox"/> Oláh, András L. 2 Oláh, András Oláh, András Oláh, A.	11 Computer Science ; Pamany Peter Katolikus Budapest Hungary Engineering ; Energy, ... Egyetem
<input type="checkbox"/> Oláh, Andras Bela 3 Oláh, Andras Bela Oláh, A. B.	3 Agricultural and Budapesti Corvinus Budapest Hungary Biological Sciences ; Engineering Egyetem

Display 20 results per page < Page 1 >

2/31. ábra A keresett szerző kiválasztása a találati listából

Scopus Register | Login

Search | Alerts | My list | Settings Live Chat | Help and Contact | Tutorials

The Scopus Author Identifier assigns a unique number to groups of documents written by the same author via an algorithm that matches authorship based on a certain criteria. If a document cannot be confidently matched with an author identifier, it is grouped separately. In this case, you may see more than 1 entry for the same author.

Back to results | 1 of 3 Next > Print | E-mail

Oláh, András About Scopus Author Identifier | View potential author matches
Pecs Tudomanyegyetem, Faculty of Health Sciences, Pecs, Hungary
Author ID: 9243195400
Other name formats: Oláh, Andras, Oláh, András, Oláh, András, View More

Documents: 24 Citations: 132 total citations by 102 documents
h Index: 7 The h Index considers Scopus articles published after 1995.
Co-authors: 109 Subject area: Medicine, Pharmacology, Toxicology and Pharmaceutics View More

View Author Evaluator View citation overview View h-Graph

24 Documents | Cited by 102 documents since 1996 | 109 co-authors

24 documents View in search results format Sort on: Date Cited by ...

Export all Add all to my list Set document alert Set document feed

Economic burden of long-term care of rheumatoid arthritis patients in Hungary	Horváth, Z., Sebestyén, A., Ósterle, A., (...), Bagosi, G., Boncz, I.	2014	The European Journal of Health Economics	0
---	---	------	--	---

Article in Press

View at Publisher

Follow this Author Receive emails when this author publishes new articles

Get citation alerts Add to ORCID Request author detail corrections

Author History
Publication range: 2004 - Present
References: 557
Source history: Canadian Medical Association Journal View documents, Journal of Advanced Nursing View documents

2/32. ábra A szerző tudományos munkájának összegzése

2.6.7. SpringerLink

A kutatók körében igen népszerű az interdiszciplináris SpringerLink adatbázis, melyben a Springer Kiadó gondozásában megjelent könyveiben és folyóirataiban lévő tanulmányok teljes szövege elérhető. A szolgáltatás tudományos minőségét támasztja alá, hogy az adatbázis magas impakt faktorral rendelkező folyóiratokat és tudományos társaságok gondozásában megjelent folyóiratok is tartalmaz. Tárgykörei az alábbi tudományterületekre fókuszálnak: orvostudomány, élettudomány, kémia, földrajztudomány, számítástechnika, fizika és csillagászat, műszaki tudományok, környezetvédelem, jogtudomány, közgazdaságtan, társadalomtudomány. Az adott témától függően angol mellett német, francia, olasz, spanyol és holland nyelvű publikációk is vannak ebben az adatbázisban. A bibliográfiai adatok (tartalomjegyzékek, absztraktok) előfizetés nélkül, a teljes szövegű publikációk csak előfizetéssel vagy a felsőoktatási intézmények könyvtárainak számítógépes hálózatáról érhetők el az EISZ alprogram segítségével.

A keresési lehetőség

A modern kezdőfelületen a Google típusú keresőszolgáltatás révén egyszerű a keresés. A felhasználónak a keresősávba kell begépelnie a kereső-kifejezést/kifejezéseket. A kifejezések között a Boole Operatorok használatosak. A pontos kifejezés keresése során célszerű az exact (") jelet használni. A kapott találati lista mellett lévő bal oldali oszlopban (*Refine Your Search*) a dokumentum típusa (*Content Type*), a cikk (*Article*), könyvfejezet (*Chapter*), kézikönyv (*Reference Work Entry*), protokoll (*Protocol*) és könyv (*Book*), 24 tematikus besorolás szerint a tudományág (*Discipline*), a tudományágakon belüli szűkebb szakterületek (*Subdiscipline*), a forrás-dokumentumok (*Published In*) és nyelv (*Language*) szerint a találatokat tovább lehet szűkíteni (2/33. ábra és 2/34. ábra).

A böngészés során a nyitólap bal oldalán a „Browse by discipline” címszó alatt 24 felsorolt tudományterület között lehet választani, melynek további szűkítése is adott. A kiadványokhoz tárgymutató, bibliográfia és bizonyos esetekben szerzői index tartozik.

Találatok megjelenítése

A találati eredményeknél az adott dokumentum típusa és címe olvasható. Könyv esetében a könyv címe és megjelenés ideje, a publikációk esetében a szerzők és a folyóirat neve kerül feltüntetésre. A találati listában vannak olyan dokumentumok, amelyek csak a megvásárolhatóak, de a Springer Kiadócsoport könyveiben és folyóirataiban megjelent

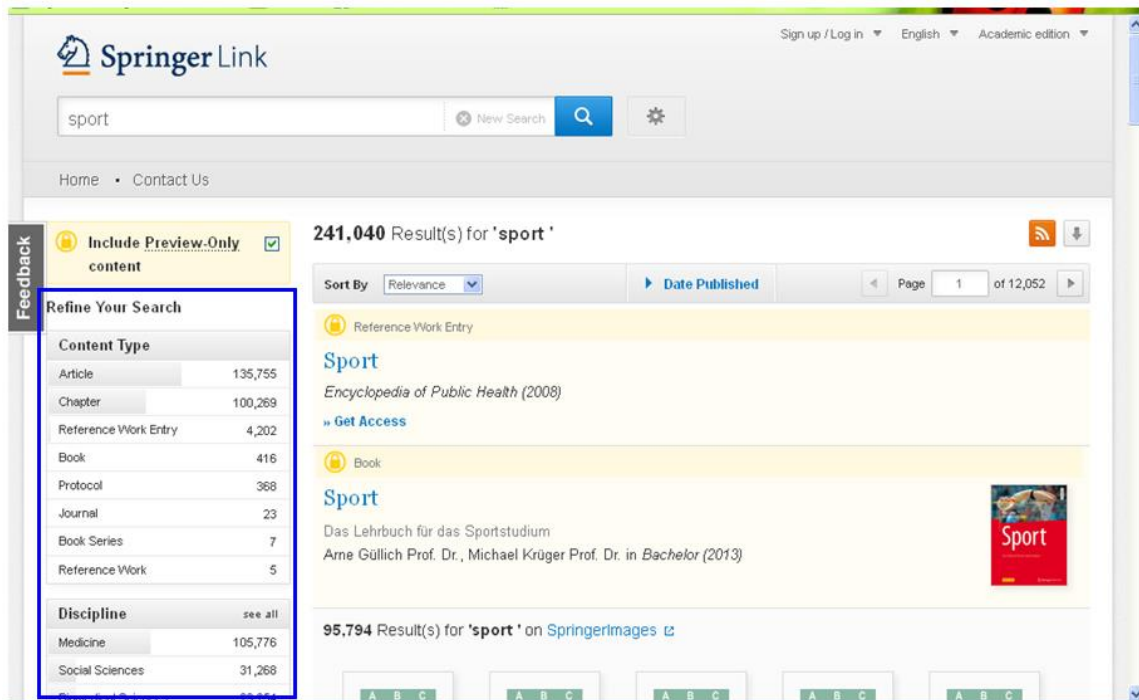
tanulmányok teljes szövege PDF formában letölthető (*Download PDF*) és olvasható (2/34. ábra és 2/35. ábra).

Hozzáférés:

A SpringerLink adatbázis <http://link.springer.com> oldalon érhető el.



2/33. ábra Keresőfelület



2/34. ábra Találati lista

135,755 Result(s) for 'sport'
within Article

Sort By: Relevance | Date Published | Page 1 of 6,788

Content Type: Article

Discipline: see all

Medicine	71,965
Biomedical Sciences	23,786
Life Sciences	19,103
Social Sciences	12,493
Psychology	10,280

Subdiscipline: see all

Orthopedics	23,731
Internal	21,403
Surgery	10,687
Human Physiology	9,210
Rheumatology	5,758

Published In: see all

Knee Surgery, Sports Traumatology, Arthroscopy	4,238
European Journal of Applied Physiology	3,856

Article: **Sport voor de huisartsboekbespreking huisarts sport**
Dit is een informatief en goed leesbaar boekje voor huisartsen die tijdens het spreekuur niet goed raad weten met problemen en vragen over sport. Volgens de auteurs is het boekje bedoeld voor de in sport geïntere...
Sjoerd Hobma in *Huisarts en Wetenschap* (2002)
» Get Access

Article: **Der neue Range Rover Sport: Technische Daten**
ATZextra (2013)
» Get Access

Article: **Josef Hackforth an der TU München Ordinarius für Sport, Medien und Kommunikation**
Michael Schaffrath in *Publizistik* (2001)
» Download PDF (215 KB)

Article: **Bericht vom SFMS (Société Française de Médecine du Sport)-**

2/35. ábra Találati lista további szűkítéssel

2.7. Szakadatbázis

2.7.1. MEDLINE

A Medline a National Library of Medicine által előállított legismertebb, legnagyobb és szakmailag irányadó orvosi információkat nyújtó bibliográfiai adatbázisok közé tartozik. Témakörei több mint 4800 tudományos folyóiraton alapulnak, melyek magukba foglalják az orvostan (*medicine*), betegápolás (*nursing*), fogászat (*dentistry*), állatorvostan (*veterinary medicine*), egészségügyi rendszer (*allied health*), pre-klinikai tudományok (*pre-clinical sciences*) területeit. Kiemelendő, hogy 13 magyar folyóiratot indexel, ezek közül négy az *Acta Veterinaria Hungarica*, *Magyar Onkológia*, *Orvosi Hetilap*, *Pathology Oncology Research* teljes szöveggel is olvasható. A tudományos igényű közleményeket 1965-től dolgozza fel. Jelenleg az indexelt folyóiratok negyede teljes szöveggel is megtalálható. Az adatbázis a MeSH (Medical Subject Headings-orvosi tárgyszórendszer) indexelését alkalmazza, melynek előnye, hogy nem veszik el az azonos tartalmú, de más kifejezéseket tartalmazó publikáció sem. Bibliográfiai adatokat az alábbi források nyújtanak: Index Medicus, International Nursing Index, Index to Dental Literature, PreMedline, AIDSline, BioethicsLine, HealthSTAR. Az EBSCO Publishing, EBSCOhost nevű szolgáltatása révén a Pécsi Tudományegyetem számítógépes hálózatáról ingyenesen elérhető, de az OVID adatbázisban is megtalálható. Az adatbázisban való keresés megegyezik az EBSCOhost adatbázisnál leírtakkal.

2.7.2. PubMed

Az Amerikai Egyesült Államokban működő National Library of Medicine hozta létre a világ legelterjedtebb, a kutatók között is népszerű orvostudomány és határterületeit feldolgozó bibliográfiai adatbázisát, amely mindenki számára ingyenesen hozzáférhető. A szolgáltató az 1950-es évekig visszamenőleg feldolgozott publikációk bibliográfiai adatai mellett jelentős részében csak absztraktot közöl, teljes szöveg (*Full Text*) olvasására csak az egyetemi/orvosi könyvtárak IP tartományán belül kínál lehetőséget.

Keresési módok

A PubMed keresőfelületére való belépést követően az egyszerű és összetett keresésre van lehetőség angol nyelven. Az egyszerű keresés során az egysoros keresőmezőbe a keresett szerző neve(i), a publikáció címe, a keresett folyóirat címe adható meg. A keresés során egy ún. betűzés ellenőrző kiegészítő funkcióval (*Spell Checking Feature*) is segítik a kutató munkáját úgy, hogy a keresődoboz alatt egy legördülő listában a megadott kereső-kifejezéshez tartozó egyéb betűzési variációt is javasol, ezáltal a gépelési hibák kiküszöbölhetőek. A keresőmező mellett lévő görgető menüből lehetőség van a NCBI (National Center for Biotechnology Information) hatáskörébe tartozó valamelyik adatbázist kiválasztani. A keresést a *Search* gombra klikkelve indítható el (2/36. ábra).

Az összetett keresést a keresődoboz alatt található *Advanced* gombbal indítható. A keresőkérdéseket egymás alatt megjelenő keresőmezőbe lehet beírni. A keresés megadható szerző, könyv, MeSH tárgyszó, folyóirat stb. szerint. A kereső-kifejezéseket összekötő Boole-operátorokat a mezők mellett megtalálható görgető menüből lehet kiválasztani, majd a *Search* gombbal a keresés indítható.

A keresőmező alatt található Search History összefoglaló táblázat segítségével mutatja meg, hogy az adott kereső-kifejezések megadása mellett, mikor, mennyi találat volt.

Találatok megjelenítése

A találati listában a megjelent adatok az alábbi információt nyújtják a felhasználó részére: cikk címe, szerző(k) neve(i), a megjelenéssel kapcsolatos adatok. Miután a felhasználó kiválasztotta az érdeklődésének megfelelő publikációt, rákattint a címre és megjelenik a cikk absztraktja. A teljes szöveg elérést a jobb oldali oszlopban megjelenő *Full text links* is mutatja. Az elérhető publikációkat a felhasználó ki tudja nyomtatni, elmentheti, és e-mailen is elküldheti (2/37. ábra).

Hozzáférés

Az adatbázis a <http://www.ncbi.nlm.nih.gov/pubmed> honlapon érhető el.

NCBI Resources How To Sign in to NCBI

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed

PubMed comprises more than 24 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.

PubMed Commons

Featured comment - Aug 14
Author K Asehounne contends comparison of #ClinicalTrials for steroids in traumatic brain injury leave open questions 1.usa.gov/1osgdFU

Using PubMed
PubMed Quick Start Guide
Full Text Articles
PubMed FAQs
PubMed Tutorials
New and Noteworthy

PubMed Tools
PubMed Mobile
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
Topic-Specific Queries

More Resources
MeSH Database
Journals in NCBI Databases
Clinical Trials
E-Utilities (API)
LinkOut

You are here: NCBI > Literature > PubMed

PDF Architect - [apolas_magyar]

Write to the Help Desk

2/36. ábra Keresőfelület

NCBI Resources How To Sign in to NCBI

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed

chronic stress

RSS Save search Advanced Help

Show additional filters

Display Settings: Summary, 20 per page, Sorted by Recently Added

Send to: Filters: Manage Filters

Article types
Clinical Trial
Review
More ...

Text availability
Abstract
Free full text
Full text

PubMed Commons
Reader comments

Publication dates
5 years
10 years
Custom range...

Species
Humans
Other Animals

Clear all

Show additional filters

Results: 1 to 20 of 45318

1. [High-Salt Intake Induces Cardiomyocyte Hypertrophy in Rats in Response to Local Angiotensin II Type 1 Receptor Activation.](#)
Katayama IA, Pereira RC, Dopona EP, Shimizu MH, Furukawa LN, Oliveira IB, Heilmann JC.
J Nutr. 2014 Aug 13. pii: jn.114.192054. [Epub ahead of print]
PMID: 25122644 [PubMed - as supplied by publisher]

2. [The Low-Expression Variant of Fatty Acid-Binding Protein 4 Favors Reduced Manifestations of Atherosclerotic Disease and Increased Plaque Stability.](#)
Saksi J, Ijäs P, Mäyränpää M, Nuotio K, Isoviita PM, Tuimala J, Lehtonen-Smeds E, Kaste M, Jula A, Sinisalo J, Nieminen MS, Lokki ML, Perola M, Havulinna AS, Salomaa V, Kettunen J, Jauhainen M, Kovanen PT, Lindsberg PJ.
Circ Cardiovasc Genet. 2014 Aug 13. pii: CIRCGENETICS.113.000499. [Epub ahead of print]
PMID: 25122052 [PubMed - as supplied by publisher]

3. [Epigenome: Biosensor of Cumulative Exposure to Chemical and Nonchemical Stressors Related to Environmental Justice.](#)
Olden K, Lin YS, Gruber D, Sonawane B.
Am J Public Health. 2014 Aug 14:e1-e6. [Epub ahead of print]
PMID: 25122010 [PubMed - as supplied by publisher]

New feature
Try the new Display Settings option - Sort by Relevance

Results by year

Related searches
chronic stress depression
chronic stress hippocampus
acute chronic stress
cortisol chronic stress
chronic stress immune

PMC Images search for chronic

2/37. ábra. Találati lista

2.7.3. SPORTDiscus

Az EBSCOhost platformján elérhető el a sportorvostannal és rehabilitációval foglalkozó szakirodalmak összefoglaló adatbázisa, a SPORTDiscus. Az adatbázis által indexelt angol, francia, spanyol, olasz nyelven megjelenő folyóiratok publikációinak jelentős részének teljes szövege olvasható az alábbi témakörökben, mint például: a biomechanika, sportorvostan, testmozgás, kineziológia, sporttal, illetve testmozgással kapcsolatos pszichológia, táplálkozás, egészségüggyel és terápiás programok, fizikai fitness, fizikai terápia és rehabilitáció. Jelenleg a PTE Könyvtárából ez a szolgáltatás nem érhető el.

Keresési lehetőség

Az EBSCOhost kereső felületéről érhető el az adatbázis. A nyelvkiválasztás után az egyszerű vagy összetett keresés az EBSCOhost adatbázisnál leírtakkal megegyezően történik.

2.7.4. Semmelweis Egyetem Testnevelési és Sporttudományi Kar Digitális Könyvtár

A Semmelweis Egyetem Testnevelési és Sporttudományi Kar Digitális Könyvtár gyűjteménye a testnevelési és sporttudományi területen megjelenő releváns szakirodalmak közzétételére fókuszál. A feldolgozott szakirodalmat a jobb áttekinthetőség céljából gyűjteményekbe rendezi, melyben folyóiratcikkek, könyvek, konferencia kiadványok, PhD-értekezések, régiségek (sportlapok), válogatott TF szakdolgozatok is olvashatóak, emellett sporttudománnyal kapcsolatos videókat és képeket is tartalmaz. A Magyar Testnevelési Egyetem feldolgozott folyóiratait, mint a Kalokagathia közleményei és a Magyar Sporttudományi Szemle teljes szöveggel is elérhetőek digitális könyvtár jóvoltából.

Keresési lehetőségek

Alkalmazása során a keresés egyszerű és összetett módon történhet. Az egyszerű keresés a keresődobozba beírt kulcsszavak megadásával kezdődik. A keresés során alkalmazni lehet a csonkolást, a Boole-operátorokat, a zárójelet és proximity operátorokat. Összetett keresés történhet cím, szerző, tárgy, földrajzi fogalom, csak metaadat, és teljes szöveg szerint. Emellett a dátum intervallum, a gyűjtemények köre és a média típusa (kép, szöveg, hang, videó) is meghatározható (2/38. ábra).

Találatok megjelenítése

A találatok megjelenítése *Rövid formátum*, *Táblázatos megjelenítés* vagy *Teljes megjelenítés* formátumban tekinthető át. A rövid- és táblázatos formátumban jól látható a releváns találatok

címe, a szerző(k) neve(i), és a PDF ikon. A címre kattintva a szerző mellett a tárgyszavakat, földrajzi fogalmat, megjelenés helyét és idejét, a szakirodalom terjedelmét és a kapcsolódó gyűjteményt tárja az érdeklődők elé (2/39. ábra).

Hozzáférés

Semmelweis Egyetem Testnevelési és Sporttudományi Kar Digitális Könyvtár a <http://tf.hu/oktatas/konyvtar/tf-digitalis-konyvtar/digitalis-dokumentumok/> oldalon keresztül érhető el.



2/38. ábra Keresőfelület

2/39. ábra Találati lista

2.8. Elektronikus könyvek, elektronikus folyóiratok

Napjainkban a kutató munkát nagymértékben elősegítik az online módon elérhető elektronikus folyóiratok, melynek száma évről évre nő. Alkalmazásának számos előnye van, többek között a nyomdai munkálatok kihagyása miatt a publikáció átfutási ideje lényegesen lerövidül; a hozzáférés helytől és időtől független jóval gyorsabb és olcsóbb a számítógépes hálózaton keresztül, így bárhol és bármikor olvasható, szerkezeti felépítése megegyezik a papíralapú változattal; a kutató által talált dokumentum letölthető, elmenthető és nyomtatható is. Az e-folyóiratok egy része ingyen elérhető regisztrációval vagy anélkül. Sok esetben azonban csak előfizetéssel lehet a kívánt folyóiratot megtekinteni.

2.8.1. Open Access folyóiratok

Az elmúlt évtizedben egyre komolyabb szerepet kapnak az információ-szabadságot támogató, számtalan pozitív értékkel rendelkező, a világhálón nyílt hozzáféréssel (*Open Access*) elérhető teljes szövegű (*Full Text*) tudományos publikációk, dokumentumok.

A nyílt hozzáférés lehetővé teszi azt, hogy a releváns szakirodalom mind a szakemberek mind a laikus érdeklődők számára térben és időben korlátlanul, kényelmesen, egyenlő eséllyel elérhetővé váljon. További előnye, hogy hozzájárul a hivatkozások számának emelkedéséhez, mivel minél többen látnak és olvasnak egy adott publikációt a világhálón, annál nagyobb esélye van annak felhasználására is. Nemcsak a hivatkozások számát emeli, de mellette a

kutatók és az intézmények ismertsége is növeli. Mivel az elérése mindenki számára ingyenes, a felhasználót nem terheli többletköltség. Az elérést az is segíti, hogy megfelelő keresési feltételek mellett egyszerű keresőkkel, mint a Google, gyorsan megtalálhatóak. A digitális dokumentumokat a felhasználók elmenthetik, kinyomtathatják és e-mailen is elküldhetik. Sajnálatos módon az emelkedő előfizetési költségek miatt a könyvtárakban egyre kevesebb folyóirat áll rendelkezésre a kutatók számára, amit folyóirat krízisnek is neveznek. A kialakult információhiány megoldásának egyik lehetséges útja a nyílt hozzáférés elterjedése. Szakmai oldalát tekintve a nemzetközi és multidiszciplináris együttműködések is hatékonyan fokozná, így a szegényebb országok is könnyebben juthatnak releváns, bizonyítékokon alapuló kutatások eredményeihez. A gyors elérhetőség hozzájárul a kutatás eredményességéhez is, így tudományos viták is könnyebben kezdeményezhetők, ezáltal a kutatási problémák megoldása gyorsabb és hatékonyabbá válik. Nem szabad arról sem megfeledkezni, hogy a szerzői jogok ebben az esetben ugyanúgy érvényesek. Ma már vannak olyan licencek, amely a szerzők számára biztosítják, hogy a használati jogok határát ő maguk határozzák meg. Végezetül kiemelendő, hogy a folyóirat cikk „preprint” (nyomtatott formában történt megjelenés előtt) előtt megjelenő nyílt hozzáférése akár az elsőbbségi közlésekben is fontos szerepet játszik, így a szabadalmaztatási kérdésekben is könnyebben eldönthető az elsőbbség kérdése. Az Open Access dokumentumait a digitális könyvtárak biztosítják az olvasók számára (Molnár – Németh 2009; Bánhegyi 2003, Bánhegyi 2009).

2.8.2. Open Access Folyóiratok Könyvtára (Directory of Open Access Journals)

Az Open Access Folyóiratok Könyvtára (*Directory of Open Access Journals*) különböző tudományterületek Open Access szakmai folyóiratait gyűjti össze és biztosítja a szabad hozzáférést. Az adatbázisban a folyóiratok az angol nyelv mellett spanyol, portugál, francia, német, olasz, orosz, török, japán, kínai, de még maláj nyelven is olvashatóak.

A felhasználók számára az adatbázisban egyszerű- és összetett keresésre van mód. Egyszerű keresés alkalmával a keresőmezőbe kell a kereső-kifejezést beírni, majd a *Search* gombbal továbblépni.

Az összetett keresésben cím (*Title*), kulcsszó (*Keywords*), tárgy (*Subject*), ISSN, DOI, ország (*Journal Country*), nyelv (*Journal Language*), kiadó (*Publisher*), absztrakt (*Abstract*), szerző (*Author*), év (*Year*), újság cím (*Journal title*), újság alternatív cím (*Journal Alternative Title*) alapján lehet megadni a kereső-kifejezéseket.

Az összekapcsolására Boole-operátorokat lehet használni. Az eredménylista különböző módon rendezhető jelentőség (*Relevance*), a Doaj-hoz való hozzáadás dátuma (*Date added to*

Doaj, cím (*Title*), cikk (*Article*) publikálás dátuma (*Publication date*) alapján. A találati halmaz finomítására itt is van lehetőség. A szűkítés történhet folyóirat vagy publikáció szerint, továbbá tárgykör (*Subject*), a folyóirat nyelve (*Journal Language*), a folyóiratot kiadó ország neve (*Journal Country*), Kiadó neve (*Publischer*), közzétételi költség (*Publication charge*), folyóirat licenz (*Journal License*), a publikáció megjelenés éve (*Year of publication-Articles*) és a folyóirat neve (*Journal title-Articles*) alapján.

A találati lista valamennyi tételét a felhasználók szabadon olvashatja, de letöltheti vagy ki is nyomtathatja (2/41. és 2/42. ábra).

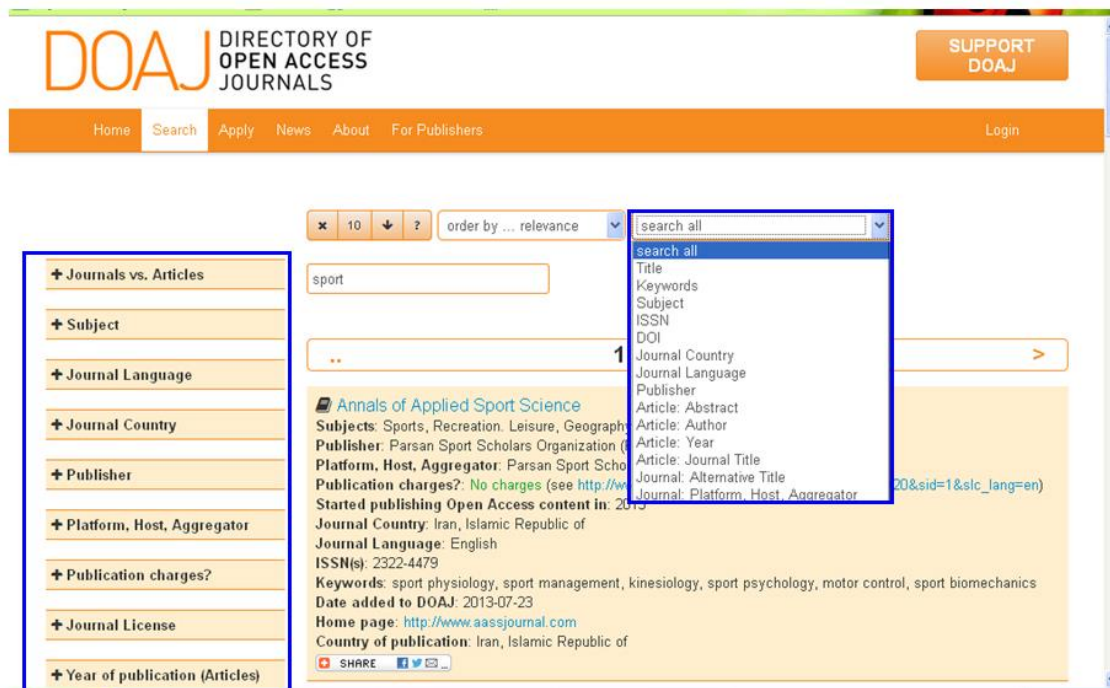
Hozzáférés

A DOAJ a <http://www.doaj.org> címen érhető el.



The screenshot shows the DOAJ website's search interface. At the top left is the DOAJ logo and the text 'DIRECTORY OF OPEN ACCESS JOURNALS'. To the right is a 'SUPPORT DOAJ' button. Below this is a navigation bar with links for 'Home', 'Search', 'Apply', 'News', 'About', 'For Publishers', and 'Login'. The main content area features a search box with the placeholder 'Search DOAJ' and a search icon. Below the search box are checkboxes for 'journals' and 'articles', and a link for '[Advanced Search]'. To the right of the search box is a statistics box showing: '9,958 Journals', '5,844 searchable at Article level', '134 Countries', and '1,705,751 Articles'. Below the search box is a section titled 'Directory of Open Access Journals (DOAJ)' with a brief description and a link to the 'Application Form'. To the right of this section is a box with links for 'FAQs', 'Features', 'Open Access Information', 'Download metadata', and 'New Journals Feed'. Below the 'Directory of Open Access Journals (DOAJ)' section is a 'Latest News' section with a sub-heading 'Proactive not reactive' and a short news snippet. To the right of the 'Latest News' section is a box with links for 'Our sponsors', 'Our members', and 'Our publisher members'. At the bottom right of the page are social media icons for Facebook, Twitter, and LinkedIn. At the bottom left of the page is a small URL: 'www.linkedin.com/company/directory-of-open-access-journals-doaj-'.

2/40. ábra Keresőfelület



2/41. ábra. A keresési lehetőségek

2.9. Internetes keresőrendszerek használata a világhálón

A világhálón található, különböző dokumentációk feltárása céljából használt keresőgépek a kutató munkát végzők körében is igen népszerű. A keresőgépek segítségével különböző dokumentumok, képek, programok egyaránt megtalálhatóak, hiszen az egész világot átszövő számítógép hálózaton történik a keresés. A kulcsszavas keresés történhet egyszerű vagy összetett keresés formájában. Kedvelt keresőgépek közé tartozik az angol eredetű, de magyar keresőfelülettel rendelkező Google (<http://google.com>). Emellett portálszolgáltatásokkal kibővülő keresőgépek a teljesség igénye nélkül: Excite (<http://excite.com>), Yahoo (<http://yahoo.com>), Microsoft Internet Start (<http://msn.com>), HotBot (<http://hotbot.com>), Lycos (<http://lycos.com>). Felhívjuk a figyelmet arra, hogy a weben keresztül hatalmas információ mennyiség érhető el, ezért meglehetősen nehéz kiválasztani a valódi, releváns szakirodalmakat.

2.10. Idézetek és szövegek közötti hivatkozások

Egy tudományos igényű dolgozat elkészítése során más szerző(k) gondolata, releváns munkája is egyértelműen feldolgozásra kerül. Alapvető elvárásként jelenik meg, hogy ezekre a szószerinti vagy szabadon megfogalmazott idézetekre hivatkozni kell. A tudományos pontosság és tisztasága megőrzése és a szerzői jogok megsértésének elkerülése miatt a felhasznált forrás eredetét mindig pontosan meg kell adni. Plágiumnak minősül más szerző

szavainak, gondolatainak, kutatási eredményeinek saját gondolatként történő felhasználása. Hivatkozást akkor érdemes alkalmazni, ha alátámasztja a kutató állítását vagy valamilyen ötlet, elképzelés megcáfolása miatt szükséges a dolgozatba beépíteni, vagy olyan fontos adatot, megállapítást tartalmaz, amely nélkülözhetetlen a téma feldolgozása szempontjából, vagyis fel lehet használni más ember szavait, gondolatait, kutatási eredményeit, de pontosan leírt bibliográfiai adatokkal hivatkozni kell a felhasznált információ forrására. A hivatkozás másik alapelve, hogy a legyen egyértelmű és a felhasznált irodalom megbízható forrásból származzon. A dolgozat készítése során mindig törekedni kell arra, hogy a felhasznált idézet elsődleges forrásból származzon és nem szabad átvenni ellenőrzés nélkül mások által történt hivatkozásokat. A szakirodalmak feldolgozásának munkafolyamatába érdemes egy listát vezetni a felhasznált szakirodalmakból úgy, hogy pontos feljegyzést készít a mű bibliográfiai adatairól. Ez a jegyzet az irodalomjegyzék/bibliográfia készítésénél a későbbiekben nagy segítséget fog nyújtani, és a visszakeresést is nagymértékben megkönnyíti. A források megadása történhet lábjegyzetbe vagy a szöveg végén a felhasznált irodalom/bibliográfia címszó alatt.

A különböző forrásokra történő hivatkozás három alaptípusa ismeretes:

- **Szó szerinti idézés** esetében idézőjelbe kell tenni a szerzőtől átvett szövegrészt. A zárójelbe történő hivatkozásban fel kell tüntetni a szerző nevét, a publikáció megjelenési évét, és az oldalszámot. Meg kell jegyezni, ha három vagy több szerzője van az adott szakirodalomnak, csak az első szerzőt kell kiírni, a társszerzőket az „et al” (jelentése: és mások) vagy „és mtsai” rövidítéssel jelölik.
 - **Példa:** Selye professzor a következőképpen határozta meg a stressz fogalmát:
„a stressz egy fajlagos tünetcsoportban megnyilvánuló állapot, mely magába foglal minden nem fajlagosan elő idézett elváltozást egy biológiai rendszeren belül.”...(Selye 1936, 32-45.)
- A **parafrázis** (tartalmi idézés) azt jelenti, hogy egy adott szerzőtől átvett gondolatot a dolgozat készítője saját szavaival írja le. Ebben az esetben arra is ügyelni kell, hogy a dolgozat írójának gondolatmenete pontosan kövesse a másik szerző gondolatait. A zárójelben történő hivatkozás - ebben az esetben is - a szerző nevét, a publikáció megjelenési évét tartalmazza, az oldalszám megadása azonban nem előírás.
- **Közvetett/kereszthivatkozás** alatt értendő, ha az eredeti publikáció nem került elolvasásra, de egy másik szerző publikációjában arra a műre hivatkozás történt. Ilyenkor a hivatkozást úgy kell korrekt módon megadni zárójelben, hogy a forrás

szerző adatai után fel kell tüntetni, hogy mely szerző művében található a felhasznált idézett szövege.

- **Példa:** Dr. Karsai István egy - a VIII. Országos Sporttudományi Kongresszuson tartott - előadásában hangsúlyozta, hogy ... (Karsai 2012)

A hivatkozásnak formai szabályai vannak, közlésének két leggyakrabban alkalmazott módszere ismeretes.

A számozásos módszernél a felhasznált forrásokat első előfordulásuk szerint kell sorrendbe állítani. A "név-év" módszer (Harvard szisztéma) esetében a szerző neve és a forrás megjelenési évét kell feltüntetni. A szerzők betűrendi sorrendjében összeállított irodalomjegyzék sorszámozását ebben az esetben nem kell figyelembe venni.

Az irodalomjegyzék/bibliográfia készítése a tudományos munka során

A tudományos dolgozat készítése során a felhasznált és hivatkozott releváns szakirodalom adatait a dolgozat főszövege után egy összesítő jegyzékbe kell megadni. A magyar terminológia elkülöníti egymástól az irodalomjegyzék és a bibliográfia fogalmát. Az irodalomjegyzék (felhasznált irodalom) a ténylegesen elolvasott, áttekintett releváns szakirodalmat tartalmazza, függetlenül attól, hogy a dolgozatban történ-e rá utalás. Míg a bibliográfia a kutatás témájához kapcsolódó alapvető és nélkülözhetetlen irodalmat rendszerezi. Elkészítésük célja nemcsak az azonosíthatóság, de ezáltal biztosított a visszakereshetőség is. A bibliográfiai tétel a felhasznált dokumentumok, mint a hivatkozások, a jegyzetek és a bibliográfiák alapegységeként értendő. A bibliográfiai hivatkozások formáját szabvány írja elő:

- Magyar szabvány: MSZ ISO 690 (1991-től),
- Nemzetközi szabványok, mint például APA (American Psychological Association), Harvard British Standard, AMA (American Medical Association), Vancouver/ICMJE.

Felhívjuk a figyelmet arra, hogy felsőoktatási intézményenként, intézetenként, szakmai folyóiratokként az irodalomjegyzékkel kapcsolatos formai követelmények szabványa eltérhet egymástól. Ezeket mindig alaposan és körültekintően kell áttanulmányozni, mert pontosan a hivatkozásokkal és az irodalomjegyzék készítésével kapcsolatban pontosan meghatározzák az elvárásaikat. Ezek eltérhetnek az alábbiakban tárgyalt, de általánosságban elfogadott hivatkozási normáktól.

A bibliográfiai tételekkel szemben támasztott általános előírások

- A felsorolt szakirodalmakat alfabetikus sorrendben kell rakni és minden egyes tételt külön sorba kell írni.
- A betűrendi sorrend felállítása az első szerző vezetékneve alapján történik.
- Magyar szerzők esetében a vezetéknev megadását követően (ebben az esetben nem kell vessző), a keresztnév első betűje következik, ami után egy pontot kell írni.

Példa:

- Ács P. (2009): A sportolók területi mozgásai, avagy a sportolói vándorlás. Tér és társadalom. 23(3),147.
- A szerzők sorrendje a hivatkozott mű címlapja szerint kerüljön megadásra.
- Nem minősül bibliográfiai adatnak a tudományos fokozat, egyetemi rang és a társadalmi rangra vonatkozó névelőzések, ezért az irodalomjegyzékben/bibliográfiában ne szerepeljen (dr., prof., PhD., br., gr.,).
- A külföldi szerzőkre vonatkozó írásmód: a szerző vezetékneve, majd vessző és a keresztnév kezdőbetűje, ami után pontot kell tenni.

Példa:

- Baxter, P. – Jack, S. (2008): Qualitative case study methodology: Study design and implementation for novice researchers. The Qualitative Report, 13(4), 544-559.
- Ha a szerző több keresztnévvel rendelkezik, a keresztnévek kezdőbetűje között egy szóközt kell tenni.

Példa:

- Adams, J.M.G. – Tyson, S. (2000): The Effectiveness of Physiotherapy to Enable an Elderly Person to Get up from the Floor. Physiotherapy, 86(4), 185-189.o.
- Magyar szerző idegen nyelven megjelenő művére a külföldi szerzőkre érvényes írásmódot kell alkalmazni, vagyis a szerzővezeték neve és a keresztnév közé vesszőt kell írni.

Példa:

- Oláh, A., Katona, Gy., Gál, N., Müller, A., Damasdi, M., Boncz, I., Betlehem, J. (2012): The comparison of two minimal invasive surgeies, the tension-free vaginal tape (TVT) and the transobturator tape (TOT) in terms of efficiency

and the complications. South Eastern Europe Health Sciences Journal, 2(2), 82-87.

- Szerkesztett könyvek esetében először a szerkesztő vezeték és keresztnéve kerül megadásra, majd zárójelben utalást kell tenni a szerkesztőváltóra (szerk.), (ed.), (eds.).

Példa:

- Oláh A. (szerk.) (2012): Az ápolástudomány tankönyve. Budapest, Medicina Kiadó.

- A többszerzős mű szerzőinek nevét hosszú kötőjellel és szünetek beiktatásával kell megadni. Kettősnév előfordulása során a kötőjelet szünet nélkül kell kitenni.

Példa:

- Pakai A. – Kivés Zs. (2013): Kutatásról ápolóknak. 2. rész: Mintavétel és adatgyűjtési módszerek az egészségtudományi kutatásokban. Nővér, 26 (3), 20-43.

- Ugyanazon szerző(k) munkáját a megjelenés évszáma alapján kell sorrendbe állítani. Ha a szerzőnek több műve is megjelent, akkor azokat a forrás címe alapján kell rendezni és abc betűivel elkülöníteni.

- Többkötetes mű esetén a cím után mindig arab számmal kell megadni a kötetszámot.

- **Példa:** (1. köt.)

- A könyv kiadását tekintve, ha nem az első kiadás kerül felhasználásra, akkor a mű címének megadását követően a kiadás számát is le kell írni.

- **Példa:** 2. átdolgozott kiadás

- Idegen nyelven megjelenő munkák esetén az adott nyelv szabályait kell követni.
- Idegen nyelvű bibliográfiai tételbe nem szabad magyar nyelvű betéteket elhelyezni. A címet a forrásban szereplő nyelven kell megadni, de a magyar nyelvű fordítását szögletes zárójelben meg lehet adni.
- Ha a felhasznált mű nem rendelkezik szerzővel, akkor a zárójelbe foglalva kell a mű címét és a megjelenés évét feltüntetni.
- Az irodalomjegyzékben általános rövidítések is használnak, melyet a 2. táblázat szemléltet.

2/2. táblázat. A hivatkozásokban és irodalomjegyzékben megtalálható általános rövidítések összefoglaló táblázata

kiadás jelölésénél	
átdolgozott	átdolg.
bővített	böv.
edition	ed.
javított	jav.
kiadás	kiad.
kiegészített	kieg.
a megjelenés adatcsoportnál	
Akadémiai Kiadó	Akad. K.
Budapest	Bp.
hely nélkül	h. n.
Kiadó	K.
kiadó nélkül	k. n.
Könyvkiadó	Kvk.
folyóiratnál	
évfolyam	évf.
szám	sz.
volume	vol.
number	no.
terjedelmi adatoknál	
oldal	o. p. (a latin pagina szóból),
oldaltartományok	pp.

A bibliográfia formája a különböző dokumentumtípusok szerinti az alábbiak lehetnek:

Könyv:

Szerző neve (Kiadás évszáma): Teljes cím. Kiadás helye: Kiadó.

Ács P. (2009): Sporttudományi kutatások módszertana. Pécs: Pécsi Tudományegyetem Temészettudományi Kar Testnevelés- és Sporttudományi Intézet.

Háromnál több szerzős könyv esetében:

Első szerző neve és mások/et al. (évszám): Könyv címe. Kiadás helye: Kiadó.

Csermely P. és mások (1999): Kutatás és közlés a természettudományokban. Budapest: Osiris Kiadó.

Könyvfejezet:

Szerző neve (kiadás évszáma): A könyvrészlet, tanulmány vagy fejezet címe. In:

Szerkesztő(k) neve (szerk. vagy ed(s.)): Könyv címe. Kiadás helye: Kiadó, Oldalszám (től-ig)

Staunder A. (2007): Stressz és stresszkezelés. In: Kállai J., Varga J., Oláh A. (szerk.): Egészségpszichológia a gyakorlatban. Budapest: Medicina Könyvkiadó Zrt, 153-176. p.

Szerkesztett könyv:

Szerkesztő(k) neve (szerk. v. ed(s.)) (évszám): Könyv címe. Kiadás helye: Kiadó.

Oláh A.(szerk.)(2012): Az ápolástudomány tankönyve. Budapest: Medicina Kiadó.

Szerző nélküli könyv:

Könyv címe (Kiadás évszáma) Kiadás helye: Kiadó.

Magyar Statisztikai Évkönyv, 2010 (2011) Budapest: KSH.

Folyóirat:

Szerző(k) neve (megjelenés éve): Cikk címe. Folyóirat címe, Évfolyam szám. Kötetszám, Oldalszám (től-ig)

Melczer Cs., Melczer L., Szabados S., Ács P.(2012): Szívelégtelen betegek életminőségét mérő validált kérdőívek összehasonlító vizsgálata. In: Egészség-Akadémia, 3.1.54-60. p.

Külföldi cikk esetében (hasonlóan a magyar nyelvű cikk hivatkozására):

Oláh, A., Katona, Gy., Gál, N., Müller, A., Damasdi, M., Boncz, I., Betlehem, J. (2012): The comparison of two minimal invasive surgeies, the tension-free vaginal tape (TVT) and the transobturator tape (TOT) in terms of efficiency and the complications. In: South Eastern Europe Health Sciences Journal, 2.2. pp.82-87.

Konferencia valamely előadására való hivatkozás:

Előadó(k) neve: Az előadás címe. Helye. Ideje.

Pakai A. (2012): A tudományos közlések módszertani alapjai a táplálkozástudomány területén: Előadás. Budapest, „A tét a jövőnk: A táplálkozás és a mozgás összhangjában”.: A Magyar Dietetikusok Országos Szövetsége XIII. Szakmai Konferenciája. 2012.11.17.

Hivatkozás jogszabályokra:

A jogszabály címe (Kiadás évszáma). Folyóirat címe, Évfolyam szám. Kötetszám, Oldalszám (től-ig)

2010. évi XCII. törvény egyes egészségügyi és szociális tárgyú törvények jogharmonizációs célú módosításáról (2010). In. Egészségügyi Közlöny 60. 20. 2986-2989. p.

Hivatkozás elektronikus forrásokra, dokumentumokra:

Szerző (k) (vagy szerkesztő, szervezet, forrás neve) (évszám): Tanulmány, cím vagy weboldal címe: alcím (ha van). A honlap neve, a honlap címe/URL címe, utolsó letöltés ideje

Fidy, J., Makara, G. (2005). Biostatisztika. Retrieved from:

URL:<http://www.tankonyvtar.hu/hu/tartalom/tkt/biostatisztika-1/ch11.html> {2013.06.05}

Szabvány

A szabványkibocsátó országban használatos betűjel és a szabvány száma. Szabvány címe. Megjelenés helye: Kiadó, Megjelenés éve. Oldalszám.

MSZ ISO 690 (1990) Bibliográfiai hivatkozások. Budapest: Magyar Szabványügyi Testület, 22. p.

3. STATISZTIKAI ALAPFOGALMAK, AZ ADATOK FAJTÁI, VÁLTOZÓK ÉS ISMÉRVVÁLTOZÓK TÍPUSAI (Ács Pongrác)

3.1. A statisztika meghatározása

Ebben a részben a gyakorlati adatelemzések során elengedhetetlen alapfogalmakat fogjuk tisztázni. A statisztika, mint minden más sajátos önálló nyelvezettel rendelkezik. A gyakorlati kutatások során gyűjtött információk feldolgozásának, elemzésének statisztikai alapmódszereit kívánjuk a későbbiekben bemutatni, melyhez egyfajta statisztika alapismeretre van szükségünk. Napjainkban már a kutatások adatainak értékelése, elemzése, elképzelhetetlen a számítógépek használata nélkül - és valljuk meg roppant mód, meg is könnyíti a munkát- ezért az alapfogalmak tisztázását és a módszerek bemutatását az SPSS program nomenklatúra felosztásában fogjuk tárgyalni.

A *statisztika* a tömegesen előforduló jelenségekre, folyamatokra vonatkozó információk összegyűjtésének, leírásának, elemzésének, értékelésének és közlésének tudományos módszertana. A nemzetközi irodalomban elterjedt csoportosítás szerint megkülönböztethetjük a *leíró statisztikát*, a *következtetési statisztikát* és a *statisztikai döntésméleletet*.

A *leíró statisztika* alapvetően a numerikus információk összegyűjtését, az információk összegezését, és tömör jellemzését szolgáló módszereket foglalja magában. A leíró statisztika legfontosabb területei: az adatgyűjtés, az adatok ábrázolása, az adatok csoportosítása, osztályozása, az adatokkal végzett egyszerűbb aritmetikai műveletek, az eredmények megjelenítése. E területen jellemzően egyszerűbb statisztikai módszereket használunk és az alapsokaságra vonatkozó, a vizsgálat szempontjából fontos adatokat maradéktalanul használjuk fel. A leíró statisztikai eljárásokat gyakran az Excel program segítségével számítják, hiszen kezelése könnyű és világos.

A *következtetési statisztika* segítségével a jelenségekre, folyamatokra vonatkozóan olyan megállapításokat tehetünk, amelyek nem csak a közvetlen megfigyelésen alapulnak. Igen leegyszerűsítve, alkalmazásával közvetlenül nem mérhető, csak összetett statisztikai, matematikai eljárásokkal megszerezhető számszerű információkat nyerhetünk. A következtetési statisztika szorosan épít a matematikai- statisztikára és a valószínűségelméletre, fontos, hogy a következtetések itt mindig valamely mintából (részsokaságból) származnak. Két általunk tárgyalandó része: a becslés és a hipotézisellenőrzés.

A *statisztikai döntésmélet* a véletlen környezet által bekövetkező események figyelembevétele mellett, több cselekvési lehetőség közül az optimálisnak vélt kiválasztásához ad számszerű információkat. Az empirikus statisztikai megfigyeléseken, és a következtetéseken túl szubjektív, szakértői értékítéleteknek is teret enged. A valószínűségelmélet és a játékelmélet elemeit kombinálja a statisztikai megfigyelések eredményeivel.

3.2. Statisztikai adatok

A statisztikai adatok legtöbbször mérések útján keletkező eredmények (pl.: a statisztikai sokaság elemeinek száma, vagy valamely számszerűsíthető jellemző).

Szekunder adatnak nevezzük a miénktől leggyakrabban eltérő valaki más által mért, gyűjtött adatokat, míg *primer adat* az adott kutató által történő adatgyűjtés saját kutatási célnak megfelelően.

Megkülönböztetünk *alapadatokat*, melyeket elsődleges mérések és számlálások útján nyerünk, illetve *származtatott adatokat*, melyekhez kettőnél több alapadaton elvégzett művelet eredményeképpen jutunk. Az állandósuló statisztikai származtatott adatokat a statisztikában gyakran *mutatószámokként* is emlegetik (pl.: BMI, népsűrűség).

Az adatok jellegük szerint megkülönböztethetők:

- *megállapítható*, azaz minőségi/*kvalitatív* vagy
- *mérhető*, vagyis mennyiségi/*kvantitatív* adatok.

A két típust az adatfelvétel és a mérhetőség különbözteti meg. Elmondható, hogy a mennyiségi adatokból mindig képezhetők minőségi jellegű adatok (kategóriák, osztályok, rangsorok, stb.), míg fordítva ez nem vagy nagyon nehezen lehetséges. A mennyiségi adatokkal matematikai műveletek könnyen elvégezhetők (pl. átlagolhatók, összevonhatók stb.), a minőségi adatoknál - a csoportok és kategóriák esetében - ez értelmetlen.

Ozsváth és Ács szerint az adatok értékük/értékkészletük szerint lehetnek:

- alternatív vagy bináris (dummy),
- diszkrét,
- folytonos adatok.

A *bináris adatoknál* mindig csak két (tetszőleges) érték fordulhat elő (férfi/nő). Igen gyakori a „0-1” („nem - igen”) érték – bár ez a későbbi számításoknál gondot is okozhat, mivel nullával nem lehet osztani. Vannak olyan statisztikai módszertanok, amelyeknél 1-re való kódolást kell megvalósítani, és egy ilyen dummy változót kell létrehozni.

A **diszkrét adatok** jellemzője, hogy „pontoszerűek”, az értékek között nincs folytonosság, a legtöbb esetben az értékek közötti tartomány nem is értelmezhető. Tipikus példái a rangsorok, darabszámok, évszámok, kategóriába sorolások stb.. A diszkrét változók véges számú ismérvváltozatokkal rendelkeznek (pl.: az elmúlt évben előforduló betegségtípusok vagy gyermekek száma családonként).

A **folytonos adatok** tetszőleges pontossággal megadhatók és bármely két érték közötti tartomány is értelmezhető. A „folytonosság” mérhető adatok értékészletének jellemzője, amelyben végtelen pontossággal mérhetünk (pl.: súly, vérnyomás stb.).

Statisztikában a megfigyelés és mérés tárgyát képező megfigyelési egységeket **egyedek**nek nevezzük. A statisztikai megfigyelések tárgyát képező egyedek összességét statisztikai **sokaságnak** is nevezzük. A statisztikai sokaságra használatos még a populáció elnevezés is. Kiemelt fontossággal bír a statisztikában a sokaság egyedeire vonatkozó tulajdonságok és jellemzők, melyeket **változóknak** vagy **ismérveknek** nevezünk. **Ismérvváltozatoknak** hívjuk az ismérvek vagy változók lehetséges kimeneteit, módozatait (Ozsváth-Ács 2011).

3.3. A változók csoportosítása, skálátípusok

Általánosságban az ismérvek/változók lehetnek: *mennyiségi, minőségi, időbeli, és területi* ismérvek. Léteznek statisztikai módszertanok pl. regresszió- analízis, amelyek megkövetelik a változók más jellegű definiálását is. Ennek megfelelően megkülönböztetünk **függő** és **független** változókat. A függő változót minden esetben a független változó határozza meg, ok és okozat kapcsolat áll fenn közöttük.

Független változók: azok a változók, amelyekről úgy véljük, meghatározó szerepet játszanak a minket érdeklő problémában. Ezen változók értékeit (attribútumait) mi változtatjuk (illetve választjuk) meg.

Függő változók: azok a változók, amelyek „viselkedésére” (eloszlására) kíváncsiak vagyunk. Megfigyeljük, hogy a független változók változtatása hogyan hat a függőváltozókra (így próbáljuk felderíteni, milyen kapcsolat van közöttük). Például vizsgálni szeretnénk az ápolók körében a feladathelyzetben történő szorongást és depressziót. A szorongás és depresszió mértéke lesznek a függő változóink, amit mérnünk kell. A független változók lehetnek azon változók, amiktől ez függhet, például az életkortól, nemtől vagy a feladat nehézségétől. Ezek lesznek a független változók, melyeket nekünk kell módosítani. A feladat nehézségi fokát a lehetőségekhez mérten változtathatjuk, azonban az életkort és nemet változtatni nem tudjuk, csak megválasztani.

A gyakorlati adatelemzéshez, főleg a számítógéppel támogatott feldolgozások esetén a változók mérési szintjét, illetve a mérési skálákat pontosan meg kell határozni.

A mérési skálák típusai:

- nominális (névleges) skála,
- ordinális (sorrendi) skála,
- intervallum skála,
- arányskála.

Nominális (névleges, nominal, categorical) skála a legegyszerűbb skála, mely kevés információt szolgáltat. Az egyedek osztályozása, csoportosítása csak megkülönböztetésre szolgál. A skálán az értelmezhető, hogy a megfigyelési egyedek egyenlők, vagy különbözők és a skálához tartozó értékekkel műveleteket elvégezni (pl. kivonás, osztás) nem lehet. A megfigyelési egységhez rendelt kódokat önkényesen választják. A nominális skála ennek megfelelően kategóriákat, csoportokat foglal magába, számszerűsítése csak kódolásnak tekinthető. Gyakran csak bináris adatokat tartalmaz („két kategória”, „igen-nem”, „egyezik - nem egyezik”). Több kategória esetén az értékkészlete is nagyobb, de ez esetben is mindig diszkrét értékeket képez. Összegezve fontos, hogy a nominális skála értékei egymáshoz nem viszonyíthatók, nem adhatók össze, nem rendezhetők, nem átlagolhatók, nincs „kisebb-nagyobb”, „jobb-gyengébb” stb.. A nominális skála mindig kvalitatív adatokat jelent, és ennek megfelelően soha nem tartalmaz folytonos eloszlású adatokat.

Ordinális (sorrendi, ordinal, ordered categorical) skála megkülönböztet és sorrendet is mutat. A sorrendiségre vonatkozó relációk alapján rangsorba rendezzük a megfigyelt objektumokat, egyedeket. A sorrendi különbséget megállapítja, azonban az egymástól mért pozíciókat (pl. mennyivel jobb az egyik beavatkozás hatásfoka, mint a másik) nem mutatja. Az ordinális skála is mindig diszkrét adatokat tartalmaz és alapjaiban kvalitatív jellegű adatokat jelent.¹

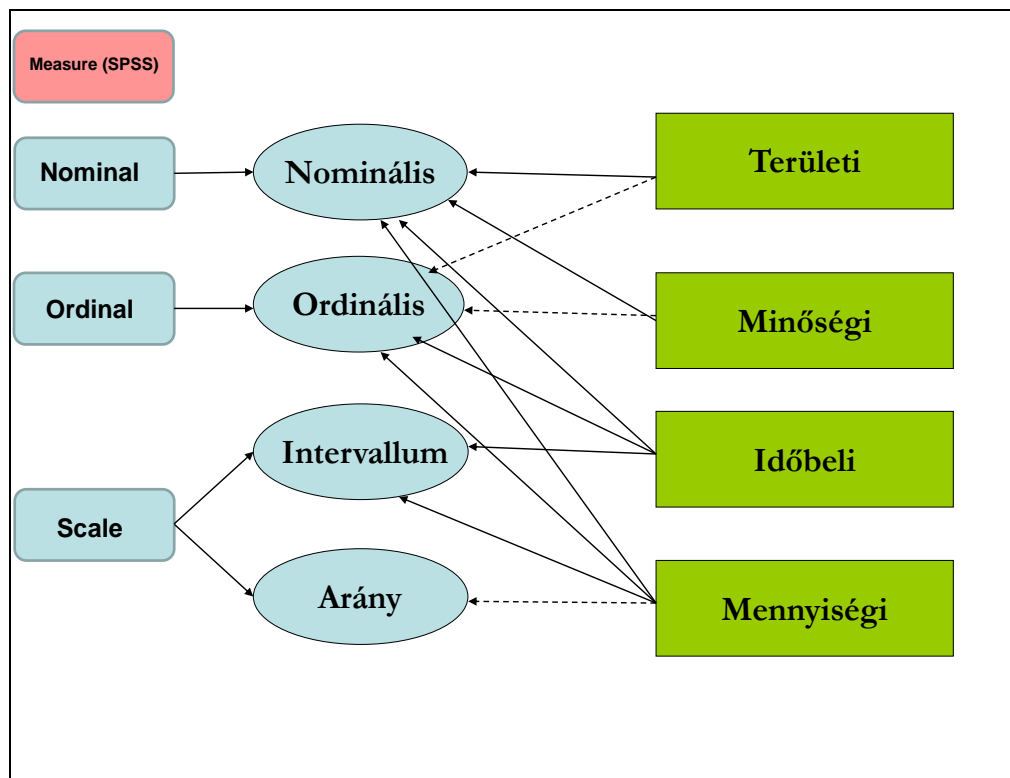
Intervallum (interval) skálát metrikus skálának is nevezik. A skála alkalmas arra, hogy mérje, a különbséget a két érték között, így megmondhatja, hogy az egyik mennyivel nagyobb, jobb, vagy szebb a másikinál. Az intervallum-skálának nagyon jellegzetes

¹ Alapjaiban tipikus ordinális skála a kérdőíves módszereknél gyakran alkalmazott, attitűd vizsgálatokhoz kidolgozott, eredetileg ötfokozatú Likert-skála. Néhány fokozatú terjedelme és diszkrét értékei miatt „alapjaiban” kvalitatív jellegű. De ez a skála lehetne százas, ezres vagy még nagyobb terjedelmű – amitől persze a „pontossága” nem feltétlenül javulna. Az adatfeldolgozás során hasonló esetekben általában már a paraméteres eljárásokat alkalmazzák, lényegében kvantitatívnak – és egyszerűen „csak” osztályba soroltnak, így intervallumskálán elhelyezkedőnek – tekintve a felmérési eredményeket.

tulajdonsága, hogy nem rendelkezik igazi zéró ponttal, vagyis a zéró érték nem jelenti a tulajdonság hiányát.

Arányskála (proportional), szolgáltatja a legtöbb információt, legmagasabb mérési szintet. Bármely két skálaérték arány értelmezhető. A skálának van igazi zéró pontja, ami azt jelenti, hogy a nulla érték a tulajdonság hiányát egyértelműen jelzi. A skálán bármely matematikai művelet elvégezhető. Az összes használatos mértékegységünk arányskálát képez. A mérési pontosság kizárólag technikai kérdés. Használatánál mindössze arra kell ügyelni, hogy az azonos jellegű, de különböző dimenziójú mértékegységek egymástól eltérő számrendszerűek lehetnek.

A két metrikus (intervallum - és arányskála) skála között különbség a zérus meghatározásából áll. Hiszen az intervallum skálán a zérus meghatározás önkényes.



3/1. ábra: Az ismervek és mérési skálák összefüggése és azok jelölése az SPSS programban

Forrás: Pintér – Rappai 2007, 31. o. nyomán saját szerkesztés

Fontos megjegyezni, hogy a legfejlettebb az arányskála, melyet az elvégezhető aritmetikai műveleteknek a száma is mutat. Belőlük bármely skálatípus transzformálható, előállítható. Minél fejlettebb egy skála, annál részletesebb elemzés, összehasonlítás elvégzésére alkalmas. „A skálatípusok azonosítása nagyon fontos, ugyanis egyértelműen meghatározza, hogy milyen elemzéstípusokat tudunk végezni, tehát az, hogy független vagy függő változók vannak metrikus vagy nem metrikus skálán mérve, jelentős eltéréseket okozhat.” (Sajtos – Mitev 2007, 25. o.)

4. ONLINE KÉRDŐÍVEK SZERKESZTÉSE A GYAKORLATBAN (Raposa László Bence)

4.1. Bevezetés

A kérdőív (*questionnaire*) több tudományterület (főleg a társadalomtudományok) adatgyűjtésére alkalmazható eszköze. A tudományos vizsgálatok mintáit illetve adatait az experimentális, analitikus, mért adatokon kívül a kérdőívek kérdéseire adott válaszok adhatják. A kérdőív fő szerepe a kitöltők által adott válaszok rögzítése, leginkább strukturált, előre meghatározott formában, adott kérdéssorok mentén. A kérdőíveknek számos formája, verziója alakult ki a telekommunikáció fejlődésével, így megkülönböztetünk postai kérdőívet, telefonos kérdőívet, online kérdőívet és személyes megkérdezést (interjú, mélyinterjú) alkalmazó kérdőívet. A személyes kérdőívek is többféleképpen jelenhetnek meg, így papíron (*P2P - paper-and-pencil*) vagy számítógépen (*CAPI - Computer-Aided-Personal-Interview*) (Lógó 2007). Az egyes kérdőívformák alkalmazhatóságának kritériumai változóak és alkalmazhatóságuk is eltérő, így fontos, hogy az adatgyűjtés céljának megfelelő formát válasszuk (Bauer és mtsai 2007).

4/1. táblázat: A különböző kérdőívek alkalmazhatóságának feltételei

	Megkérdezés típusa			
	Szóbeli	Postai	Telefonos	Online
Tartalom	Minden téma megkérdezhető	Inkább általános vélemény	Minden téma megkérdezhető	Gazdag tartalom (képek, filmek, animációk)
Költség	Magas	Alacsony	Alacsony	Nagyon alacsony
Hibaforrás	Viszonylag alacsony, kérdezői hibák	Viszonylag magas, minta hibák	Viszonylag alacsony, válaszadási hibák	Önkitöltés, mintavételi hibák
Lebonyolítás ideje	1-2 hét	8-10 hét	3-4 hét	1-2 hét

Forrás: Bauer és mtsai 2007

Napjainkban a klasszikus, papír alapú kérdőívek alkalmazása már háttérbe szorul az online kérdőívekkel szemben, hiszen az internet korunk egyik legfőbb információ forrásává, a

mindennapok részévé vált. Mindemellett számos, olyan tényezőt tudunk megemlíteni, ami az internetes, online kérdőívek mellett szól, használatukat indokolja:

- gyorsabbak és jóval hatékonyabbak a papír alapú kérdőívekkel szemben, nem beszélve az adatok gyűjtéséről, feldolgozásáról, kiértékeléséről
- finansiális szempontból jóval költséghatékonyabbak (nincsen nyomtatási és egyéb költség)
- szélesebb körben és könnyebben érhetőek el az emberek, mint személyesen – klasszikus formában
- jóval magasabb részvételi, válaszadási arány érhető el velük
- a kérdésekre adott válaszok jóval közelebb állnak a tényleges valósághoz, őszintébbek, hiszen a felhasználók anonimitás érzete az online esetben nagyobb (íráskep, azonosíthatóság nem lehetséges)
- az internet segítségével és a modern telekommunikáció eszközeivel olyan kérdésformákat is alkalmazhatunk, mely papír alapon nem lehetséges (mozgókép, stb..).²

4.2. A kérdőív előkészítése, általános tudnivalók, a készítés alapjai

Mielőtt kérdéseinket meghatároznánk, fontos hogy előzetes információt gyűjtsünk, kutatást végezzünk annak érdekében, hogy magunk is megismerhessük mélységében az adott témát, annak egyes dimenzióit. Érdemi munkát akkor tudunk végezni, ha feltérképezzük a felmerülő kérdéseket.

Ha a fent említett tényezőket ismerjük, pontosan meghatározhatunk problémákat, a későbbiekben kérdéseinket pontosan meg tudjuk fogalmazni, strukturálni.

Felmérésünk céljait már a kezdetekkor meg kell határozni. A végső céltól eltérni nem érdemes, hiszen ez torzíthatja eredményünket, ugyanakkor a továbbiakban felmerülő további vizsgálati lehetőségeket, aspektusokat fel lehet használni.

Meg kell fogalmazni azokat a nullhipotéziseket, hipotéziseket, melyeket munkánk során igazolni – cáfolni kívánunk. Nagyon fontos, hogy a hipotézis mindig kijelentés, sohasem kérdés!

Ezt követően meg kell határozni a kérdőív fő kérdéseit illetve azt, hogy azok pontosan a hipotézisekre irányuljanak.

² <http://www.kerdoivem.hu/advice/> (2014-07-16)

Meg kell szabnunk a minta nagyságot (becsült kitöltő szám) mivel egyes válaszok kiértékelésekor a nem megfelelő kitöltő szám az eredmény torzulásához és a későbbiekben téves következtetések levonásához vezethet. A kérdések meghatározása és szerkesztése előtt prognosztizálnunk kell a várt eredményt, hiszen az ahhoz szükséges információkra szükségünk van, de ha az erre irányuló kérdések hiányosak, vagy nem térnek ki teljesszűren a megismerni kívánt területre abban az esetben a kiértékelésnél problémák adódhatnak. Végezetül fontos megemlíteni, hogyha bonyolultabb statisztikai analízist kívánunk végrehajtani érdemes homogenizálni felmérésünkben a skálákat illetve azok szintjeit, fokozatait.³

Az online kérdőív szerkesztését megelőzően fontos végig gondolni, azon alapvető kritériumokat melyek szükségesek ahhoz, hogy a kérdőívvel a számunkra fontos információkat nyerjük ki és munkánk sikeres lehessen. Fontos, hogy a kitöltő számára kis erőfeszítéssel is kitölthető legyen, a kérdések megválaszolhatóak legyenek, olyan ismeretanyaghoz kötődjenek, melyek a kitöltő egyén „birtokában” vannak. Fontos még, hogy a válaszadót a kérdések motiválják, általánosan felkeltsék az érdeklődést, hogy ne válhasson érdektelenné a kitöltés. Kérdéseinket ennek értelmében a megfelelő formában kell feltenni és, ami még fontosabb a válaszadási hibák lehetőségét is minimalizálni kell, amennyire csak tudjuk.

Mielőtt a kérdéseket megfogalmaznánk, fontos meghatároznunk a kérdőív témáját és célját, szükséges megszabnunk a teljes kérdőív struktúráját, felépítését és a kérdések logikai sorrendjét is. Miután a téma, a cél, a szerkezet meghatározásra került a kérdéseket témánként több, egymástól eltérő egységre kell felosztanunk.

A kérdőívek általánosságban véve azt a logikai felépítést követik, melyben az általános, könnyen megválaszolható kérdésektől indulva jutunk el a témaspecifikus kérdésekig. Amennyiben az ettől való eltérést valami különleges ok nem indokolja, érdemes ezt a felépítést használni.

A fent említettek alapján a kérdéseket főbb szekciókba sorolhatjuk:

1. Kérdések a kitöltőről

Bár a kérdőívek kitöltése nagyrészt anonim módon történik, ezekben az esetekben sem árt megtudni általános információkat a válaszadókról. Ezen kérdéscsoporton belül érdemes

³ <http://www.kerdoivem.hu/advice/> (2014-07-16)

megkérdezni a kitöltő nemét, életkorát, iskolai végzettségét illetve minden olyan szocio-demográfiai, akár szocio-ökonómiai tényezőt, mely később hasznunkra lehet. Ezek a válaszok ugyanis alapját adhatják nemek, társadalmi rétegek, korosztályok összehasonlító elemzésének.

2. Általános információk

Ebben a kérdéscsoportban a válaszadók általános tapasztalataira illetve percepcióira, ismereteire vagyunk kíváncsiak az adott témával kapcsolatban.

Hasznos lehet ugyancsak az adott válaszok kiértékelésekor, hiszen nem mindegy, hogy a kitöltő mennyire van tisztában a feltett témakör alapismereteivel, vélemény esetén nem játszik e közre elfogultság válaszainak megadásakor.

(pl.: Mit gondol általában a sportokról? Mennyire elégedett saját sporttevékenységével? Mik a tapasztalatai általánosságban a sportlehetőségekről lakóhelyén?).

3. Témaszpecifikus kérdések

A kérdések ebben a csoportban már azokra a konkrét és kruciális ismeretekre, véleményekre koncentrálnak, melyek a kiértékelés törzsét adják majd. A kérdőív megszerkesztése szempontjából ez a legfontosabb kérdés szekció, emiatt különös figyelmet igényel és általában a legbővebb kérdés csoport.

(pl.: Mit gondol a kardio tréning élettani hatásairól? Mennyire elégedett az XY sportlétesítmény szolgáltatásaival kapcsolatban? Milyen gyakran végez kardio edzést?)

4. Problémamegoldás, javaslatok

Ebben a részben a kitöltő kifejtheti véleményét, látásmódját az adott témával kapcsolatban illetve megteheti javaslatait, melyek felhasználhatóak a későbbiekben.

Megemlítendő azonban, hogy ez a szekció nem feltétlenül képi részét minden kérdőívnek, inkább marketing jellegű, esetleges vállalatfejlesztési felmérések állandó szereplője.

A kérdéssor végén nagyon hasznos „feed-back” lehetőséget adni a válaszadóknak, melyben nyitott kérdés formájában jellemezhetik a kérdéseket, kifejezhetik véleményüket, esetleges javaslataikat. A javaslatok ugyanis új fókuszpontok kialakításához nyújthatnak segítséget a későbbiekben és az esetlegesen nem észlelt hibákra is fény derülhet⁴.

⁴ <http://www.kerdoivem.hu/advice/> (2014-07-16)

(pl.: Mit gondol, miért nem sportol heti rendszerességgel? Milyen további kérdést tartana még fontosnak a sportolással kapcsolatosan? Ön szerint mennyire méri fel a kérdőív a sportolással kapcsolatos attitűdöket, szokásokat?).

Kérdésekkel kapcsolatos általános tudnivalók

Ezen alfejezetben azokat a fontos és praktikus tanácsokat említjük meg, melyek figyelembevétele és alkalmazása szükséges a sikeresen kiértékelhető adatbázis „kinyeréséhez”. Több információ a kérdések feltevésének technikájáról, szerkezetéről szól, ezek a legfontosabbak, ha kérdőívet használunk.

1. A feltett kérdések egyszerűek, rövidek, tömörek legyenek. Próbáljunk meg közérthetően fogalmazni annak érdekében, hogy a kitöltő minden esetben tisztában legyen azzal, hogy a kérdés pontosan mire irányul. Kerüljük a többszörösen összetett, sajátos logikájú kérdéseket, a kérdésfeltevést ne komplikáljuk feleslegesen.
2. Kerüljük a túl sok kérdés egy oldalra helyezését. Bár a „Rács” kérdőív elemmel lehetőségünk van egyszerre több tényező gyakoriságát felmérni (mátrix kérdés), ugyanakkor ez esetben is csak a kifejezetten, értelmileg összetartozó faktorokat szerkesszük egy kérdésbe. Amennyiben lehetőségünk van rá több egyszerű kérdésre bontsuk szét összetett, akár „mátrix-kérdéseinket”.
3. Mindig tartsuk szem előtt, hogy a kitöltő figyelmét és érdeklődését a kitöltés elejétől a végéig fent kell tudni tartani, ellenkező esetben pontatlan, egyes esetekben spontán válaszokat vagy félig kitöltött kérdőíveket kapunk vissza, melyeket nem tudunk kiértékelni és adatbázisunk tele lesz hiányzó változókkal.
4. Fontos, hogy folyamatosan használjunk szűrést valamint továbbító válaszokat, az olyan kérdéseknél, ahol szokásokról – véleményekről érdeklődünk. Ha ezt nem tesszük a kitöltőre előbb utóbb irreleváns kérdések sorozata „zúdul”, mely ugyancsak pontatlan válaszadáshoz, hiányos kitöltéshez vezet. (pl.: Sportol-e Ön? – ha a válasz nem, akkor ne kérdezzük meg, hogy mit sportol?). Mindig legyen tovább vezető kérdés, hogy amennyiben nemleges a válasz, gördülékenyen lehessen tovább haladni a kitöltéssel.
5. Mielőtt a kérdőívet közzé tesszük meg kell győződnünk arról, hogy a kitöltők számára az összes elképzelhető válasz lehetőséget megadtuk, amennyiben ebben nem vagyunk biztosak, úgy érdemes egy „Egyéb” rubrika beszurása a válaszadási lehetőségek közé,

melyet az egyén maga tölthet ki, amennyiben úgy ítéli meg, hogy az adni kívánt válasz nem szerepel a felsorolásban.

- 6.** Kérdésünket mindig pontosan fogalmazzuk meg, hogy az ne legyen félreérhető!
- 7.** Kerüljük a szuggesztív, rávezető kérdés feltételét, mely befolyásolhatja az adott választ illetve a kitöltő objektív megítélését! (pl.: Melyiket választaná egészségének megőrzése érdekében, az egészséges gyümölcsöket vagy a zsíros húsokat? Ugye nem szokott túl zsírosan táplálkozni?, stb.)
- 8.** Fontos, hogyha gyakoriságról vagy mennyiségről kérdezünk, mindig határozzuk meg, hogy az általunk megadott válasz opció pontosan mit jelent, hiszen a gyakoriságot meghatározó szavak mindenki számára mást jelenthetnek és mindenki saját maga értelmezésében fogja használni (pl.: általában, gyakran, alkalmanként). Mindig egyértelmű mértékegységet vagy időtartamot határozzuk meg (pl.: cm, hetente 2-3 alkalommal).
- 9.** Amennyiben lehet kerüljük a „Nem tudom” és a „Nehéz megválaszolni” válaszadási lehetőségeket, mivel ezek a válaszopciók kiértékelhetőség szempontjából nem irányadóak. Természetesen, speciális esetekben alkalmazható, de csak akkor, ha a kérdésre való válaszadás valóban nehézséget okozhat a kitöltőnek.
- 10.** A feltett kérdésen belül ne használjunk ellentétes értelmű szavakat, mert az megzavarhatja a válaszadó korrekt értelmezését. Figyelmet kell még fordítanunk arra is, hogy a hasonló olvasatú szavakat valamilyen más szinonimával helyettesítsük (pl.: helyettesítsük: az alkalmazott / alkalmazatlan szópárost alkalmazott / nem használt szópárral).
- 11.** Különös figyelemmel fogalmazzuk meg az érzékeny területekre irányuló kérdéseket (pl.: jövedelemszint) illetve a válaszopciókat, még anonim kérdőív esetén is. Ellenkező esetben a kitöltők megtagadhatják a válaszadást. Ezekben az esetekben határozzuk meg kategorizált, intervallumokat (-tól -ig értékeket). Lehetőség szerint az intervallumokat úgy szabjuk meg, hogy azok szélsőséges értékeket is tartalmazzanak valamilyen formában.
- 12.** Adjuk meg a lehetőséget a válaszadás megtagadására: érzékeny, a kitöltő számára esetleg kellemetlen kérdések esetén, mely a „nem hajlandóságot” jelzi számunkra. Ez azért fontos mert elképzelhető, hogy az opció hiányában a kitöltő a kérdőívet emiatt nem tölti ki, vagy nem küldi vissza nekünk válaszait, így ezekben az esetekben jobb egy vagy néhány hiányzó adat, mint egy egész kérdőív elvesztése.

13. A válaszadók figyelméről - kérdés megfogalmazás tekintetében - már esett szó, de a kérdésszámok tekintetében még nem, pedig ez is befolyásoló tényező lehet. Fontos, hogy túl hosszú, sok kérdésből álló kérdőívet ne szerkesszünk és minden kérdés szükségességét gondoljuk át. Már a kérdőív elején tegyünk fel figyelemfelkeltő kérdéseket, azokkal jó benyomást keltve a kitöltőben. A kérdőív elején lehetőség szerint nyitott kérdéseket ne tegyünk fel, ezek helye inkább a kérdőív közepe-vége fele ajánlott.
14. Mielőtt a kérdőívet „élesben” alkalmazzuk, végezzünk próba kitöltést az esetleges hibák felderítésére és a kérdőív tesztelésére. Ügyeljünk a helyesírás szabályainak maradéktalan betartására is ⁵!

Kérdéstípusok

A következőkben bemutatjuk azokat a kérdőív elemeket, melyeket leginkább adatgyűjtésre használhatunk. Ezt megelőzően azonban tisztázni kell a kérdések két alap típusát.

Nyitott kérdés

A nyitott kérdéstípusnál a kitöltőnek saját szavaival kell megfogalmazni a választ a feltett kérdésre, nem pedig egy előre meghatározott listából kiválasztani az egy vagy több opciót. Azokban az esetekben ajánlott a használata, melyekben a válaszadó érzelmeire, megítélésre vagyunk kíváncsiak illetve arra, hogy egyes döntéseket milyen tényezők alapján hoz meg. Kerülni kell ezt a kérdés típust az önkitöltős kérdőívek esetében.

Statisztikai szempontból a nyitott kérdések feldolgozása nehézkes, sok esetben nem megoldható kódolásuk, hiszen a kitöltő saját maga fogalmazza meg válaszát.

Zárt kérdés

A zárt kérdés a nyitott kérdése ellentéte, tehát a válaszadó egy adott kérdésre előre meghatározott, rögzített válaszok közül választhatja ki a neki leginkább megfelelőt, saját válaszkategória megadására a legtöbb esetben nincs lehetőség.

Statisztikai szempontból jól feldolgozható, kódolható és könnyen kezelhető. Több formája ismert, így a választási lehetőségek száma alapján megkülönböztetünk alternatív (tipikus

⁵ <http://www.kerdoivem.hu/advice/> (2014-07-16)

<http://www.kerdoivem.hu/questions/> (2014-07-16)

„IGAZ vagy HAMIS” kérdések), szelektív (több válasz is lehetséges) skála (minősítési sorrend 1-5) kérdéseket.

Skála-kérdés

A „Skála – kérdés” érzelmek, attitűdök felmérésre alkalmazható, azok pozitív-negatív voltának mértékét határozhatjuk meg segítségükkel. A skálák több típusa ismert (ordinális skála, intervallum skála, arány skála, nominális skála). A leggyakrabban 1-5-ig terjedő skálát alkalmazunk kérdőívekben. Figyelni kell a skála egyes szintjeinek egymáshoz viszonyított szimmetriájára is, hogy a semleges attitűd (ez esetben a 3.) ugyanolyan távolságra legyen mind a pozitív, mind a negatív végponttól. Fontos még, hogy amennyiben más léptékű, fokozatszámú skálát használunk, az mindig páratlan fokozatú skála legyen. Ennek az az oka, hogy a szimmetria csak így tartható fent, hiszen csak így van teljesen elkülöníthető középpont- semleges vélemény (5- fokozatú skálán ez a 3.).

A gyakorlatban 7 szintű skálánál nagyobb tartományú változót nem szoktunk használni, mert a skála szintek növelésével csökken az egyes lépcsők közötti távolság és egyre nehezebb objektívan meghatározni véleményünket illetve különbséget tenni két egymás melletti szint között. Amennyiben egy állítással való egyetértést mértékét határozzuk meg abban az esetben a használt tartomány változót „Likert-skálának” hívjuk. Ezen skála esetében csak az állításokat kell megfogalmaznunk, maga a skála mindig ugyanaz. Ezért alkalmazható jól olyan esetekben ahol a skálák homogenizálása szükséges (bonyolultabb statisztikai próbák) (Lógó 2007).

Klasszikus 7 fokozatú „Likert-skála”:

- Teljes mértékben egyetértek
- Nagyrészt egyetértek
- Csak kismértékben értek egyet
- Közömbös számomra
- Nem túlságosan értek egyet
- Nagyrészt nem értek egyet
- Egyáltalán nem értek egyet

Alkalmazható kérdőív elemek

Az elemeket a Google Docs™ kérdőív készítőjének segítségével mutatom be. Az előbb említett kérdőív készítő több szempontból is a legoptimálisabb tudományos munkánk,

felmérésünk lebonyolításához. Könnyen létrehozható és mindenki számára ingyenesen elérhető, aki rendelkezik Google-fiókkal (Gmail). Felhasználóbarát felülettel és e-mailben megosztható linkkel bír. Kérdőívünk könnyen beágyazható weblapra, közösségi oldalra egyaránt. Adatbázisunkat Excel-táblázat⁶ formájában letölthetjük felmérésünk végeztével. Ugyanakkor, ami talán a legfontosabb, hogy minden tekintetben korlátozásmentes (kérdőív kérdések száma, kitöltők száma, megosztások száma) (Farkas 2012).

„Szöveg / Text”

Rövid szöveg beírására ad lehetőséget, bármilyen információval kapcsolatban (pl.: elérhetőség, email cím).



4/1. ábra: A „Szöveg” kérdőív elem megjelenése

„Hosszabb szöveg / Paragraph Text”

A „Hosszabb szöveg” ugyancsak saját válaszadásra ad lehetőséget, alkalmazása hosszabb terjedelmű szövegeknél indokolt (pl.: vélemény, javaslat)



4/2. ábra: A „Hosszabb szöveg” kérdőív elem megjelenése

„Feleletválasztós / Multiple Choice”

Ezt a kérdés típust akkor használjuk, ha a felsorolt verziókból azt szeretnénk, ha a kitöltő csak egyet választania, ennek lehet logikai vagy statisztikai kiértékelhetőségi oka is. Plusz kiegészítési opció, ha az „Egyéb” (*Add Other*) rubrikát bejelölve lehetőséget adunk a kitöltőnek egy általunk előre nem megjelölt választási lehetőség megadására, jelölésére.

⁶ **Excel-táblázat:** A Microsoft Office programcsomag Microsoft Excel táblázat kezelő alkalmazása, programja.

Legmagasabb iskolai végzettsége:

- Általános iskola
- Középiskola
- Főiskola
- Egyetem
- Tudományos fokozat

4/3. ábra: A „Feleletválasztós” kérdőív elem megjelenése

„Jelölő négyzet / Checkboxes”

Ez esetben ugyancsak feleletválasztós kérdéstípust választhatunk, azzal a különbséggel, hogy itt egyszerre több választ is megjelölhet a kitöltő. Megemlítenéd, hogy amennyiben szeretnénk, itt is rendelkezésre áll az „Egyéb” megadásának lehetősége.

Milyen forrásból származnak bevételei?

- Munkabér
- Családi pótlék
- Szociális segély
- Bérkompenzáció
- Másodállás

4/4. ábra: A „Jelölőnégyzet” kérdőív elem megjelenése

„Választás listából / Choose from a list”

Funkció tekintetében gyakorlatilag ugyanaz, mint a „Feleletválasztós” kérdőív elem, azzal a különbséggel, hogy itt egy legördülő listából választhatja ki a válaszadó, az általa jelölni kívánt választ.

Legmagasabb iskolai végzettsége:

Általános iskola
Középiskola
Főiskola
Egyetem
További tudományos fokozat

4/5. ábra: A „Választás listából” kérdőív elem megjelenése

„Tartomány / Scale”

Ennél az elemnél egy adott tartományt, skálát adhatunk meg –tól –ig értékek megadásával, mely 1-5-ig értékelhető (a gyakoriságot meghatározó: “mindig – gyakran – soha” típusú

kérdésekhez használható). A statisztikában, ezt a változót - amennyiben egyetértést mérünk vele – hívjuk az előzőekben már ismertetett „Likert-skálának” (egyáltalán nem – teljes mértékben egyetértek).

Milyen gyakran sportol, végez valamilyen mozgást:

1 2 3 4 5

soha mindig

4/6. ábra: A „Tartomány” kérdőív elem megjelenése

„Rács / Grid”

Hasonló a „Tartomány” (*Scale*) elemhez, ugyanakkor az a különbség, hogy itt egyszerre több tényezőt csoportosíthatunk egy kérdésen belül, téma alapján és azokat egyéni igényünk szerint súlyozhatjuk, különféle gyakorisági címkék megadásával. Nemcsak –tól –ig gyakoriságot határozhatunk meg.

Milyen gyakran „üzi” az alábbi sportágakat?

	soha	alkalmanként (havonta)	általában (2 hetente legalább 1 alkalommal)	gyakran (heti rendszerességgel, de maximum 2 alkalommal)	mindig (hetente legalább 3 alkalommal)
Labdarúgás	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Kosárlabda	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Kézilabda	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Röplabda	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Futás	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

4/7. ábra: A „Rács” kérdőív elem megjelenése

„Dátum / Date”

Ezt a kérdés formátumot akkor használjuk, ha egy adott eseményre, vagyis annak pontos dátumszerű időpontjára vagyunk kíváncsiak (pl.: születés ideje).

Születés ideje:

éééé. hh. nn. ▼

2014. július ▼

H	K	Sze	Cs	P	Szo	V
30	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31	1	2	3

4/8. ábra: A „Dátum” kérdőív elem megjelenése

„Idő / Time”

Az „Idő” (*Time*) kérdés formátum gyakorlatilag ugyanaz mint a „Dátum” (*Date*) formátum, azzal a különbséggel, hogy itt időpontot adhatunk meg, nem pedig dátumot.

Az edzés kezdete:

Óra ▼ - perc ▼ de./du. ▼

4/9. ábra: Az „Idő” kérdőív elem megjelenése

4.3. Kérdőívszerkesztés a gyakorlatban: a Google Docs alkalmazás segítségével

Az előzőekben már említettük a Google Docs™ illetve a „Google Űrlap” (*Google Form*) előnyeit és alkalmazhatóságát, ezen alfejezetben megismerkedünk a felülettel illetve a bemutatást követően egy konkrét rövid kérdéssort is elkészítünk majd.

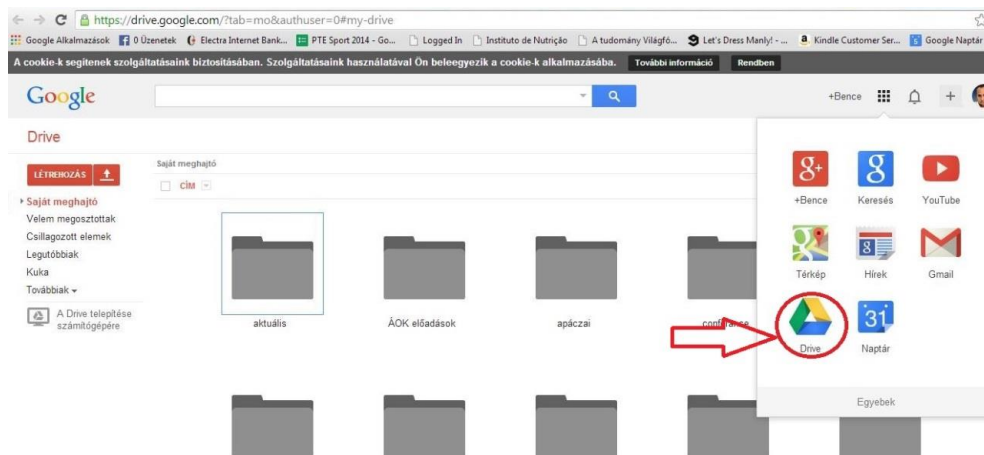
A „Google Űrlap” (*Google Form*) alkalmazásban történő adatbázis kinyerésének, négy fő lépése van és az első alkalmakkor még egy varázsló is segítségünkre van. Először létre kell hoznunk, magát a kérdőívet („Gyorsan létrehozhatja az űrlapokat a billentyűkódok és a módosítások automatikus mentése segítségével”). Fontos megemlíteni, hogy egyszerre többen is szerkeszthetjük az űrlapot teljesen valós időben, hasonlóan működik, mint egy közös, munkacsoport email címe vagy akár a Google egymással is megosztható naptára („Másokkal valós időben együttműködve hozhat létre űrlapokat”). Ezután amikor a kérdőív elkészült és „éles” meg kell osztanunk a linket, melyet forráskód formájában tudunk kinyerni. Beágyazható közösségi oldalra, honlapra és e-mailben is kiküldhető. Utolsó lépésként a visszaérkezett válaszokat összesíti, a rendszer, melyeket Excel formájában kinyerhetünk, lementhetünk számítógépünkre („A válaszokat táblázatba küldve hatékony elemzést

végezhet”). Bár a rendszerbe beépített kiértékelő rendszer is helyet kapott (grafikonok, diagramok megjelenítés) alkalmazását nem javasoljuk komolyabb tudományos munkáknál illetve olyan esetekben, ahol bonyolultabb, összehasonlító statisztikai elemzést kívánunk végezni (Farkas 2012).



4/10 .ábra: A Google Űrlapok szemléltető ismertetője annak használatáról

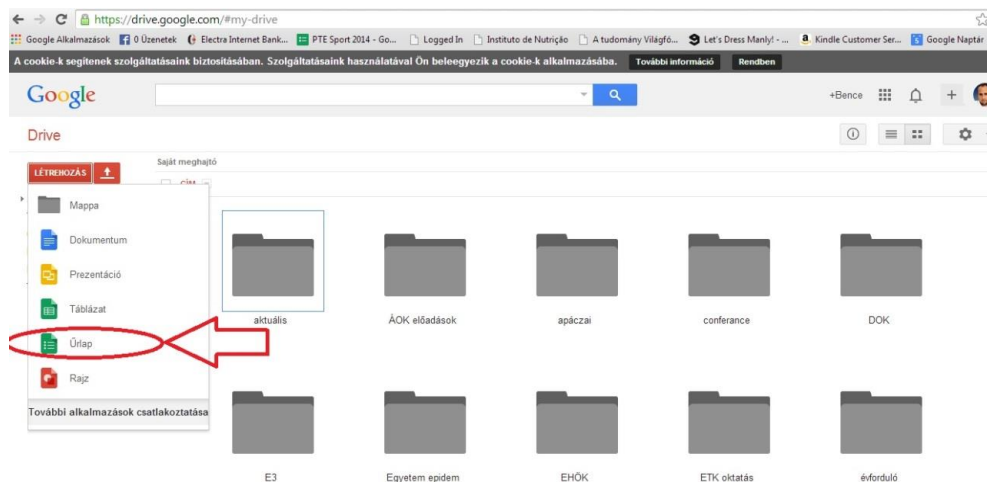
Ahhoz, hogy neki állhassunk kérdőívünk szerkesztéséhez először be kell lépünk Google-fiókunkba, majd ellátogatni a „Drive” alkalmazásba, ezt „Gmail”-es fiókunkból úgy érhetjük el, hogy email fiókunk kezelő felületének jobb felső sarkában rákattintunk (bal klikk), az alkalmazások menüpontra, melyből kiválasztjuk a „Drive” alkalmazást.



4/11. ábra: Eljutás a Google-fiókból a „Drive” felületére

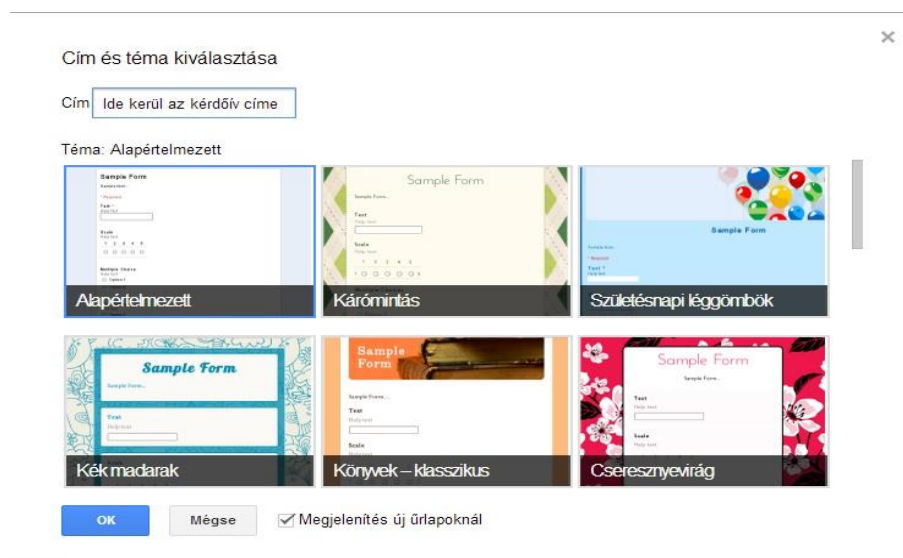
Ha ez valamiért nem sikerül más módon is el tudunk jutni a „Drive” alkalmazásba, ebben az esetben is először be kell lépniünk Google-fiókunkba, majd ezt követően új lapot nyitni böngészőnkben és oda bemásolni a következő webcímet: <https://drive.google.com/>.

Amikor valamilyen módon eljutottunk a „Drive” alkalmazásba, létre kell hoznunk egy új űrlapot, ezt úgy tehetjük meg, hogy a kezelő felület bal oldalán található „Létrehozás” (*Create*) gombra kattintunk majd ott az „Űrlap” (*Form*) menüpontot választjuk ki a legördülő listából.



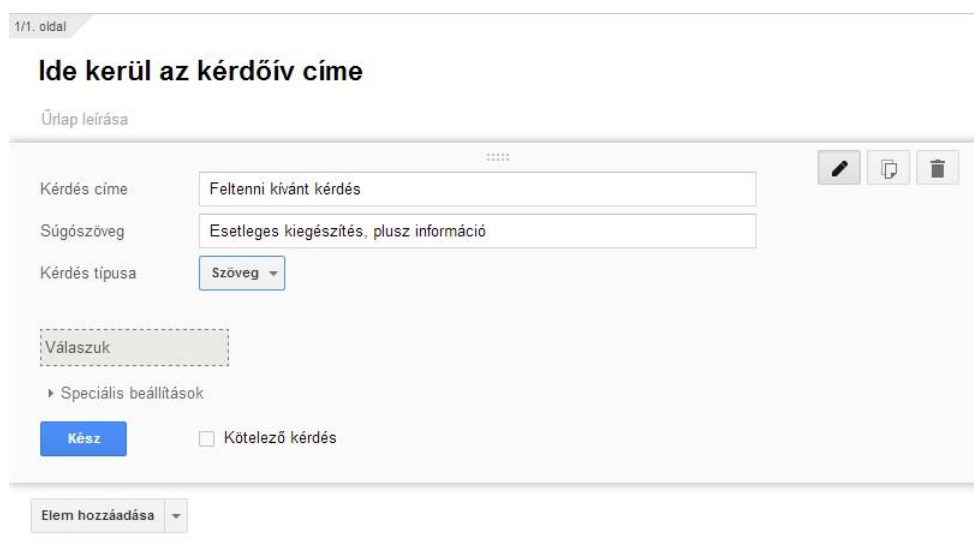
4/12. ábra: Új űrlap létrehozása a „Drive” alkalmazásban

Az űrlap kiválasztását követően, először a kérdőív címének megadására valamint stílusának kiválasztására van lehetőség. Mindenkinek saját döntése és joga, hogy kérdőívének stílusát kiválassza, ugyanakkor egy tudományos felmérés komolyságának megfelelően az „Alapértelmezett” stílus alkalmazását javasoljuk.



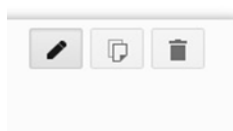
4/13. ábra: Űrlap cím és stílus megadása, kiválasztása

A cím és stílus megadását követően belépünk a kérdőív szerkesztő belső felületére. A felületen kérdésenként szerkeszthetjük össze kérdőívünket. Kérdést az „Elem hozzáadás” (*Add Item*) gombra kattintva, a felugró panelen keresztül adhatunk meg. A „Kérdés címe” (*Question Title*) sávba értelem szerűen a feltenni kívánt kérdés kerül, a közvetlenül alatta levő „Súgószöveg” (*Help Text*) részben adhatunk meg plusz információkat, esetleg kiegészítéseket, melyek segítik a korrekt válaszadást illetve a kérdés értelmezését. A „Kérdés típusánál” (*Question Type*) pedig kiválaszthatjuk a nekünk szükséges kérdés típust, ezeket már az „Alkalmazható kérdőív elemek” alfejezetnél ismertettük, bemutattuk. Az alábbi ábrán látható, hogy a kérdés egyes rubrikáiban minek, hol a helye.



4/14. ábra: Kérdés szerkesztő panel

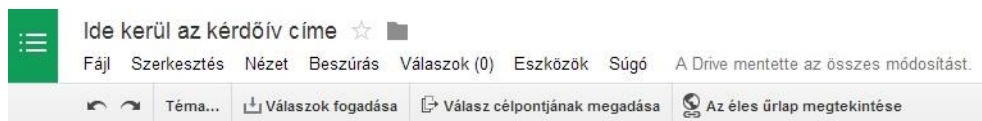
Az előbb említett módon kérdésről-kérdésre haladhatunk és állíthatjuk össze az általunk feltenni kívánt kérdéssor(oka)t, ugyanakkor utólagos változtatásra is van lehetőség. Az adott, szerkeszteni kívánt kérdésbe bármikor bele kattintva szabadon áthelyezhetjük azt, amennyiben meg szeretnénk változtatni a kérdések sorrendjét. További korrekciós lehetőségeket az előbb bemutatott panel jobb felső sarkában találunk még.



4/15. ábra: A korrekcióra szolgáló ikonok

A fent látható három ikonnak természetesen eltérő funkciója van. Balról az első ikon (ceruza) segítségével bármikor beleszerkeszthetünk az adott kérdésbe, esetleges helyesírási hibát javíthatunk. A középső ikon (dokumentumok) segítségével az adott kérdésről másolatot készíthetünk. Például ha egy szövegezésben vagy kérdéstípusban hasonló dologra vagyunk kíváncsiak, akkor nem szükséges az egész kérdést újra írni vagy a kérdéstípust újra beállítani. A harmadik ikon (kuka) értelemszerűen a kérdés teljes törlését jelenti, ha időközben kiderülne, hogy egy adott kérdésre nincs szükségünk.

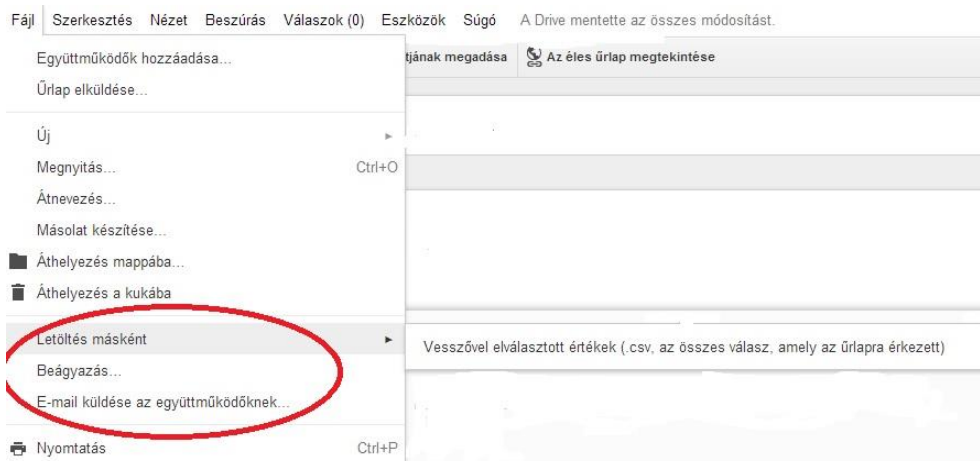
Ha a szerkesztéssel megvagyunk további lehetőségeket még az alapfelület felső részében találhatunk. Ezeket a következőkben ismertetjük, ugyanakkor, mivel ezek egy része a legtöbb más programnál is hasonló funkciót tölt be, így leginkább azokat a pontokat magyarázzuk meg melyek ezektől eltérőek.



4/16. ábra: A kérdéssor fő, szerkesztő sávja

A „Fájl” menüpont alatt általános opciókat találunk, melyek közül kiemelt fontosságúnak a „Letöltés másként”, a „Beágyazás” és az „E-mail küldése az együttműködőknek” menüpontot említhetjük. A „Letöltés másként” menüpontban letölthetjük magát az összes kérdőívre érkezett választ, a fájl formátuma .csv. A „Beágyazás” segítségével megkaphatjuk azt a „html kódot”⁷, melyet honlapokra történő beágyazásra tudunk használni. Az „Email küldése az együttműködőknek” menüpont, nem szorul bővebb magyarázatra, érdemes megemlíteni, hogy a képernyő jobb sarkában található kék „Űrlap elküldése” gorsgomb ugyanezt a célt szolgálja.

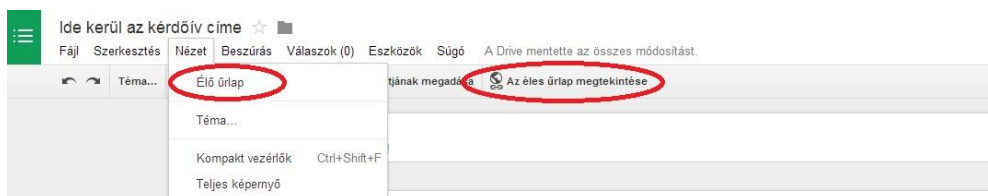
⁷ **HTML (HyperText Markup Language):** hiperszöveges jelölőnyelv



4/17. ábra: A „Fájl” menüpont egyes funkciói

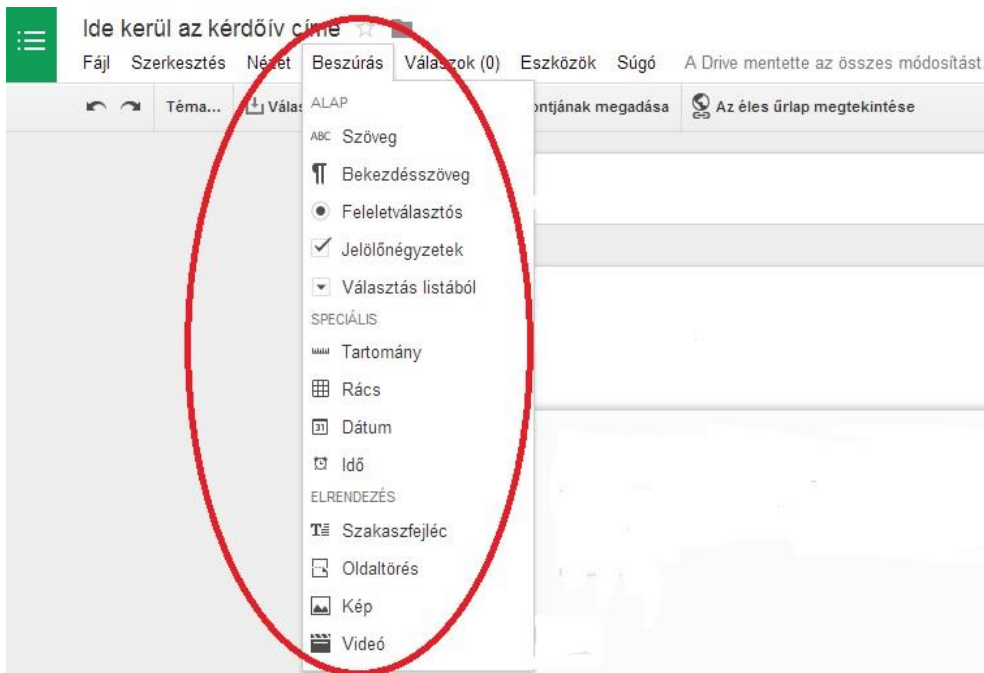
A „Szerkesztés” menüpontban az egyes elvégzett műveletek között léphetünk előre, illetve vonhatjuk vissza azokat.

A „Nézet” menüpontban azt állíthatjuk be, hogy milyen nézetben kívánjuk a kérdőívet megtekinteni többek között teljes képernyős verzió is található a lehetőségek között, de ami talán még fontosabb az „Élő űrlap” opciót kiválasztva megnézhetjük kérdőívünk végleges formáját, ahogyan majd a kitöltők is látják, ugyanerre a célra szolgál az „Az éles űrlap megtekintése” gyorsgomb is.



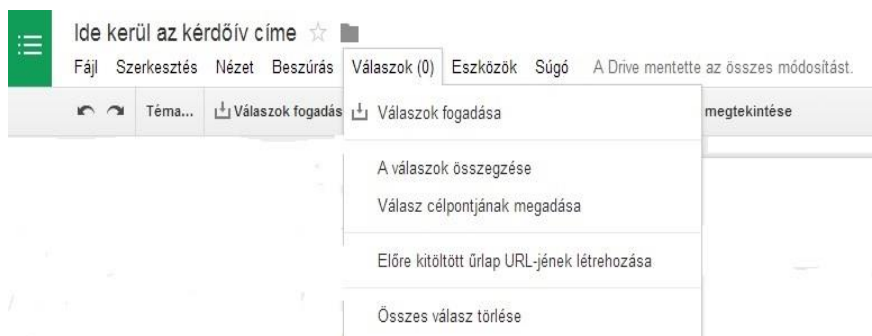
4/18. ábra: A „Nézet” menüpont opciói

A beszúrás menüpontban szereplő opciók egyrészt a már említett kérdéstípusok beszúrására alkalmasak, másrészt formai szerkesztéseket is eszközölhetünk velük, mint a „Szakaszfejléc”. Ez az opció akkor bír jelentőséggel, hogyha valamilyen okból szeparálni szeretnénk az egyes kérdéscsoportokat és ennek látható jelét szeretnénk adni a kitöltő felé annak logikai vagy tartalmi jelentősége miatt. Az „Oldaltörés” hasonló funkciót tölts be, mint az általános szövegszerkesztő programokban. A „Kép” illetve „Video” opció - kép, mozgóképfelrakás kérdőívbe történő beágyazására szolgál.



4/19. ábra: A „Beszúrás” menüpont opciói

A „Válaszok” menüpontban kérdőívünkre beérkező válaszokat kezelhetjük. A „Válaszok fogadása” gomb azt jelenti, hogy kérdőívünk aktív, közzé tehető illetve fogadjuk az érkező válaszokat. Egyet kattintva erre a gombra inaktívvá tehetjük azt (továbbiakban nem lesz elérhető más számára, amíg mi vissza nem állítjuk).



4/20. ábra: A „Válaszok” menüpont opciói

A „Válasz célpontjának megadása” menüpontban megadhatjuk, hogy a beérkezett válaszokat új Excel-táblába vagy már egy létező tábla új munkalapjára kívánjuk importálni.

Válasz célpontjának kiválasztása

Új táblázat

Ide kerül az kérdőív címe (válaszok)

Új munkalap meglévő táblázatban...

Mindig hozzon létre új táblázatot ?

Létrehozás

Csak az Űrlapokban lévő válaszok megőrzése

További információ



4/21. ábra: A válaszok célpontjának megválasztása

Az „Összes válasz törlése” segítségével ugyanazt a kérdőívet új mintán, csoporton is tesztelhetjük, előző adataink teljes mértékben törlésre kerülnek. A „Válaszok összegzése” menüpontot részletesen a következőkben tárgyaljuk.

A további menüpontok „Eszközök” illetve a „Súgó”. Az eszközök menüpontban szkript⁸ kezelőt és szerkesztőt találunk, míg a súgó a hibajelentésekre, gyors billentyűk bemutatására valamint hasznos információk elérésére használhatjuk.

4.4. Válaszok összegzése, kiértékelése

Mint azt már az előzőekben említettük, a beérkezett válaszokat illetve eredményeket a program összesíti számunkra, ugyanakkor azt ajánljuk, hogy a további statisztikai próbákat, szűrést illetve az adatbázis feldolgozását magunk végezzük az eredményeinkből kinyert Excel-tábla segítségével.

A beérkezett válaszokat illetve eredményeinket a „Válaszok megtekintése” (*See responses*) gombra kattintva tekinthetjük meg, ekkor a program automatikusan az adatainkkal feltöltött táblázatot adja ki, ahonnan további lehetőségünk nyílik műveletekre.



4/22. ábra: Válaszok megtekintése

⁸ **szkript**: rövid program, ami gyakran egy-egy részfeladat automatizálására szolgál

Még mielőtt a további műveleteket ismertetném, a fent látható ábrán is látszódó „Válaszok” menüpont alatt található a „Válaszok összegzése” (*Summary*) menüpont. Ezen menüpont alkalmazásával készít a program a kérdésekre adott válaszokból grafikonokat, diagramokat, de különösebb szelektációs vagy fókuszpont meghatározási lehetőségünk nincs, így ezt tényleg csak információ megtekintés céljára, nem pedig statisztikai kimutatásokra ajánljuk.

Milyen gyakran sportolsz, amióta egyetemre jársz (kötelező testnevelés órákon kívül)?



Mennyi időt fordítasz egy alkalommal sportolásra?



4/23. ábra: Eredmények képes összefoglalása

További lehetőség nyílik szerkesztésre, rendezésre, a válaszainkat már Excel-táblába tömörítő felületen. Mint ahogyan eddig is, itt is azokat a részeket említjük, amik az eredmények rendszerezése szempontjából fontosak számunkra illetve eltérőek az alapprogramokban alkalmazott műveleteknél.

A bal felső, első, szürke sorban a „*Timestamp*” automatikusan megjelenik mindig, ez arról az időpontról ad tájékoztatást, hogy mikor töltötte ki a válaszadó kérdőívünket. A többi oszlop legfelső cellájába a kérdésünk kerül, hogy tudjuk, mire adták az adott választ a kitöltők. Ha ezekbe a szürke oszlopcímekbe jobb egérgombbal belekattintunk, akkor a „Rendezés” (*Sort*) felirattal ábécé sorrendbe helyezhetjük az adatainkat (két irány van: A -> Z, és Z -> A), ugyanezt a funkciót látja el az Adatok → „Munkalap rendezés A-Z-” vagy „Munkalap rendezés Z-A”.

Timestamp	A PTE melyik karán folytatás tanulmányokat?	Hány éves vagy?	Nemed?	Milyen gyakran sportolsz, amióta egyetemre jársz (kötelező testnevelés órákon kívül)?	Mennyi időt fordítasz egy alkalommal sportolásra?	Milyen gyakran sportolnál ideális esetben?
10/29/2013 13:50:05	TTK	27	Nő	legalább heti 4 alkalommal	61-90 percet	minden nap

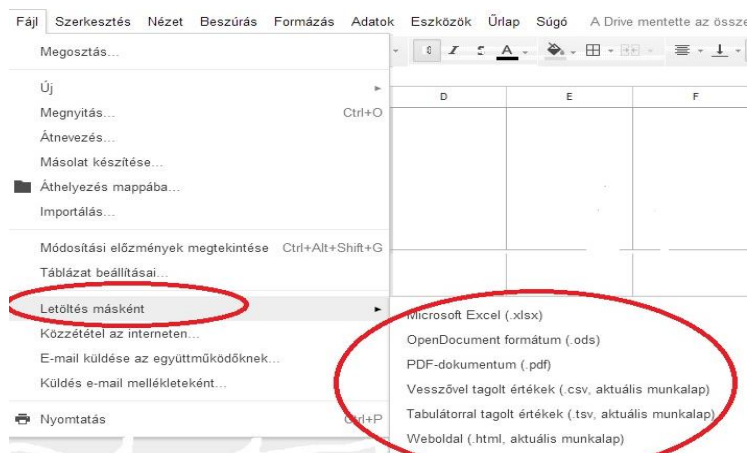
4/24. ábra: Adatbázis felépítése és elrendezése

A „Letöltés másként” (*Save as*) azért fontos menüpont, mert itt tölthetjük le különálló és általunk kiválasztott fájlformátumban, kiterjesztésben az eredményeinket tartalmazó fájlt.

Többféleképpen is letölthetjük adatbázisunkat:

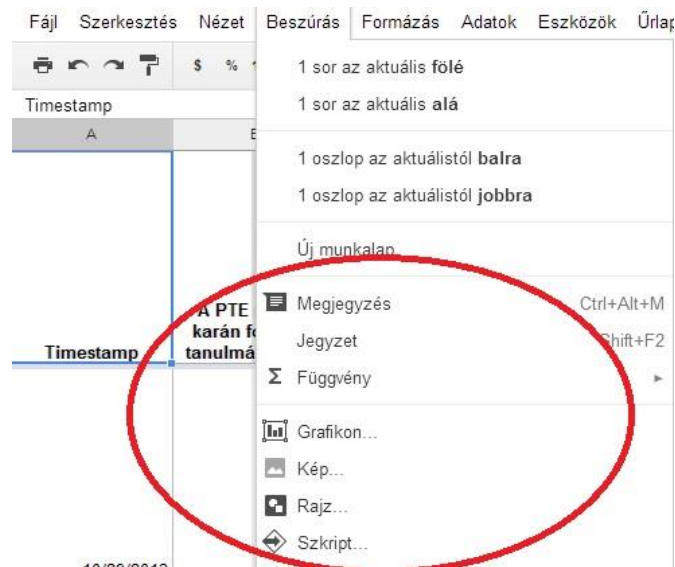
- Microsoft Excel formátum (.xlsx)
- OpenDocument Formátum (.ods)
- PDF dokumentum (.pdf)
- Vesszővel tagolt értékek (.csv)
- Tabulátorral tagolt értékek (.tsv)
- Web (.html)

További szerkeszthetőségi és kiértékelési szempontból a leginkább a Microsoft Excel (.xlsx) formátumot ajánljuk.



4/25. ábra: A „Letöltés másként” menüpont

Érdeemes még megemlíteni a „Beszúrás” menüpontot, melyben több lehetőség adódik egyes plusz elemek beszúrára, közvetlenül - direkt módon - az adatbázisba. Megjegyzést, különálló jegyzetet, függvényt stb. szúrhatunk be az egyes cellákhoz.



4/26. ábra: A „Beszúrás” menüpont

A „Formázás” menüpontban ugyanazokat a kinézet stilisztikai és betűtípusra vonatkozó opciókat találjuk, mint egy általános szövegszerkesztő programban.

Bejegyzésül még egy hasznos elemet kell megemlíteni, ez pedig az Adatok → Szűrő. Ugyanis még az online adatbázisban is van lehetőségünk különféle szűrők beállítására, alkalmazására, (pl.: nemek közötti különbségek összehasonlítása).

4.5. Kérdőívszerkesztés egy gyakorlati példán keresztül

Az eddigiekben tisztáztuk a kérdőív készítés alapjait képező fogalmakat, megismerkedtünk azokkal a módszertani irányelvekkel, melyek szükségesek egy szakmailag helytálló kérdőív szerkesztéséhez és egy tudományos munka alapját képező adatbázis kinyeréséhez. Ezt követően megismerkedtünk a „Google Űrlap” programjával, annak elemeivel és szerkesztési lehetőségeivel. Ebben az alfejezetben egy rövid, tényleges kérdőív összeállítását mutatjuk be, pontról pontra.

Az általunk kiválasztott téma illetve kérdőívvel feltárni kívánt terület az egyes terheesség alatti és utáni testsúlyváltozások és azok okainak felmérése.

A kérdéseket úgy válogattuk össze, hogy a metodikai részben szereplő legtöbb kérdés típus szerepeljen kérdőívünkben és az ismertetett irányelveknek is megfelelő legyen, tehát effektíve

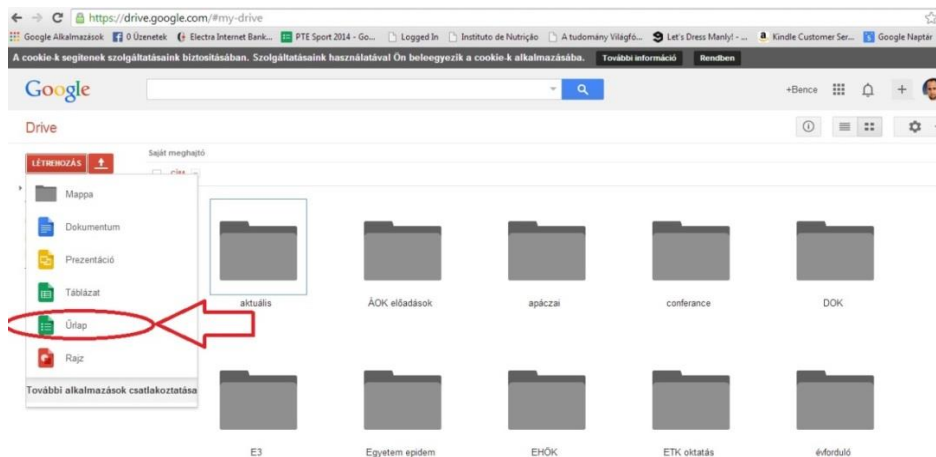
egy valós helyzetet kívánunk demonstrálni. Ennek értelmében kérdőívünk 14 kérdést tartalmaz, melyeket az alábbi sorrendben kívánunk feltenni:

- Betegkód
- Hány éves Ön?
- Mi az Ön legmagasabb iskolai végzettsége?
- Honnan származik Ön?
- Testmagassága?
- Szülés előtti testsúlya?
- Szülés napján mért testsúlya?
- Gyermekeének születési súlya?
- Szülés után 6 hónappal mért testsúlya?
- Korszerűnek ítéli meg az egészséges táplálkozással kapcsolatos ismereteit?
- Hányadik terhességi héten szült?
- Volt e krónikus betegsége?
- Szülésének kimenetele?
- Szülés közben érzett fájdalom (1-5 skálán)?

A kérdéseket az alábbiak szerint sorba rendeztük, most pedig szerkesszük meg kérdőívünket a „Google Űrlap” segítségével.

Első lépésként lépünk be Google fiókunkba, majd a Drive alkalmazásba, ahogyan azt az általános részben bemutatunk.

Amikor a „Drive” szolgáltatásban vagyunk, hozzunk létre új űrlapot, ezt a már ismertetett módon a képernyő bal felső sarkában található piros „Létrehozás” (*Create*) gomb segítségével tehetjük meg, ahol a legördülő listából a „Űrlap” (*Form*) opciót választjuk.



4/27. ábra: Új űrlap létrehozása a „Drive” alkalmazásban (előzőekben ismertetve)

Ha ezzel elkészültünk és előttünk van az űrlap szerkesztő felület, először válasszuk ki az általános témát valamint adjuk meg a kérdőív címét, esetünkben a cím a „Terhesség alatti és utáni testsúlyváltozások”.

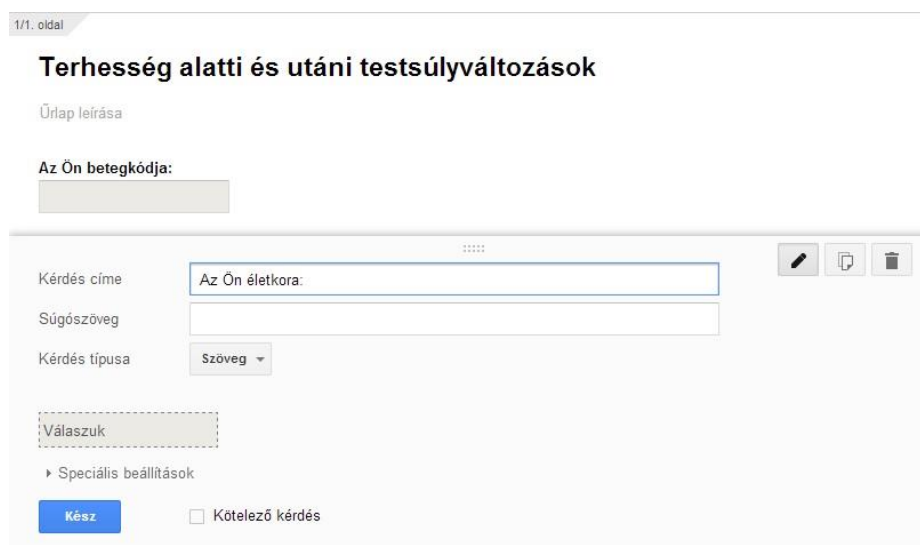
Első kérdésünk a betegkód, ezt előzetesen kiküldjük a felmérésben résztvevők számára így feladatunk csupán annyi, hogy az „Elem hozzáadása” gyorsgombra kattintva egy rövid, szöveges kérdést tegyünk fel, melyben ezt az információt meg tudják majd adni a válaszadók. A sávokat az alábbi ábra alapján töltjük ki, a kérdés típusa pedig a már előzőekben említett Szöveg (*Text*). Amennyiben számunkra ez a kérdés elengedhetetlen információt hordoz a kérdőív szempontjából, abban az esetben jelöljük a jobb alsó sarokban található „Kötelező kérdés” rubrikát (későbbiekben többször nem említjük, ahol kötelezőnek ítéljük meg a kérdést, mindenhol elvégzendő).

4/28. ábra: Betegkód megismerésére irányuló kérdőív elem

Ha a feljebb látható ábrán beállítottunk és kitöltöttünk mindent, a kék „Kész” gombra kattintva már létre is hoztuk első kérdésünket.

További kérdéseket mindig az „Elem hozzáadás” gyorsgombbal adhatjuk hozzá kérdőívünkhöz, ezen lépést minden egyes kérdésnél a továbbiakban nem említjük.

Második kérdésünk az életkor, mivel az életkor minden résztvevőnél eltérő, statisztikai szempontból folytonos változó és az összes életkort felsorolni listába értelmetlen volna, így az előző kérdéstípust alkalmazzuk. A sávokat az alábbi ábra alapján töltjük ki, a kérdés típusa pedig a már előzőekben említett „Szöveg” (*Text*) ismét. A kérdést ebben az esetben is a „Kész” gombra kattintva véglegesíthetjük (későbbiekben többször nem említjük, minden kérdésnél elvégzendő).



The screenshot shows a web-based survey creation tool. At the top, it says '1/1. oldal'. Below that is the title 'Terhesség alatti és utáni testsúlyváltozások' and a sub-label 'Űrlap leírása'. There is a field for 'Az Ön betegkódja:'. The main question editor shows 'Kérdés címe' with the text 'Az Ön életkora:', 'Súgószöveg' (empty), and 'Kérdés típusa' set to 'Szöveg'. There is a 'Válaszuk' section with a 'Speciális beállítások' link. At the bottom, there is a blue 'Kész' button and a checkbox for 'Kötelező kérdés'.

4/29. ábra: Életkor megismerésére irányuló kérdőív elem

Következő kérdésünk a kitöltő legmagasabb iskolai végzettségére irányul, ebben az esetben, válaszadás tekintetében öt lehetőséget kívánunk megkülönböztetni. A válaszadási lehetőségek a következők: 8 általános, érettségi, OKJ, felsőfokú szakképzés, főiskola, egyetem. Ez esetben „Feleletválasztós” (*Multiple Choice*) típusú kérdést alkalmazunk, hiszen azt szeretnénk, ha a kitöltő csak egy választ jelölne. Igény szerint és a válaszok megadásával alkalmazható a „Választás listából” (*Choose from a list*) mivel a válaszadás metodikája ezen két kérdés típusnál megegyezik. A példában a feleletválasztós verziót mutatjuk be. A kitöltést továbbra is az ábra alapján hajthatjuk végre.

The screenshot shows a question editor with the following fields and options:

- Kérdés címe:** Az Ön legmagasabb iskolai végzettsége:
- Súgószöveg:** (Empty)
- Kérdés típusa:** Feleletválasztós Ugrás a válasznak megfelelő oldalra
- Options:**
 - 8 általános
 - érettségi
 - OKJ
 - felsőfokú szakképzés
 - főiskola, egyetem
 - Opció hozzáadásához kattintson ide az egér vagy „Egyéb” hozzáadása
- Speciális beállítások:**
 - Képzés
 - Kötelező kérdés

4/30. ábra: A legmagasabb iskolai végzettség megismerésére irányuló kérdőív elem

Ezt követő kérdésünk a lakóhely típusára, vonatkozik. Válaszok tekintetében két opciót határoztunk meg: „város” illetve „falú”. Mivel a kérdést úgy fogalmazzuk meg, hogy „Honnan származik?” és nem kívánunk félreértést generálni, így azt egy súgószöveggel látjuk el, melyből már mindenki számára egyértelmű, hogy az mire vonatkozik. Azt szeretnénk, ha a kitöltő csak egy választ jelöljön ebben az esetben így az előző kérdéshez hasonlóan itt is a „Feleletválasztós” valamint a „Választás listából” kérdéstípust alkalmazhatjuk, de a szemléltetés valamint a válaszok számából adódóan most az utóbbit választjuk.

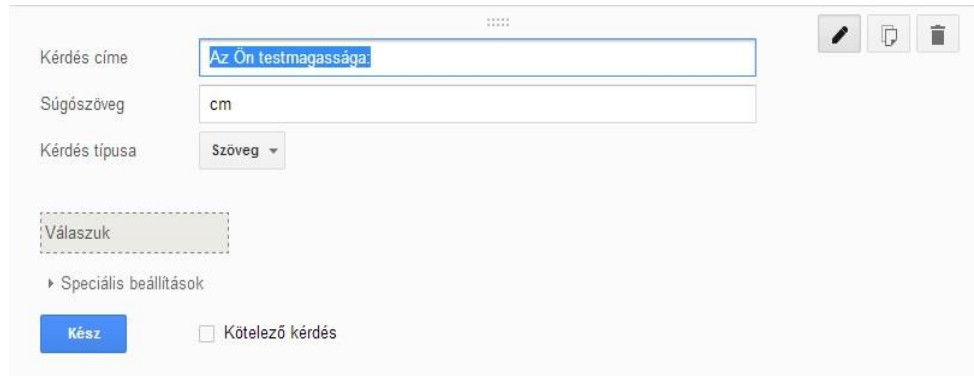
The screenshot shows a question editor with the following fields and options:

- Kérdés címe:** Honnan származik?
- Súgószöveg:** Milyen település típusban él jelenleg?
- Kérdés típusa:** Választás listából Ugrás a válasznak megfelelő oldalra
- Options:**
 1. falu
 2. város
 3. Opció hozzáadásához kattintson ide az egér
- Speciális beállítások:**
 - Képzés
 - Kötelező kérdés

4/31. ábra: Lakóhely típusára irányuló kérdőív elem

Mivel az 5, 6, 7, 8, 9-es számú kérdések mind folytonos változókra vonatkoznak, mint testsúly, testmagasság így ezeket egy kiválasztott kérdés megszerkesztésének bemutatásával kívánjuk demonstrálni, mivel a fent említett változó lévén „Szöveg” (*Text*) kérdéstípust kell alkalmaznunk minden esetben (1, 2 kérdéseknél bemutattuk ennek részletes szerkesztését).

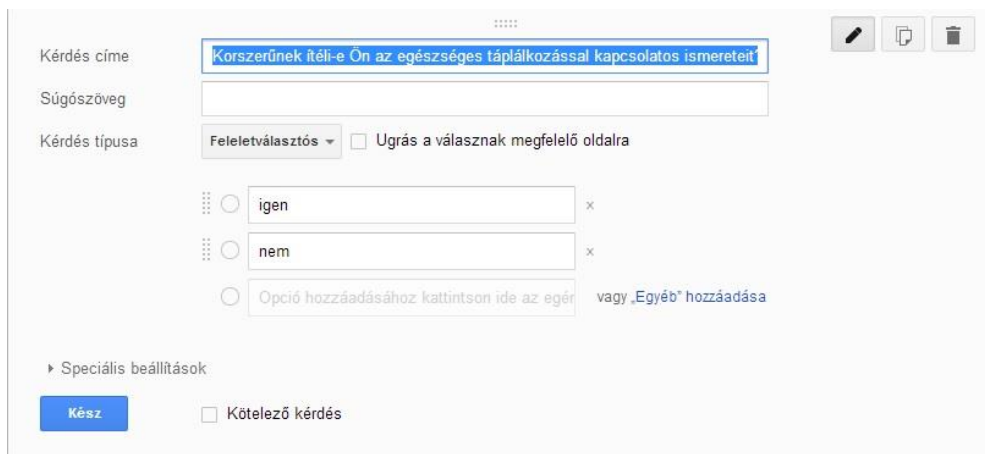
Fontos ugyanakkor megemlíteni, hogy a folytonos változók esetében, amennyiben szükséges a mértékegységet érdemes feltüntetni a súgószövegben, hiszen, ha valaki centiméterben, valaki pedig méterben adja meg testmagasságát az az adatok értékelhetőségét negatívan befolyásolhatja.



The screenshot shows a question editor interface. The 'Kérdés címe' (Question title) field contains 'Az Ön testmagassága'. The 'Súgószöveg' (Help text) field contains 'cm'. The 'Kérdés típusa' (Question type) dropdown is set to 'Szöveg'. Below the question type, there is a 'Válaszok' (Answers) section with a dashed border and a 'Speciális beállítások' (Special settings) link. At the bottom, there is a 'Kész' (Done) button and a checkbox for 'Kötelező kérdés' (Required question).

4/32. ábra: Testmagasságra irányuló kérdőív elem

A 10. kérdés, melyben arra vagyunk kíváncsiak, hogy az édesanyák, korszzerűnek ítélik-e meg egészséges táplálkozással kapcsolatos ismereteiket, az „igen” és a „nem” válaszopciót adtuk meg. Ebben az esetben „Feleletválasztós” (*Multiple Choice*) vagy „Választás listából” (*Choose from a list*) típusú kérdést alkalmazunk, hiszen azt szeretnénk, ha a kitöltő csak egy választ jelölne meg ennél a kérdésnél, hasonlóan a 3, 4 –es kérdésekhez.



The screenshot shows a question editor interface for a multiple choice question. The 'Kérdés címe' (Question title) field contains 'Korszzerűnek ítéli-e Ön az egészséges táplálkozással kapcsolatos ismereteit'. The 'Súgószöveg' (Help text) field is empty. The 'Kérdés típusa' (Question type) dropdown is set to 'Feleletválasztós'. There is a checkbox for 'Ugrás a válasznak megfelelő oldalra'. Below the question type, there are three radio button options: 'igen', 'nem', and 'Opció hozzáadásához kattintson ide az egér' (with a link to 'Egyéb* hozzáadása'). At the bottom, there is a 'Kész' (Done) button and a checkbox for 'Kötelező kérdés' (Required question).

4/33. ábra: Az egészséges táplálkozással kapcsolatos ismeretekre irányuló kérdőív elem

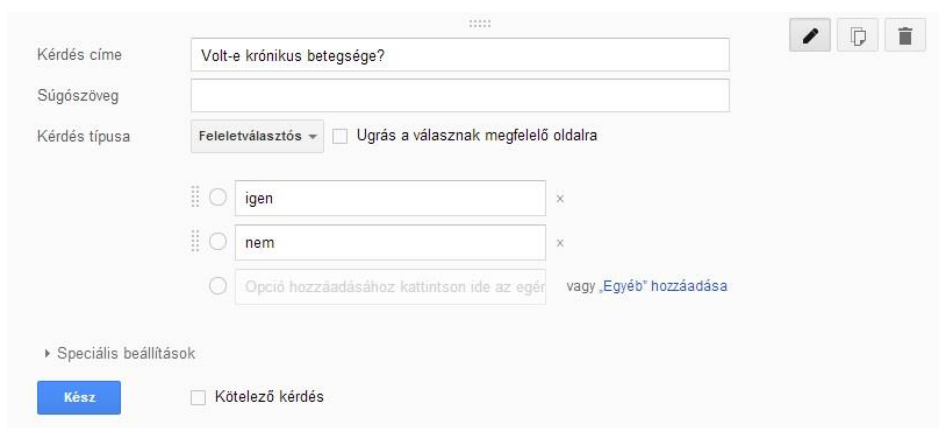
A 11. kérdés: Hányadik terhességi héten szült? - ugyancsak folytonos változóra irányul, így a már előzőekben bemutatott (5, 6, 7, 8, 9. kérdések) rövid, szöveges kérdés típust kell alkalmaznunk, ennek részletei a felül hivatkozott kérdéseknél megtekinthetőek. Szemléltetésképpen az ábra bemutatja a kitöltéssel kapcsolatos információkat.



The screenshot shows a question editor interface. At the top right, there are three icons: a pencil (edit), a document (copy), and a trash can (delete). Below these is a question title field containing "Hányadik terhességi héten szült Ön?". There is a placeholder text field for "Súgószöveg". The question type is set to "Szöveg" (Text) in a dropdown menu. Below this is a dashed box labeled "Válaszok" (Answers). There is a link for "Speciális beállítások" (Special settings). At the bottom left is a blue "Kész" (Done) button, and at the bottom right is a checkbox for "Kötelező kérdés" (Mandatory question).

4/34. ábra: Szülés időpontjára (terhességi hét) irányuló kérdőív elem

A 12. és 13. kérdés – Volt e krónikus betegsége („igen”, „nem”) illetve a Szülésének kimenetele („hüvelyi”, „császármetszés”) – válaszként mind a két esetben 2-2 válasz közül választhatunk. Ebben az esetben is hasonlóan járunk el mint a 3, 4, 10. kérdések esetében, így „Feleletválasztós” (*Multiple Choice*) típusú kérdést alkalmazunk. Igény szerint és a válaszok megadásával alkalmazható a „Választás listából” (*Choose from a list*) mivel a válaszadás metodikája ezen két kérdés típusnál megegyezik. A szemléltetés miatt az előbbi kérdést „Feleletválasztós” formában míg az utóbbit a „Választás listából” kérdéstípus segítségével mutatjuk be.



The screenshot shows a question editor interface. At the top right, there are three icons: a pencil (edit), a document (copy), and a trash can (delete). Below these is a question title field containing "Volt-e krónikus betegsége?". There is a placeholder text field for "Súgószöveg". The question type is set to "Feleletválasztós" (Multiple choice) in a dropdown menu. There is a checkbox for "Ugrás a válasznak megfelelő oldalra" (Jump to the page corresponding to the answer). Below this are two radio button options: "igen" (yes) and "nem" (no), each with a text input field and a delete icon (x). There is also a radio button for "Opció hozzáadásához kattintson ide az egér vagy „Egyéb” hozzáadása" (Click here to add an option or "Other"). There is a link for "Speciális beállítások" (Special settings). At the bottom left is a blue "Kész" (Done) button, and at the bottom right is a checkbox for "Kötelező kérdés" (Mandatory question).

4/35. ábra: Előzetes krónikus betegségekre irányuló kérdőív elem

The screenshot shows a question editor with the following fields and options:

- Kérdés címe:** Szülés kimenetele:
- Súgószöveg:** (empty)
- Kérdés típusa:** Választás listából Ugrás a válasznak megfelelő oldalra
- Options:**
 - 1. hüvelyi x
 - 2. császármetszés x
 - 3. Opció hozzáadásához kattintson ide az egér
- Speciális beállítások:**
 - Kötelező kérdés
- Buttons:** Kész

4/36. ábra: A szülés kimenetelére irányuló kérdőív elem

A 14. egyben utolsó kérdésünkben arra voltunk kíváncsiak, hogy a szülés alatt mennyire érzett fájdalmat az anya, ezt az eddig még a konkrét példán be nem mutatott „Tartomány” (*Scale*) kérdéstípust alkalmazva szemléltetjük.

Ennél az elemnél egy adott tartományt, skálát adhatunk meg –tól –ig értékek megadásával, mely 1-5-ig értékelhető, attól függően, hogy a fájdalom érzet milyen erősnek ítélte meg a szülő anyja. Két végpontunk az „egyáltalán nem” és az „abszolút, teljes mértékben”. A szükséges beállításokat az alábbi ábra szemlélteti.

The screenshot shows a question editor with the following fields and options:

- Kérdés címe:** Mennyire volt fájdalmas szülése:
- Súgószöveg:** Szülés alatt érzett fájdalom érzet erősségét határozza meg
- Kérdés típusa:** Tartomány
- Tartomány:**
 - 1: egyáltalán nem
 - 5: abszolút, teljes mértékben
- Buttons:** Kész Kötelező kérdés

4/37. ábra: A szülés alatt érzett fájdalomérzetre irányuló kérdőív elem

Ezzel előre meghatározott példa kérdőívünk szerkesztésének végére értünk, ezután további feladatunk a kérdőívek kiküldése, esetleges beágyazása valamint a válaszok fogadás és kiértékelése.

Mivel a bemutatás és szemléltetés során a kész kérdőív elemeket nem mutattuk be, mert a szerkesztési lépések prioritást élveztek. Ezért ezt a következőkben a teljes mértékben elkészült kérdőíven lehet megtekinteni.

Űrlap címe

Terhesség alatti és utáni testsúlyváltozások

Az Ön betegkódja:

Az Ön életkora:

Az Ön legmagasabb iskolai végzettsége:

- 8 általános
- érettségi
- OKJ
- felsőfokú szakképzés
- főiskola, egyetem

Honnan származik?

Milyen település típusban él jelenleg?

Az Ön testmagassága: cm

Az Ön várandósság előtti testsúlya: kg

Az Ön testsúlya a szülés napján: Szülést megelőzően, kg

Az Ön gyermekének születési súlya: kg

A Ön testsúlya a szülés utáni 6. hónapban: kg

Korszerűnek ítéli-e Ön az egészséges táplálkozással kapcsolatos ismereteit?

- igen
- nem

Hányadik terhességi héten szült Ön?

Volt-e krónikus betegsége?

- igen
- nem

Szülés kimenetele:

Mennyire volt fájdalmas szülése: Szülés alatt érzett fájdalom érzet erősségét határozza meg

1 2 3 4 5

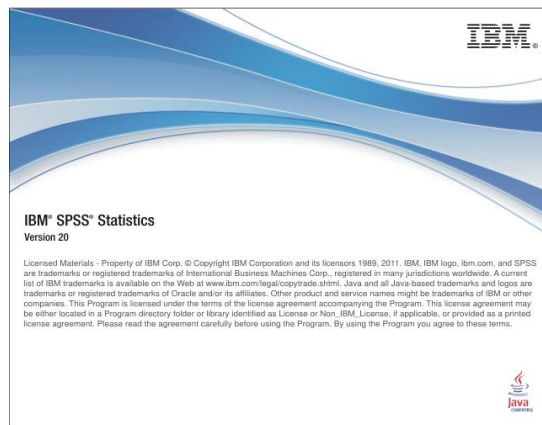
egyáltalán nem abszolút, teljes mértékben

5. AZ SPSS PROGRAM FELÜLETE, AZ ADATOK BEVITELE ÉS EXPORTÁLÁSA, MENÜPONTOK BEMUTATÁSA (Ács Pongrác)

5.1. Az SPSS program felülete

Az SPSS (Statistical Package for the Social Sciences) megjelenése amerikai és kanadai gyökerekkel 1968-ra tehető. Jelenleg elmondható, hogy talán a legismertebb piacvezető statisztikai szoftverként emlegetik világszerte, mely a korai felismerés és megjelenésnek is köszönhető. Alapjában a társadalomtudományi kutatásokat célozta meg és a tudományos kutatások alapeszközévé vált, melyet az egyetemi oktatásban is integráltak. Mindezekon felüli előnye a felhasználóbarát kialakítása, valamint a statisztikai analízisek megfelelő csoportosítása és vizuális megjelenítése. Miután ez egy valóban statisztikai elemzések céljából készített program, így nem okoz nehézséget – az Excel programmal ellentétben – az óriási adatbázisok kezelése és a velük való számítások elvégzése. A program magyar nyelven nem elérhető, viszont a könnyű áttekinthetősége révén használata nem igényel emelt szintű angol nyelvtudást. Beállítási lehetőségei rendkívül változatosak, használatát mégis hamar meg lehet szokni. Oktató és súgó moduljai jól használhatók és nagyon részletesek. Grafikája nem túl látványos, nem tartozik az erősségei közé, de tág határok között állítható. 2005-ig a magyar felsőoktatás oktatási-kutatási célokra – dátumkódos korláttal – ingyenesen használhatta a programcsomagot. 2005 őszén ezt a rendszert megszüntették, illetve teljesen átszervezték, de az egyetemek továbbra is kedvezményesen juthatnak hozzá a programcsomaghoz. Jelenleg elmondható, hogy szinte minden évben egy újabb verzióval jelentkeznek, azonban ezeknek az alapjai azonosak, csak apróbb „esztétikai” módosítások, időnként egy új módszertan integrálása történik. Számtalan előnye mellett kiemelendő, hogy más programokkal készített és különböző formátumokban mentett adatbázisokat képes konvertálni, megnyitni (Excel, dBase, Lotus, SAS, Stata stb) és menteni. Az érdeklődő a programcsomagról és a cégről további információkat a www.spss.com és a www.spss.hu weboldalakon találhat.

A továbbiakban a számunkra elérhető SPSS 20 verzióval fogunk dolgozni, így a statisztikai elemzéseket is ezen keresztül fogjuk bemutatni. Ismételten jelezni kívánjuk, hogy az elkövetkezőkben bemutatásra kerülő eljárások mindegyik verzióban megtalálhatóak, legfeljebb a beállítási modulokban fogunk apróbb különbségeket találni.



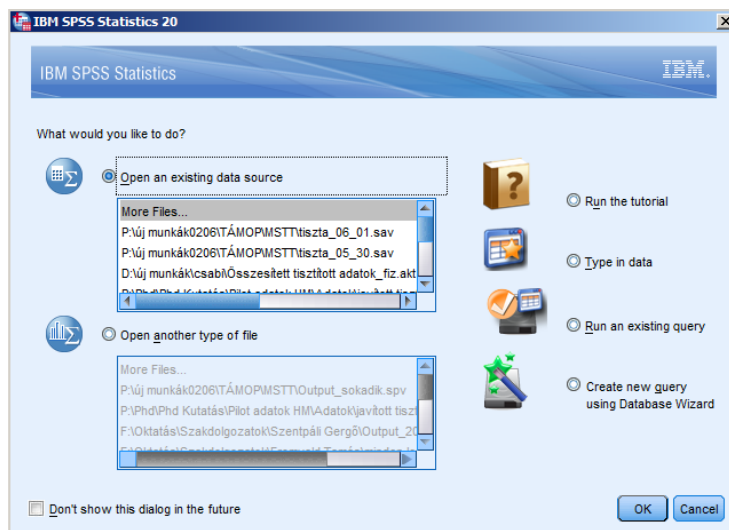
5/1. ábra: A program bejelentkezése során megjelenő információ

Az SPSS program a Windows alapú programokhoz hasonlóan többablakos technikával került kialakításra, ami azt jelenti, hogy külön ablakban és formátumban kezeli az adatbázist, az eredményeket (output) és külön ablakban szerkeszthetők a synatoxok is, amely az SPSS belső nyelve és segítségével utasításokat tárolhatunk és futtathatunk. A tankönyv írása során nem a belső nyelv és utasítások ismertetését célozzuk, erről az olvasó részletesen olvashat a Túlélőkészlet az SPSS-hez című könyvben (Székelyi – Barna 2005). Mi a program menürendszerének megismerését követően (*File, Edit, Data, Analyze*) a dialógusablakok kitöltésén és opciók kipipálásának segítségével fogjuk az alapvető gyakorlati statisztikai módszertanokat megismerni. Megpróbáljuk első lépésben a menürendszerben legfontosabbnak vélt utasításokat - a teljesség igénye nélkül - bemutatni. A terjedelmi korlátokra hivatkozva az esetlegesen kimaradó opciók alkalmazásának megismerését az SPSS program kiváló *HELP* menüjén keresztül az érdeklődő megteheti.

Az SPSS program indításakor az első párbeszéd ablak négy lehetőséget kínál az első lépés meghatározására:

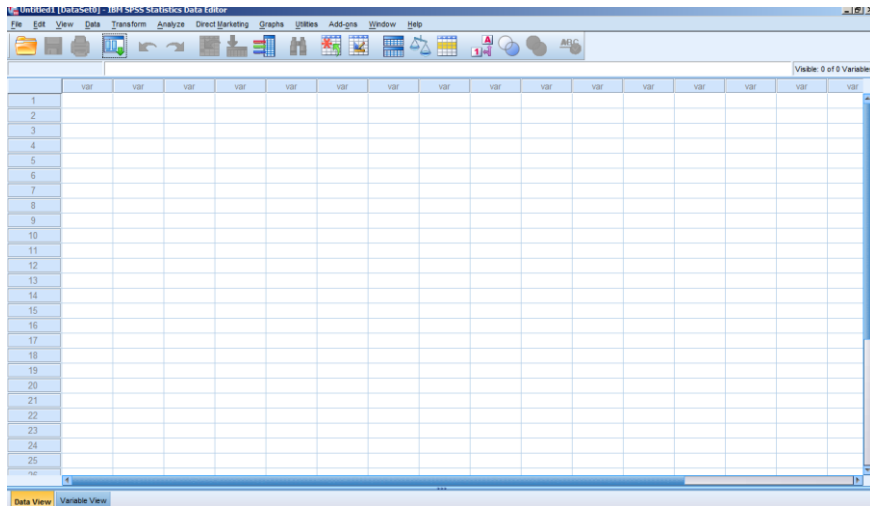
- Egy oktatóprogram indítása
- Adatok begépelése
- Egy létező lekérdezés nyitása
- Adatok egy másik adatbázisból való megnyitásával.

Az *Open an existing data source* választását követően egy korábban használt adatbázist nyithatunk meg. Az SPSS jelen verziójával használt adatbázisaink kiterjesztése: *.sav, míg az SPSS által készített *OUTPUT* eredményeket tartalmazó *.spv.

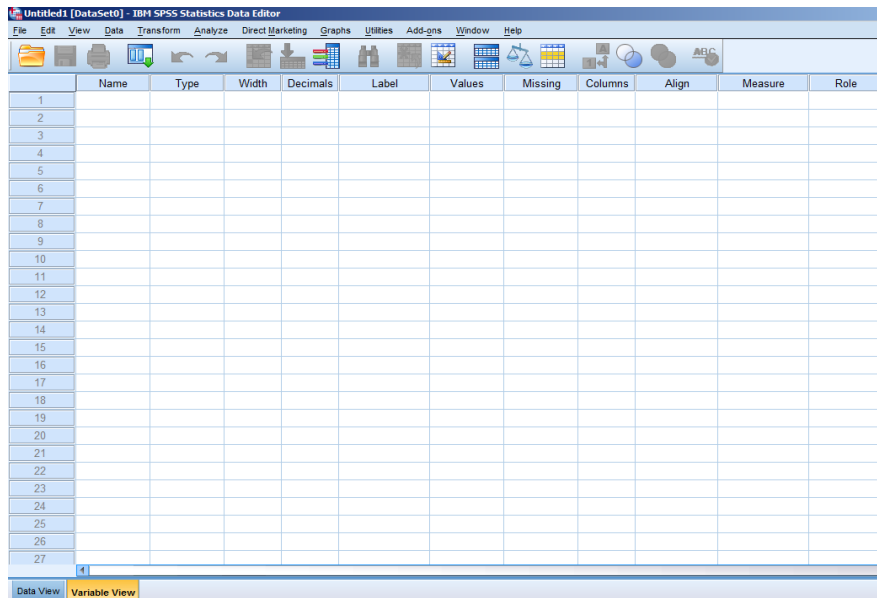


5/2. ábra: Bejelentkezést követő párbeszédpanel

Amennyiben a *Cancel* gombra klikkelünk egy üres adatbázist kapunk, amely menürendszer felépítés tekintetében teljesen hasonlít a Microsoft Office programcsomagokban megismert struktúrához. A már ott megismert alkalmazások (pl.: kivágás, beillesztés, másolás, stb.) itt is megtalálható és alkalmazható. Az SPSS-ben az adatszerkesztő (*DATA EDITOR*) modulban két megjelenítési formát találunk (*DATA VIEW*: adat-, és *VARIABLE VIEW* változónézet). Az adatszerkesztő két nézete között az ablak bal alsó sarkában található fülekre való kattintással, illetve a CTRL+T billentyűk megnyomásával válthatunk. Az adatbázisunk primer adatokkal való feltöltése az Excel programhoz hasonlóan történik az adatnézet (*DATA VIEW*) ablakban (5/3a. ábra). Az adatbázisunkat esetekkel (rekordokkal) sorokkal (*CASE*), illetve oszlopokkal, vagyis változókkal (*VARIABLE*) jellemezhetjük. Ez jelentheti, hogy a sokaság egy egyedét a sorok (*CASES*), míg az egyes változókat (*VARIABLE*) az oszlopok jelentik. Az adatszerkesztő két nézete között az ablak bal alsó sarkában található fülekre való kattintással, illetve a CTRL+T billentyűk megnyomásával válthatunk.



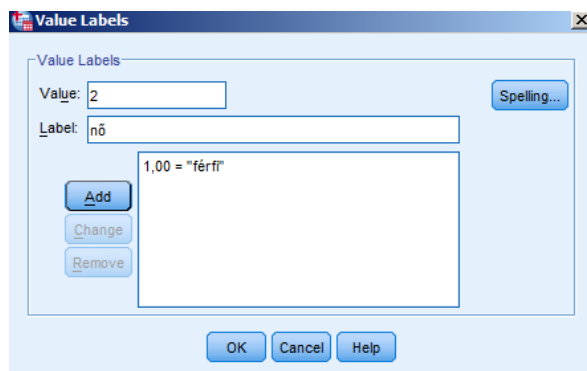
5/3a. ábra. Adatnézet



5/3b. ábra: Változónézet

A változónézet (5/3b. ábra) panelen tudjuk az ismérveink paramétereit megadni, változtatni, szerkeszteni. A sorok a változókat, míg az oszlopok a változók paramétereit jelölik. *NAME*: a változó nevének rövidített változatát tartalmazza. A változónév betűvel kezdődik és maximálisan 64 karakter hosszúságú lehet, és nem érdemes ékezetes és speciális karaktereket alkalmazni. Amennyiben mi nem definiáljuk a változót, akkor a program automatikusan VAR0000 kezdéssel, folyamatos növeléssel ezt megteszi. Érdemes az áttekinthetőség kedvéért is rövid neveket adni, melyet bővebben a címke (*LABEL*) opcióban definiálhatunk. A *TYPE* oszlopban a változónk típusát tudjuk meghatározni. A típust a cella jobb sarkából klikkelésre előjövő megadott nomenklatúrából kell, minden változónál külön-külön megadni. Leggyakrabban a szám (*NUMERIC*) és szöveg (*STRING*) formátumot szoktuk jelölni, de ezen kívül lehetőség van még dátum, pénzügyi és egyéb formák kiválasztására is. *WIDTH*

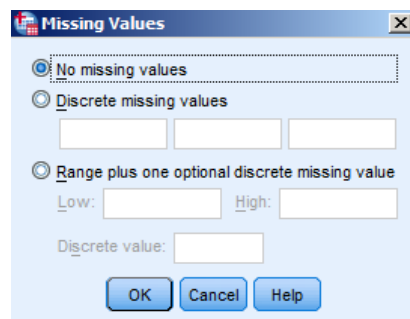
oszlopban a változók karakterszámát adjuk meg, ami szöveges változónál maximum 255, és számoknál maximum 40 karakter hosszúságot engedélyez. *DECIMALS* oszlopban a megjelenítésre kerülő tizedesek számát adjuk meg. A *LABEL*, címke oszlopban a változók bővebb definiálása történik, amit az eredményközlésben megjelenő *OUTPUT* táblákban vizionlátunk. Itt lehetőség van, akár a kérdőívben megjelenő teljes kérdés begépelésére is. A *VALUES* oszlopban a nominális és ordinális adatainknak címet adhatunk. Itt adhatunk pl. a Likert skála 1-5 értékeinek elnevezést (1-5). Amennyiben az egyik kérdés a nemre vonatkozót, akkor itt nevezhetjük meg az egyes ismérvtértekhez tartozó címkéket.



5/4. ábra: Az ismérvtértek elnevezése, címkézése

A *VALUE* cellában tudjuk az értéket megadni, míg a *VALUE LABEL* cellában a szöveges elnevezésre van lehetőség. Minden egyes ismérvtértek szöveges rögzítését követően az *ADD* gombot megnyomni szükséges. A szöveges megjegyzést a *CHANGE* gombbal tudjuk változtatni, míg a *REMOVE*-val törölhetünk.

A *MISSING* oszlopban a hiányzó értékek definiálása történik, a megfelelő feltételek megadását követően. Fontos, hogy a különböző okokból hiányzó vagy tévesen felvitt adatokat megjelöljük, hogy a statisztikai adatfeldolgozás és eredmény valid lehessen. Definiálnunk kell azokat az értékeket, amelyek nem fordulhatnak elő.

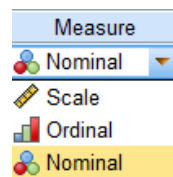


5/5. ábra: A hiányzó értékek beállításai

A felnyíló párbeszédpanelben három lehetőség áll rendelkezésre:

- *No missing values* jelenti, hogy nem adunk meg hiányzó értékeket. Ilyenkor a kihagyott értékek helyén egy pont jelenik meg. Ez az alapbeállítás, ha nincsen különösebb okunk a definiálásra, akkor ezt javasoljuk használni.
- *Discrete missing values*: itt maximálisan három konkrét értéket, hiányzó adat értéket tudunk megadni (999 nincsen válasz).
- *Range plus one optional discrete missing value* esetében lehetőségünk van egy tartomány megadására, illetve még diszkrét érték megadásra is.

COLUMNS (oszlopok) határozhatjuk meg az adatok adatképernyőjében (*DATA VIEW*) megjelenített oszlopszélességét. *ALIGN* oszlop az ismérvértékek cellában való jobbra, balra vagy középre igazítását szolgálja. A *MEASURE* (mérési skála) oszlopban a már korábban bemutatott (5/1. ábra) és kiemelten fontos skála típusokat (metrikus, sorrendi, nominális) kell megadni. Ismételten meg kell jegyezni, hogy az SPSS program nem különbözteti meg az intervallum és arányskálát, közösen *SCALE*-ként kezeli.



5/6. ábra: Skálátípusok megadása

Az *OUTPUT VIEW* az SPSS programba az eredmények megjelenítésre, mentésére és szerkesztésére szolgáló nézet. Ehhez a nézethez saját menü és eszközsor tartozik. A megjelenítésnél két rész különül el. A baloldalon a vizsgálatokat, elemzéseket tartalmazó struktúra-fa jelenik meg. Jobb oldalon a számítási eredményeket és részeredményeket tartalmazó táblázatok, ábrák, eredmények jelennek meg. Amennyiben a megjelenő táblázatra, ábrára kétszer klikkelünk az egyedi formázás veheti kezdetét. Amennyiben az eredményeket tartalmazó objektumokat másolni szeretnénk, akkor (CTR+C) és (CTR+V) billentyűk használhatóak.

5.2. Adatbevitel

Az SPSS programban az adatok bevitelének két módját különböztethetjük meg:

- *Elsődleges adatbevitel*, amikor az adatokat beírjuk a cellákban. Első lépésben meg kell határozni a változóinkat a változónézetben, majd ezt követően az adatképernyőn begépeljük az adatokat.

- *Másodlagos adatbevitel* során a már létező adatbázis SPSS-be történő importálása történhet.

Az *elsődleges adatbevittelt* egy egyszerű példánál maradva kívánjuk szemléltetni. Egy betegosztályon megjelenő új betegek adatait kell rögzíteni néhány kiemelt ismérv (nem, életkor, lakhely, szisztolés vérnyomás érték felvételkor, diasztolés érték felvételkor) alapján.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	nem	Numeric	8	2	A beteg neme	{1,00, férfi}...	None	8	Right	Nominal	Input
2	kor	Numeric	8	2	A beteg életkora	None	None	8	Right	Scale	Input
3	lakhely	Numeric	8	2	A beteg lakhelye	{1,00, megy}...	None	8	Right	Ordinal	Input
4	diast	Numeric	8	2	Diasztolés vér...	None	None	8	Right	Scale	Input
5	systol	Numeric	8	2	Szisztolés vér...	None	None	8	Right	Scale	Input
6											
7											
8											
9											
10											
11											
12											
13											
14											

5/7. ábra: Elsődleges adatbevitel során történő változók definiálása

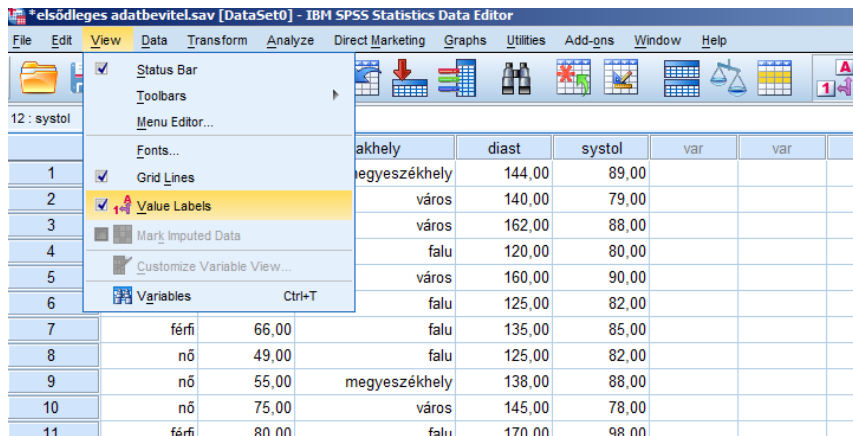
A paraméterek meghatározása (első lépés) során a nem és lakhely változókat elneveztük (*VALUES*). A nem változó alkalmával 1-es kóddal a férfiakat, 2-essel a nőket jelöltük. A lakhely vonatkozásában három kategóriát képeztünk, mely a lélekszámot is tükrözi (1-es érték megyeszékhely, 2-es város, 3-as falu). Jelen esetben nem definiáltuk a hiányzó értékeket. A nem és lakhely diszkrét változók, melyeknél a mérési skálát nominálisként (nem) és ordinálisként definiáltuk. A többi folytonos változónk végtelen számú ismérvváltozóval rendelkezhet, így a mérési szintet skála típusban határozzuk meg.

Ezt követően váltunk adatnézetre (*DATA VIEW*) és töltjük fel esetekkel (rekordokkal) az előbbieken definiált változóinkat.

	nem	kor	lakhely	diast	systol	var	var	var	var	var
1	1,00	47,00	1,00	144,00	89,00					
2	2,00	65,00	2,00	140,00	79,00					
3	1,00	72,00	2,00	162,00	88,00					
4	1,00	51,00	3,00	120,00	80,00					
5	2,00	60,00	2,00	160,00	90,00					
6	2,00	63,00	3,00	125,00	82,00					
7	1,00	66,00	3,00	135,00	85,00					
8	2,00	49,00	3,00	125,00	82,00					
9	2,00	55,00	1,00	138,00	88,00					
10	2,00	75,00	2,00	145,00	78,00					
11	1,00	80,00	3,00	170,00	98,00					

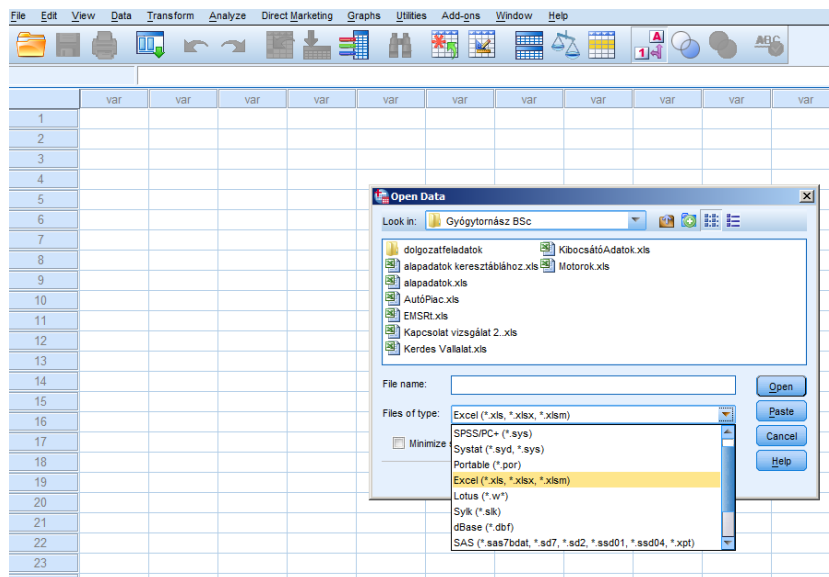
5/8. ábra: Az adatok feltöltése

Amennyiben a *VIEW* menüpontban a *VALUE LABELS* opciót jelöljük, akkor azt követően a számérték helyén az általunk megadott címkék fognak megjelenni.



5/9. ábra Value label opció kiválasztása

A **másodlagos adatbevitel** során már meglévő adatbázist nyitunk meg az SPSS program segítségével. A korábbiakban már említettük, hogy nem kizárólag a programmal lementett adatbázisokat tudunk importálni, hanem más formátumú (leggyakrabban Excel) állományokat is. Amennyiben xls kiterjesztésű (Excel) file szeretnénk betölteni, abban az esetben a bal felső sorban található File menü *Open Data* ablakának legördülő *File of type* beállítása során válasszuk a számunkra megfelelő típust, (pl.: Excel).

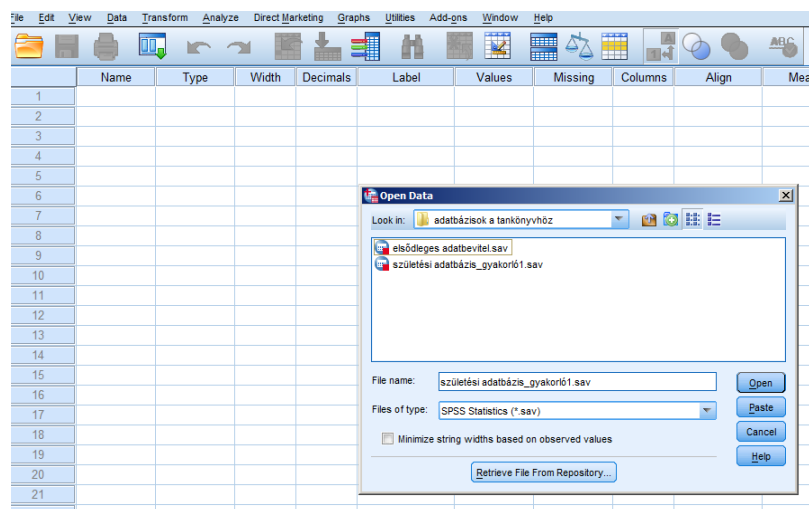


5/10. ábra: A adatok importálásának beállításai

Fontos megemlíteni, hogy mielőtt az Excel file-okat importáljuk ellenőrizzük le, hogy éppen az Excel program nem használja-e, mivel ebben az esetben az SPSS nem képes megnyitni. Abban az esetben, ha szeretnénk egy időben mindkét programmal az adatbázis használni,

akkor először az SPSS programmal kell megnyitni, melyet az Excel követhet. Az adatokat tartalmazó file kiválasztását követően meg kell adni, hogy melyik munkalap adatai kerüljenek az SPSS-be. Alapértelmezésben természetesen az SPSS *sav formátumát látjuk.

A következőkben importáljuk a *sav formátumban elmentett adatbázisunkat (születési adatbázis_gyakorlól1_teljes_67), melyet a folytatásban, mint szemléltető példa fogunk felhasználni. A következő adatbázis a felhasználók számára elérhető a www.etk.pte.hu weboldalon.



5/11. ábra: A gyakorló adatbázis importálása

Az *OPEN* gomb lenyomását követően az adatbázis importálása megtörténik és a *VARIABLE VIEW* nézetben a változóink tulajdonságai láthatóvá válnak. Az adatbázisban 13 változó szerepel, mely 67 kismama válaszait tartalmazza a szülés után 6. hónapban. Az adatok a betegazonosító kódot (sorszám), életkort, végzettséget, származási helyet, szülési kimenetet és egyéb a szüléshez kapcsolható információkat hordozzák.

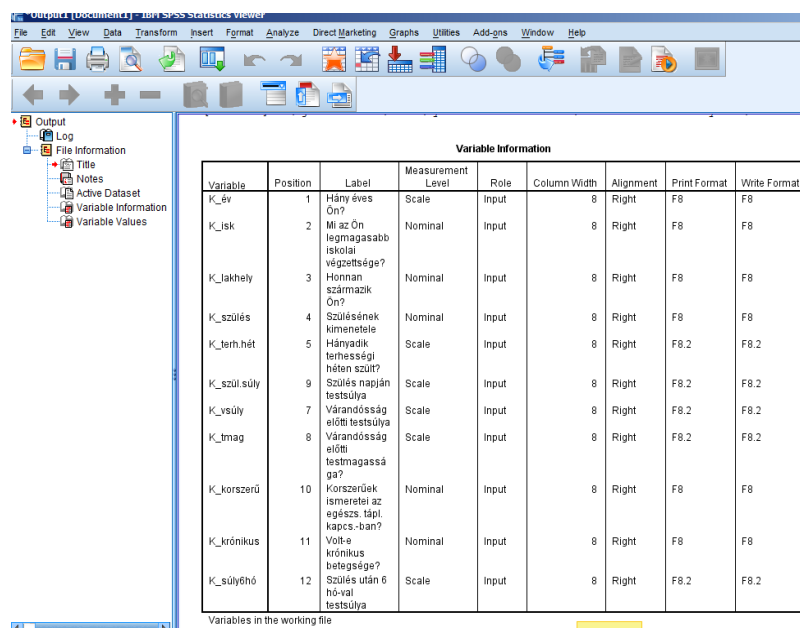
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	K_kód	numeric	8	0	Betegkód	None	None	8	Right	Nominal	Input
2	K_év	Numeric	8	0	Hány éves Ön?	None	None	8	Right	Scale	Input
3	K_isk	Numeric	8	0	Mi az Ön legmagasabb iskolai végzettsége?	{1, egyetem...	None	8	Right	Nominal	Input
4	K_lakhely	Numeric	8	0	Honnan származik Ön?	{1, város}...	None	8	Right	Nominal	Input
5	K_szülés	Numeric	8	0	Szülésének kimenetele	{1, hüvelyi}...	None	8	Right	Nominal	Input
6	K_terh_hét	Numeric	8	2	Hányadik terhességi héten szült?	None	None	8	Right	Scale	Input
7	K_szül_súly	Numeric	8	2	Gyermekeének születési súlya?	None	None	8	Right	Scale	Input
8	K_vsúly	Numeric	8	2	Várandósság előtti testsúlya	None	None	8	Right	Scale	Input
9	K_tmag	Numeric	8	2	Várandósság előtti testmagassága?	None	None	8	Right	Scale	Input
10	K_sszul	Numeric	8	2	Szülés napján testsúlya	None	None	8	Right	Scale	Input
11	K_korszerü	Numeric	8	0	Korszerűek ismeretei az egészsz. tápl. kapcs.-ban?	{1, igen}...	None	8	Right	Nominal	Input
12	K_krónikus	Numeric	8	0	Volt-e krónikus betegsége?	{1, igen}...	None	8	Right	Nominal	Input
13	K_súlybó	Numeric	8	2	Szülés után 6 hó-val testsúlya	None	None	8	Right	Scale	Input

5/12. ábra: születési adatbázis_gyakorlól1_teljes_67 változóinak paraméterei

5.3. Menüpontok bemutatása

A következő rövid részben a teljesség igénye nélkül rövid áttekintést aduk a legfontosabbnak vélt menüpontokról.

File menüben a szokásos adatokkal kapcsolatos feladatokat végezhetjük el, mint új és meglévő adatbázisok megnyitása (adat, output, text file), mentése, átnevezése. A **DISPLAY DATA FILE INFORMATION** segítségével a *.sav formátumban mentett file változóiról ad egy összesített információ táblát az **OUTPUT** eredmények között. Jelen beállításoknál a **WORKING FILE** opciót választottuk, vagyis a jelenleg betöltött adatbázis változóiról (születési adatbázis_gyakorló1_teljes_67.sav) kapunk összesítést.



The screenshot shows the SPSS Variable Information dialog box. The table below is a reproduction of the data shown in the dialog box.

Variable	Position	Label	Measurement Level	Role	Column Width	Alignment	Print Format	Write Format
K_év	1	Hány éves Ön?	Scale	Input	8	Right	F8	F8
K_isk	2	Mi az Ön legmagasabb iskolai végzettsége?	Nominal	Input	8	Right	F8	F8
K_lakhely	3	Honnan származik Ön?	Nominal	Input	8	Right	F8	F8
K_szülés	4	Születésének kimenetele	Nominal	Input	8	Right	F8	F8
K_terh.hét	5	Hányadik terhességi hetén szült?	Scale	Input	8	Right	F8.2	F8.2
K_szul.súly	9	Születés napján testsúlya	Scale	Input	8	Right	F8.2	F8.2
K_vsúly	7	Várandósság előtti testsúlya	Scale	Input	8	Right	F8.2	F8.2
K_tmag	8	Várandósság előtt testmagassága?	Scale	Input	8	Right	F8.2	F8.2
K_korszerű	10	Korszerűek ismeretei az egészsz. tápl. kapcs.-ban?	Nominal	Input	8	Right	F8	F8
K_krónikus	11	Volt-e krónikus betegsége?	Nominal	Input	8	Right	F8	F8
K_súlyöho	12	Születés után 6 hó-val testsúlya	Scale	Input	8	Right	F8.2	F8.2

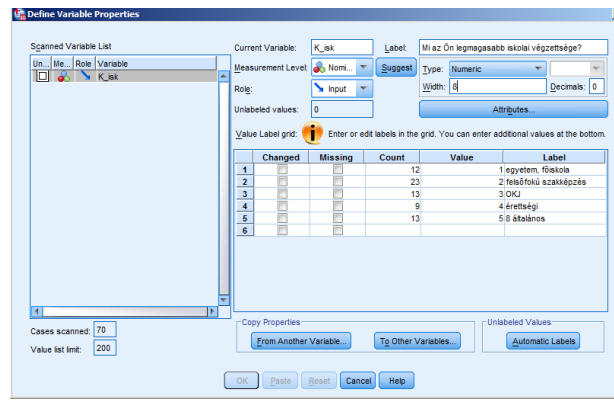
5/13. ábra: Az adatbázis változóinak összefoglaló jelentése

Az összefoglalóban információt kapunk a változó nevééről (**VARIABLE**), pozíciójáról (**POSITION**), címkéjéről (**LABEL**), mérési szintjéről (**MEASUREMENT LEVEL**), stb.

Edit menü segítségével adatszerkesztési opciókat találunk. A Microsoft Office programcsomagokból jól ismert elemek találhatók itt: másolás, beillesztés, kivágás, törlés, keresés. Az **INSERT VARIABLE** segítségével új változót (oszlopot) illeszthetünk be, míg az **INSERT CASE** segítségével új eset (új sor) beillesztésére nyílik lehetőségünk. Az **OPTIONS** menüpont segítségével az ablakok és változók beállítása történhet. A legfontosabb a **GENERAL** fül, ahol lehetőségünk van eldönteni, hogy a statisztikai elemzések és eredmények közlésénél a változók neve (**DISPLAY NAME**), vagy azok teljes jelentése (**DISPLAY LABELS**) jelenjen meg. Amennyiben a statisztikai adatok elemzésekor nem tudjuk, hogy változó milyen jelentéssel bír, akkor javasoljuk a **DISPLAY LABELS** opciót alkalmazni.

View menü beállításainak segítségével tudjuk az aktív ablakot saját igényünk szerint kalibrálni. Itt van lehetősége a felhasználónak a saját munkájához leginkább használt ikonok, feliratok megjelenítésre.

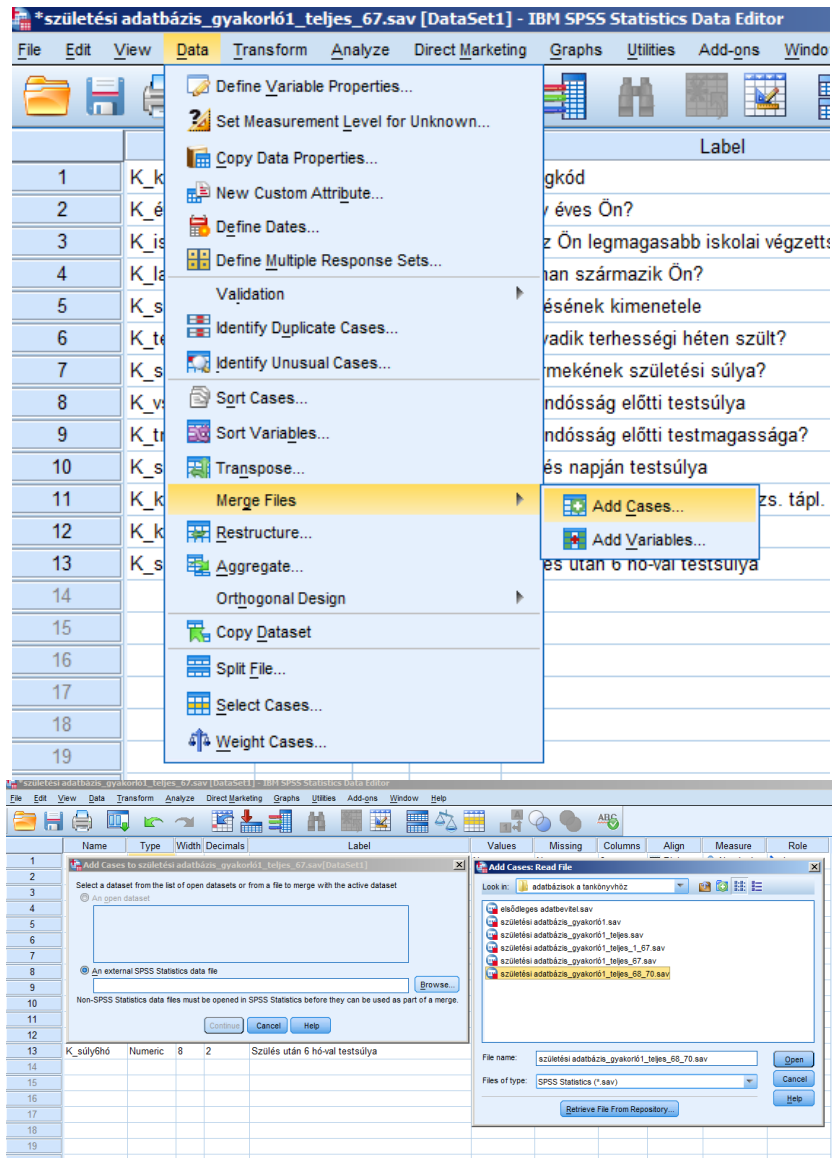
Data menü az adatok kezelésére és tagolására szolgál. A **DEFINE VARIABLE PROPERTIES** az ismérvváltozatok meghatározására és gyakoriságaik (**COUNT**) megjelenítésre szolgál.



5/14. ábra: A változók és ismérvváltozataik beállításai

SORT CASES az esetek általunk kalibrált szempontjai az egy vagy több változó csökkenő vagy növekvő sorba rendezését lehet végrehajtani. **TRANSPOSE** segítségével az adatbázis transzponálása, vagyis az oszlopok és sorok felcserélésére van lehetőség. **MERGE FILES** segítségével lehetőség nyílik több állomány vagy kérdőív eseteinek és változóinak összefűzésére, illetve bővítésére. Abban az esetben, ha változóink megegyeznek - csak az esetekkel kell bővíteni az adatbázisunkat -, ilyenkor az **ADD CASES** lehetőséget válasszuk. Abban az esetben, ha az esetek egyeznek, de más változókat szerepeltetünk róluk, akkor az esetek helyett az új változókat illesztjük be.

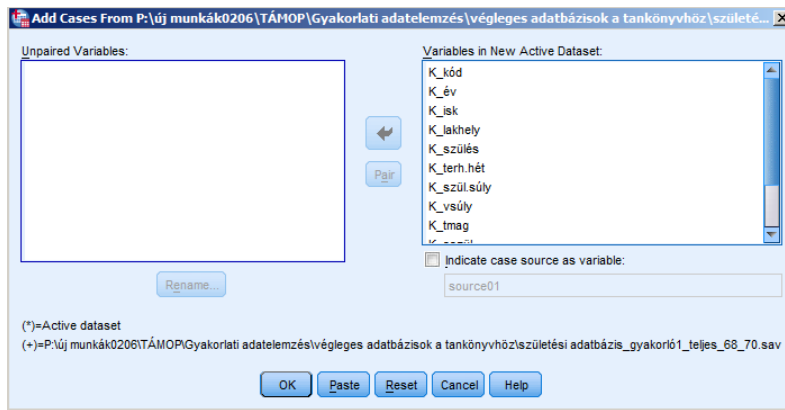
Gyakorló feladatként három kismama utólag felmért adatait (születési adatbázis_gyakorló1_teljes_68_70) illesszük a már meglévő adatbázisunkhoz.



5/15. ábra: Az adatbázisok (esetek) összefűzésének beállításai

Mivel három kismama változói megegyeznek ezért az esetek hozzáadását választottuk. A *FILE NAME* cellába kerül a hozzáfűzni kívánt adatbázis. Ezt követően megjelennek az ablakban azon változók, amelyeket az adatbázisok egyformán⁹ tartalmaznak.

⁹ Nagyon fontos, hogy a változó nevekben nem lehet karakter eltérés, különben a program nem tudja összefűzni.



5/16. ábra: A létrejövő új adatbázis közös változóinak halmaza

Az elfogadást követően (OK gomb) létrejön a közös adatbázis, amelyik immáron 70 kismama adatait tartalmazza.

	K_kód	K_év	K_isk	K_lakhely	K_szülés	K_telh.hét	K_szül.súly	K_vsúly	K_tmag	K_szül	K_koraszoró	K_korónkus	K_súlyfőh.
46	46	37	OKJ	falu császárme...	39.00	3500.00	65.00	158.00	75.00	nem	Igen	Igen	74.00
47	47	37	OKJ	falu császárme...	38.00	3300.00	64.00	155.00	76.00	nem	Igen	Igen	65.00
48	48	38	felsőfokú s...	falu császárme...	38.00	3200.00	63.00	154.00	73.00	Igen	Igen	Igen	70.00
49	49	39	felsőfokú s...	falu császárme...	38.00	3400.00	62.00	154.00	75.00	Igen	Igen	Igen	70.00
50	50	33	felsőfokú s...	falu császárme...	38.00	3560.00	61.00	149.00	70.00	Igen	Igen	Igen	73.00
51	51	32	felsőfokú s...	város császárme...	39.00	3460.00	60.00	166.00	75.00	Igen	Igen	Igen	65.00
52	52	22	felsőfokú s...	falu	hüvelyi	40.00	3470.00	59.00	166.00	78.00	Igen	Igen	62.00
53	53	23	8 általános	város	hüvelyi	37.00	3740.00	58.00	168.00	80.00	Igen	Igen	75.00
54	54	24	8 általános	város	hüvelyi	37.00	3850.00	57.00	170.00	80.00	Igen	Igen	60.00
55	55	25	8 általános	város	hüvelyi	38.00	3900.00	56.00	170.00	66.00	Igen	nem	58.00
56	56	26	8 általános	város	hüvelyi	39.00	4050.00	54.00	171.00	70.00	Igen	nem	60.00
57	57	27	8 általános	város	hüvelyi	39.00	2900.00	52.00	171.00	72.00	Igen	nem	60.00
58	58	28	egyetem. f...	város	hüvelyi	36.00	2850.00	53.00	165.00	68.00	Igen	nem	55.00
59	59	29	egyetem. f...	falu	hüvelyi	37.00	2650.00	51.00	165.00	65.00	Igen	nem	65.00
60	60	30	egyetem. f...	város császárme...	37.00	2400.00	50.00	166.00	68.00	Igen	nem	nem	50.00
61	61	30	egyetem. f...	falu császárme...	38.00	2300.00	49.00	166.00	60.00	Igen	nem	nem	105.00
62	62	31	felsőfokú s...	falu császárme...	39.00	2550.00	48.00	165.00	60.00	Igen	nem	nem	102.00
63	63	32	felsőfokú s...	falu császárme...	40.00	2910.00	100.00	166.00	110.00	Igen	nem	nem	111.00
64	64	33	felsőfokú s...	falu császárme...	40.00	3000.00	110.00	168.00	112.00	Igen	nem	nem	120.00
65	65	29	érettség	falu császárme...	40.00	3000.00	120.00	168.00	122.00	Igen	nem	nem	121.00
66	66	25	érettség	falu császárme...	39.00	4000.00	119.00	169.00	123.00	Igen	nem	nem	119.00
67	67	26	érettség	falu császárme...	39.00	4100.00	118.00	167.00	125.00	Igen	nem	nem	100.00
68	68	25	érettség	falu császárme...	40.00	3900.00	99.00	168.00	102.00	Igen	nem	nem	103.00
69	69	27	OKJ	falu császárme...	39.00	3800.00	98.00	168.00	105.00	Igen	nem	nem	70.00
70	70	31	OKJ	falu császárme...	39.00	2770.00	65.00	165.00	78.00	Igen	nem	nem	

5/17. ábra: Az új „összevont” adatbázis

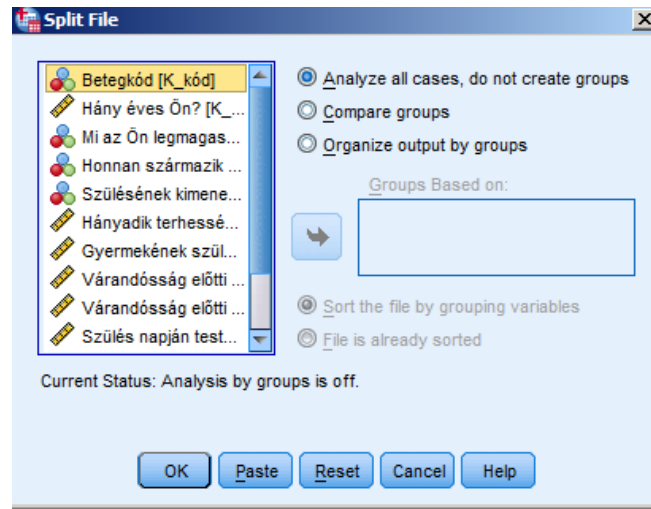
Ezt követően érdemes az új adatbázis elmenteni, melyet most mi is **születési adatbázis70.sav** néven elmentünk.

AGGREGATE opció az adatok aggregálására szolgál, vagyis az általunk választott funkció (összeadás, átlag, stb.) alapján összevonást végez.

SPLIT FILE opció segítségével lehetőségünk van az adatbázist egy általunk meghatározott és fontosnak vélt ismérv alapján csoportokra bontani, és így a további statisztikai elemzéseket ezen csoportokon végezzük. Három lehetőség közül tudunk választani:

- Minden lehetséges esetet vizsgálunk, és nem képezünk csoportokat
- Összehasonlítjuk a csoportokat
- A számítások eredményeit, outputjait csoportonként jelenítjük meg.

Amennyiben ezt a lehetőséget használjuk, az adatszét jobb alsó sarkában megjelenik a felirat „*SPLIT FILE ON*”, mindaddig, míg ki nem kapcsoljuk, vissza nem állítjuk az alapbeállítást.

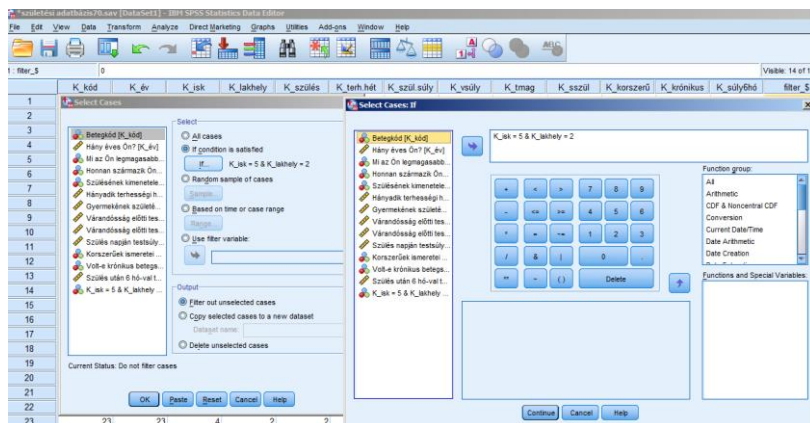


5/18. ábra: Az adatok csoportokra bontásának beállításai

SELECT CASES FILE segítségével megadott szűrőfeltételek mentén lehet az adatbázisból eseteket kiválasztani, kizárni, és ezt követően a feltételeknek megfelelő adatokkal dolgozunk tovább. Alkalmazása során a nem megfelelő egyedeket, rekordokat a sorszámuknál egy fekete nyíllal a program áthúzza, és a „*Filter ON*” felirat jelenik meg. A beállításoknál négy opciók lehet:

- Minden eset vegyen részt az elemzésben
- Csak a feltétel(ek)-nek megfelelőek
- Véletlen minta az esetekből
- Kijelölhetünk eseteinknek bizonyos (első és utolsó eset megjelölése) feltételeket.

Gyakorló feladatként szűrjük le azon kismamákat, akik falun élnek és legmagasabb iskolai végzettségük 8 általános iskola. Ez azt jelenti, hogy csak azon kismamákkal kívánunk foglalkozni, akik a feltételeknek megfelelnek.



5/19. ábra: A szűrőfeltételek megadása

Látható, hogy a feltétel opciót választva eljutunk, ahhoz a panelhez, ahol lehetőségünk van az iskola végzettségénél az általános iskolai végzettséget (K_isk) választani (5-ös ismérvték alatt szerepel), valamint a lakhelynél (K_lakhely) a falut (2-es ismérvték) megadni.

Ezt követően az adatnézet ablakban látható, hogy a feltételeknek nem megfelelő esetek sorszámát a program fekete vonallal keresztülhúzza.

	K_kód	K_év	K_isk	K_lakhely	K_szülés	K_terh.hét	K_szül.súly	K_vsúly	K_tmag	K_sszül	K_korszerő	K_krónikus	K_súlyghó	filter_S	
1	1	40	1	1	1	2	38,00	2500,00	55,00	155,00	68,00	1	2	60,00	0
2	2	25	1	1	1	1	35,00	2600,00	58,00	160,00	62,00	1	2	60,00	0
3	3	23	2	1	1	1	36,00	3000,00	59,00	160,00	70,00	1	2	59,00	0
4	4	18	2	1	1	1	38,00	3200,00	56,00	164,00	70,00	1	2	56,00	0
5	5	22	2	1	1	1	37,00	3300,00	54,00	165,00	70,00	1	2	58,00	0
6	6	25	2	1	1	1	35,00	3400,00	57,00	165,00	71,00	1	2	59,00	0
7	7	26	2	1	1	1	34,00	3500,00	58,00	166,00	69,00	1	2	60,00	0
8	8	33	2	1	2	2	36,00	4000,00	59,00	166,00	70,00	1	2	59,00	0
9	9	35	2	1	2	2	37,00	3800,00	53,00	163,00	75,00	2	2	55,00	0
10	10	36	2	1	2	2	38,00	3500,00	60,00	163,00	78,00	2	2	62,00	0
11	11	24	3	1	2	2	39,00	3400,00	60,00	163,00	80,00	2	2	62,00	0
12	12	28	3	1	2	2	40,00	3300,00	60,00	162,00	85,00	2	2	80,00	0
13	13	29	4	1	2	2	41,00	2900,00	64,00	162,00	70,00	2	2	70,00	0
14	14	27	4	1	2	2	38,00	2200,00	65,00	162,00	75,00	1	2	75,00	0
15	15	19	2	2	2	2	35,00	2150,00	80,00	165,00	89,00	1	2	89,00	0
16	16	44	5	2	1	1	37,00	3000,00	82,00	164,00	95,00	1	1	95,00	1
17	17	42	5	2	1	1	36,00	3400,00	89,00	168,00	92,00	1	1	90,00	1
18	18	41	3	2	1	1	37,00	3700,00	97,00	169,00	100,00	1	1	97,00	0
19	19	39	3	2	1	1	38,00	3800,00	94,00	169,00	100,00	2	1	95,00	0
20	20	28	3	2	1	1	39,00	3900,00	95,00	169,00	110,00	2	2	100,00	0
21	21	25	3	2	12	12	38,00	4000,00	96,00	169,00	100,00	2	2	98,00	0
22	22	26	4	2	2	2	39,00	4100,00	93,00	170,00	98,00	2	2	95,00	0
23	23	23	4	2	2	2	36,00	4000,00	92,00	170,00	96,00	2	2	96,00	0
24	24	21	4	1	2	2	40,00	4200,00	91,00	171,00	95,00	2	1	92,00	0
25	25	20	5	2	2	2	41,00	3980,00	89,00	171,00	92,00	2	1	92,00	1

5/20. ábra: A szűrés eredménye

A sorszámokon kívül megjelenik egy új oszlop ahol 0-val a szűrési feltételeknek nem megfelelő eseteket jelöli, míg a megfelelő egyedek 1-es számot kaptak. Jelen esetben a 16, 17 és a 25 kódszámú kismama felel meg a szűrési feltételeinknek. Amennyiben ismét teljes adatbázissal kívánunk dolgozni, abban az esetben állítsuk vissza az eredeti beállítást, vagyis a minden esettel történő analíziseket jelöljük.

WEIGHT CASES segítségével az esetek súlyozására nyílik lehetőség, melyet gyakran alkalmazunk, amikor a mintanagyság „finomhangolása” történik. Lehetőségünk van az alulreprezentált csoportokat nagyobb súllyal, illetve a túlreprezentált eseteket kisebb súllyal megjeleníteni.

TRANSFORM menü az adatok transzformálására, átalakítására szolgál. Az adatokat lehet átalakítani, új változókat létrehozni, akár meglévő változók segítségével. Lehetőség van a változók újrakódolására, az esetek rangszámait is kiszámíthatjuk. Az itt található opciók többsége az adatok tisztítására, „manipulálására” szolgál, ezeknek a részletes bemutatására a későbbiekben sor kerül.

COMPUTE VARIABLE opcióval új változó számítását jelenti. Ezt leginkább az adatokkal való műveletekre származtatott adatok és mutatószámok előállítására használjuk.

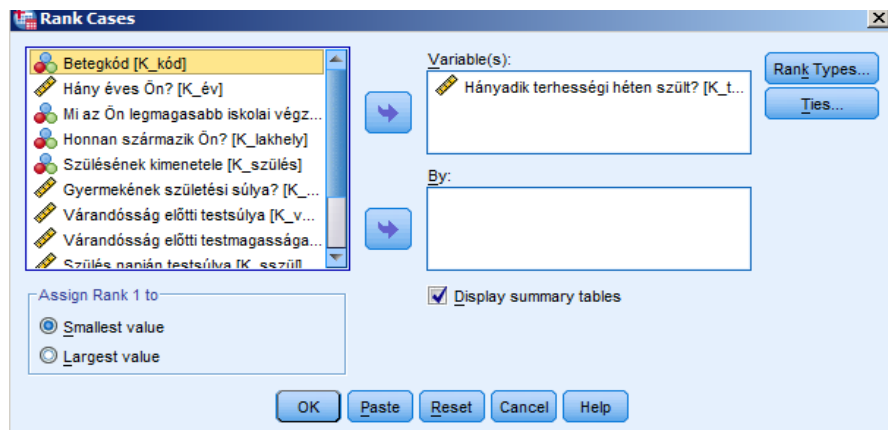
RECODE, vagyis átkódolás opciónak három különböző fajtája jelenik meg, ebben a menüben:

- **INTO SAME VARIABLES**, azt jelenti, hogy nem jön létre új változó, hanem meglévő változóban történik az átkódolás. Gyakran használjuk, amikor skálák irányultságát kívánjuk megfordítani, itt nincsen szükségünk az eredeti változó megtartására.
- **INTO DIFFERENT VARIABLES** segítségével egy új változót hozunk létre, a régi változó megtartása mellett. Ezt leggyakrabban a csoportok képzésénél, kategorizálásnál használjuk. Javaslatként ez utóbbi alkalmazása praktikus, hiszen a régi változó törlését a későbbiekben is bármikor eszközölhetjük. Ez jelenti, hogy a kódolást arra használjuk, hogy mennyiségi változóértékeket (pl. valamilyen nagyságrendi szemléletben) szövegesen definiált kategóriákba csoportosítsuk. Az keletkező nominális ismérveket tudjuk elnevezni, de lehetőség van számszerű értékekhez nominális címkék hozzárendelésére is (Jánosa 2008).
- **AUTOMATIC RECODE**, segítségével a változók automatikus átkódolása történik. A változó listából az újrakódolandó változót kiválasztva, egy új változó nevet adhatunk meg (**NEW NAME**) és az OK lenyomását követően az átkódolás megtörténik.

RANK CASES opció segítségével az eseteinket tudjuk rangsorba állítani és egy új változóban a rangorszámokat tudjuk feltüntetni. Ez egyfajta skála transzformációnak is felfogható, hiszen egy skála típusú változóból ordinális típusú változó keletkezik, ami információvesztéssel is járhat, illetve így már bizonyos statisztikai eljárások nem kivitelezhetőek velük, vagy módosulnak. A cél változónak a **VARIABLES** ablakba mozgatjuk. Amennyiben csoportokon belüli rangsorokat szeretnénk, akkor az alsó dobozt is használjuk (**BY**). Ebben az ablakban tudjuk a sorrendiséget jelölni (legkisebb vagy a legnagyobb érték

kapja az egyes rangszámot). Ezen felüli speciális rangszám beállításokra is lehetőségünk van (*RANK TYPES*), illetve az azonos értékű rangszámok módosítását is kérhetjük (*TIES*). A létrejövő új változó az eredeti nevet kapja egy R előtag illesztésével.

Gyakorló feladatként, rangsoroljuk a szülési dátumokat a terhességi hetek alapján. Az egyes rangszámmal jelöljük a legkorábbi szülést.



5/21. ábra: A rangsorolás beállításai

DI VISUAL BINNING és *OPTIMAL BINNING* opciók segítségével a változók kategorizálására van lehetőségünk. Gyakorlatilag ez jelenti, hogy a folytonos mennyiségi ismérvek diszkrét egyedi kategóriába sorolódnak. Mindkét eljárás új változó létrehozását eredményezi.

CREATE A TIME SERIES segítségével idősor jellegű változókból, más típusú idősor változókat hozhatunk létre. A *FUNCTION* segítségével az idősorokra jellemző módszertanok beállítására van lehetőségünk (szezonális, mozgóátlagolás, stb.). Az idősorok elemzése nem képezi jelen könyvünk tárgyát, az érdeklődő Pintér – Rappai (2001) könyvében részletesen olvashat az eljárásról.

REPLACE MISSING VALUES opciót használjuk, amikor a hiányzó értékeket (*MISSING VALUES*) pótolni szükséges az elemzéshez. A különböző módszerek (*METHOD*) állnak rendelkezésünkre átlag, szomszédos értékek átlaga vagy mediánja, lineáris trendértékek stb. a hiányzó értékek pótlására.

RANDOM NUMBER GENERATORS, ami a véletlen számokat generálja.

ANALYZE menü az SPSS program egyik legfontosabb és legösszetettebb menüje, ahol a statisztikai módszertanokat összegyűjtve találjuk. Ez azt jelenti, hogy innen indítva tudjuk a tényleges adatelemzéseket elvégezni. A menüpont részletes bemutatására a leggyakrabban használt módszertanok ismertetésére a könyv későbbi fejezeteiben részletesen kitérünk.

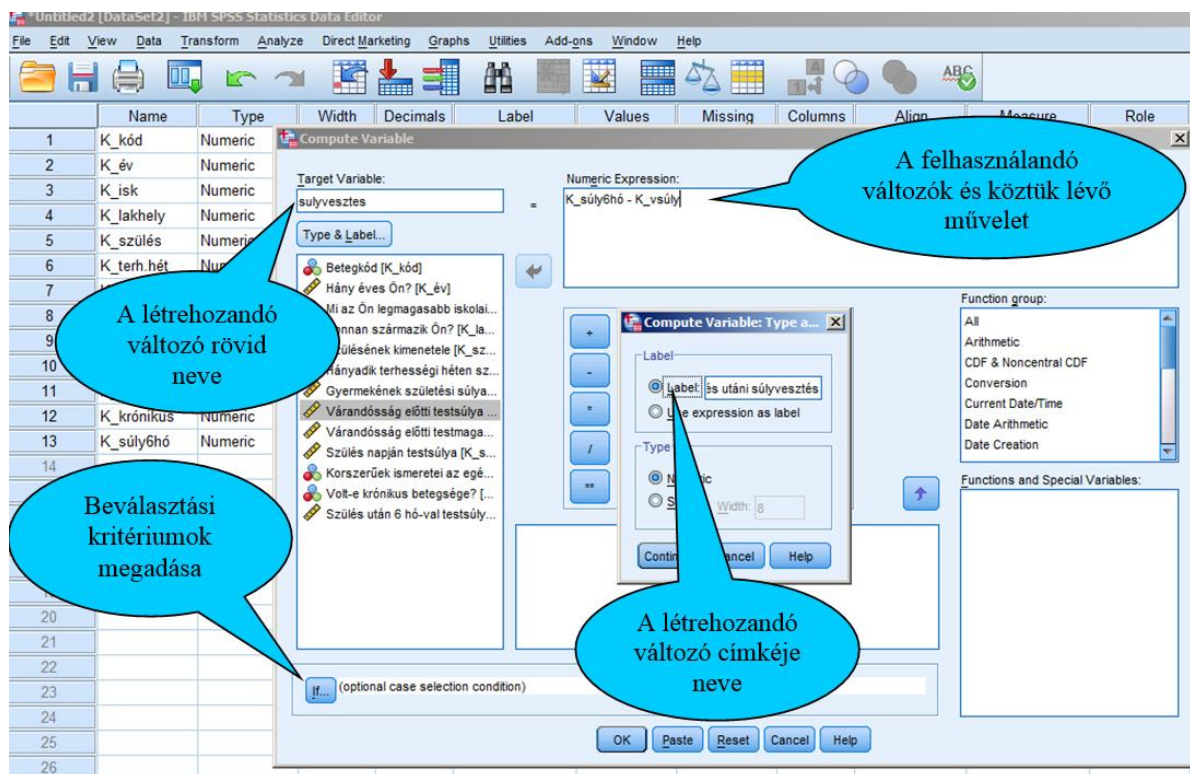
DIRECT MARKETING menü marketingkutatások során alkalmazott eljárásokat találjuk. Ez az opció a régebbi verziókban nem volt megtalálható. Marketingkampányok, piacszegmentáláshoz tartozó technikákat foglal magába, melyről az érdeklődő bővebben Jánosa (2011) könyvében olvashat.

GRAPS menü a statisztikai adatprezentáció során alkalmazott különböző ábrákat, grafikonokat, diagramokat tartalmazza. A statisztikai ábráknak egyértelmű célja, hogy a kutatások eredményeit meggyőzően bemutassa. A program bőséges lehetőséget nyújt a felhasználónak. A bőséges repertoárból a megfelelő ábra kiválasztását az alapadatokon túl az elérni kívánt cél határozza meg. Az alapábrázolások mindig egyfajta mértani elemhez (pont, vonal, oszlop, téglalap, kör, stb.) vagy azok kombinációjához vezethető vissza. A programban két diagram készítési mód van (*CHART BUILDER* és a *LEGACY DIALOGS*). Fontos, hogy a grafikonunknak egyértelmű címe legyen, feltüntessük a vizsgált időszakot és mértékegységeket. A grafikus ábrázolás részletes bemutatása a statisztika adatprezentáció eszközei részben kerül elő.

6. ADATTISZTÍTÁS, EGYSZERŰBB MŰVELETEK A PRIMER ADATOKKAL, ADATMANIPULÁCIÓ (Ács Pongrác)

A statisztikai elemzések megkezdése előtt és annak szigorú feltételeként fontos, hogy olyan adatbázissal rendelkezünk, amely minden szempontból megfeleljen a vele szemben támasztott követelményeknek. A kutatási célok, illetve a hipotézisek azok, amelyek meghatározzák és változtathatják adott esetben a változóinkat. Ezt jelentheti azt, hogy a hipotéziseink bizonyos változóink összevonását egy új változóba (pl. BMI index), vagy bizonyos adattranszformációt követelhetnek meg. Szintén priori elvárás sok módszer előfeltételeként, hogy a változóink normalitásvizsgálatát elvégezzük, esetleg az „outlier” (kiugró) értékeket is vizsgáljuk. Ebben a fejezetben szeretnénk gyakorlati szemléltetéssel bemutatni ezeket a műveleteket.

Gyakorló feladatként, adatbázisunk segítségével vizsgáljuk meg azon kismamák elmúlt 6 havi súlyvesztésének értékeit, akik saját bevallásuk szerint nem voltak krónikus betegek. Érdekes lehet, hogy 6 hónap alatt a szülés előtti testsúlyra a kismamák vissza tudnak-e térni?

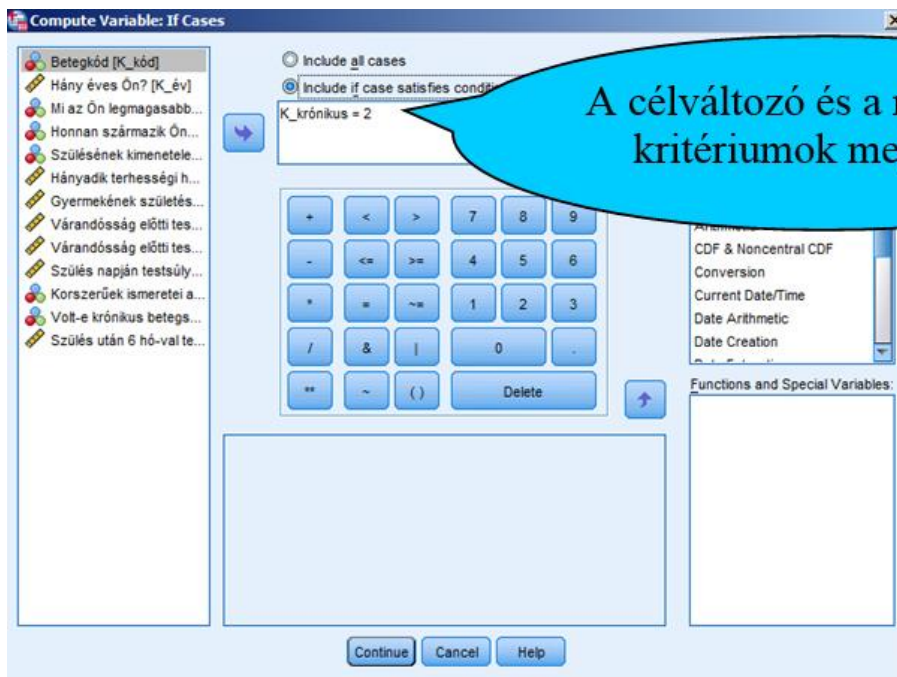


6/1. ábra: Az új változó létrehozása

Gyakorlatilag ezt az adatmanipulációt a *TRANSFORM* menü *COMPUTE VARIABLE* opciójából tudjuk indítani. Az első lépésben a létrehozandó változónak kell rövid nevet adni (sulyvesztes), ezt követően a *NUMERIC EXPRESSION* ablakba a mozgatjuk azon

változóinkat melyekkel dolgozni fogunk, és természetesen a köztük lévő matematikai műveletet vagy logikai művelet is jelezzük. Nyilvánvalóan számos előre programozott képletet is elő lehet hívni (*FUNCTIONS AND SPECIAL VARIABLES*), de számunkra erre most nincsen szükség. A *TYPE&LABEL* dobozra kattintva megadjuk az új változók típusát és címkéjét.

Ezt követően a beválasztási kritériumot is adjuk meg, vagyis a kismamák közül csak azokat vizsgáljuk, akik krónikus betegséggel nem találkoztak a vizsgált időszakban.



6/2. ábra: Beválasztási kritériumok megadása

Látható, hogy azon kismamák adataira szűkítjük az adatbázisunkat, akik a „Volt-e krónikus betegsége” kérdésre nemmel (2-es jelölést kapott az ismérvváltozó) válaszolt. Ezt követően *CONTINUE* és *OK* lenyomását követően létrejön az új változónk.

	h	K_szül_súly	K_vsúly	K_tmag	K_sszül	K_korszerű	K_krónikus	K_súlyghó	súlyvesztes
1	00	2500,00	55,00	155,00	68,00	1	2	60,00	5,00
2	00	2600,00	58,00	160,00	62,00	1	2	60,00	2,00
3	00	3000,00	59,00	160,00	70,00	1	2	59,00	,00
4	00	3200,00	56,00	164,00	70,00	1	2	56,00	,00
5	00	3300,00	54,00	165,00	70,00	1	2	58,00	4,00
6	00	3400,00	57,00	165,00	71,00	1	2	59,00	2,00
7	00	3500,00	58,00	166,00	69,00	1	2	60,00	2,00
8	00	4000,00	59,00	166,00	70,00	1	2	59,00	,00
9	00	3800,00	53,00	163,00	75,00	2	2	55,00	2,00
10	00	3500,00	60,00	163,00	78,00	2	2	62,00	2,00
11	00	3400,00	60,00	163,00	80,00	2	2	62,00	2,00
12	00	3300,00	60,00	162,00	85,00	2	2	80,00	20,00
13	00	2900,00	64,00	162,00	70,00	2	2	70,00	6,00
14	00	2200,00	65,00	162,00	75,00	1	2	75,00	10,00
15	00	2150,00	80,00	165,00	89,00	1	2	89,00	9,00
16	00	3000,00	82,00	164,00	95,00	1	1	95,00	
17	00	3400,00	89,00	168,00	92,00	1	1	90,00	
18	00	3700,00	97,00	169,00	100,00	1	1	97,00	
19	00	3800,00	94,00	169,00	100,00	2	1	95,00	
20	00	3900,00	95,00	169,00	110,00	2	2	100,00	5,00
21	00	4000,00	96,00	169,00	100,00	2	2	98,00	2,00
22	00	4100,00	93,00	170,00	98,00	2	2	95,00	2,00
23	00	4000,00	92,00	170,00	96,00	2	2	96,00	4,00
24	00	4200,00	91,00	171,00	95,00	2	1	92,00	
25	00	3980,00	89,00	171,00	92,00	2	1	92,00	

6/3. ábra: Az új számított változó adatai

Az adatnézetben jól látszik, hogy az első kismama a szülési előtti súlyához képest még öt kilóval tér el.

Fontos megemlíteni, hogy a matematikai műveleteken kívül logikai kritériumokat is meg tudunk adni, ebben az opcióban. Leggyakrabban ez akkor történik, amikor egy többtételes skálából nem minden kérdéstípusra van szükség.

6/1. táblázat: A leggyakrabban használt logikai műveletek táblázata

Jel	Jelentése
&	„És”
	„Vagy”
~	„Nem”
<	„Kisebb, mint...”
>	„Nagyobb, mint...”
<=	„Kisebb vagy egyenlő, mint...”
>=	„Nagyobb vagy egyenlő, mint...”
=	„Egyenlő”
~=	„Egyenlőtlenség”

Forrás: saját szerkesztés

A következő gyakorlati példával kettős célunk lesz, egyrészt az előbb bemutatott módon egy új változó (BMI INDEX) létrehozását kívánjuk gyakorolni, másrészt a létrejövő folytonos változóból egy kategóriákat tartalmazó diszkrét változó előállítását szemléltetjük. A BMI INDEX számítása egyszerű, a testsúlyokat elosztjuk a méterben megadott testmagasságok négyzetével (kg/m²).

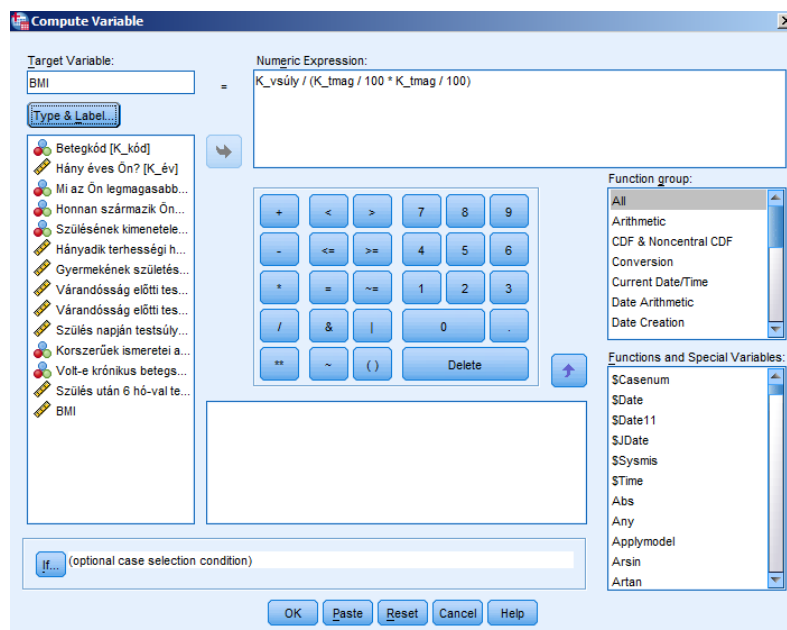
A BMI kategóriákat a következőképpen definiáljuk:

6/2. táblázat: BMI kategóriák

Érték	Jelentése
<18,49	„sovány”
18,5- 24,9	„normál testalkatú”
25-29,9	„túlsúlyos”
>30	„erősen túlsúlyos”

Forrás: saját szerkesztés

Első lépésben ismét a *TRANSFORM* főmenü első opciójára klikkelünk (*COMPUTE VARIABLE*).



6/4. ábra: A BMI index számításának beállításai

Először a *TARGET VARIABLE* ablakban az új változó elnevezését (BMI) tesszük meg, majd a *NUMERIC EXPRESSION* helyén az új változónk számítási képletét ($K_vsúly / (K_tmag / 100 * K_tmag / 100)$) adjuk meg. Ezt megtehetjük közvetlenül, vagy közvetlenül a számológépre

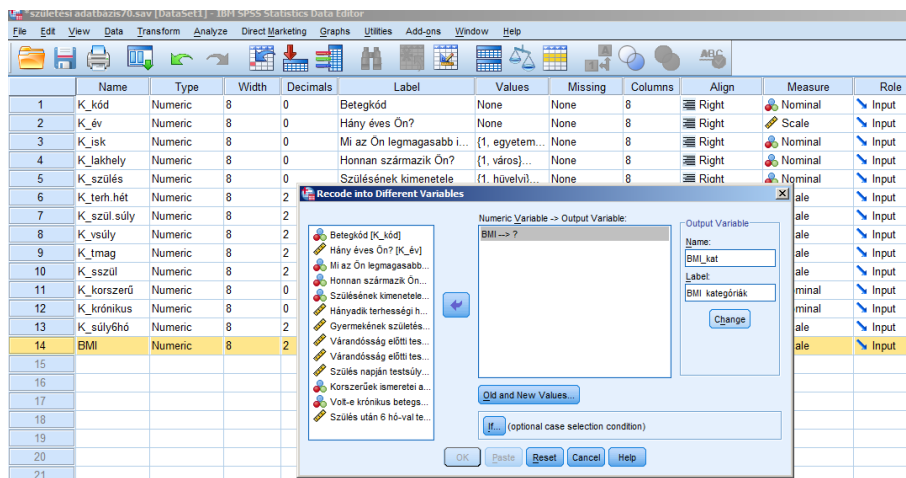
emlékeztető panel segítségével. A felhasználásra kerülő változókat az ismérveket tartalmazó ablakból a közepén található nyíl segítségével tudjuk mozgatni. A numerikus gombokat tartalmazó ablak mellett a *FUNCTION* ablakokat látjuk, melyek az előre programozott függvényeket, képleteket, utasításokat tartalmazzák. Jelen számításnál nincsen szükségünk egyéb szűrőfeltétel (IF) megadására. Miután elvégeztük a fenti beállításokat nyomjuk meg az *OK* gombot.

	K_szülés	K_terh.hét	K_szul.súly	K_vsúly	K_tmag	K_sszül	K_korszerü	K_krónikus	K_súly6hó	BMI
1	2	38,00	2500,00	55,00	155,00	68,00	1	2	60,00	22,89
2	1	35,00	2600,00	58,00	160,00	62,00	1	2	60,00	22,66
3	1	36,00	3000,00	59,00	160,00	70,00	1	2	59,00	23,05
4	1	38,00	3200,00	56,00	164,00	70,00	1	2	56,00	20,82
5	1	37,00	3300,00	54,00	165,00	70,00	1	2	58,00	19,83
6	1	35,00	3400,00	57,00	165,00	71,00	1	2	59,00	20,94
7	1	34,00	3500,00	58,00	166,00	69,00	1	2	60,00	21,05
8	2	36,00	4000,00	59,00	166,00	70,00	1	2	59,00	21,41
9	2	37,00	3800,00	53,00	163,00	75,00	2	2	55,00	19,95
10	2	38,00	3500,00	60,00	163,00	78,00	2	2	62,00	22,58
11	2	39,00	3400,00	60,00	163,00	80,00	2	2	62,00	22,58
12	2	40,00	3300,00	60,00	162,00	85,00	2	2	80,00	22,86
13	2	41,00	2900,00	64,00	162,00	70,00	2	2	70,00	24,39
14	2	38,00	2900,00	65,00	162,00	75,00	1	2	75,00	24,77

6/5. ábra: A létrejött BMI változó az ADAT nézet ablakban (részlet)

Látható, hogy a BMI névvel létrejött egy új folytonos változó. Ezt követően az átkódolás történik meg, mégpedig egy új változót szeretnénk létrehozni egy meglévő változóból. Ez azért előnyös, mivel az eredeti folytonos változónk is megmarad, a diszkrét kategóriákat tartalmazó változó mellett.

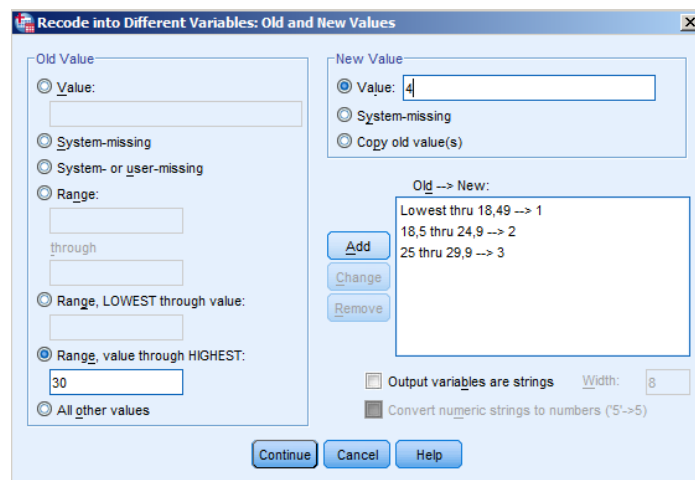
A *TRANSFORM* főmenü *RECODE INTO DIFFERENT VARIABLES*-re kattintva a következő beállításokat tartalmazó ablak jelenik meg.



6/6. ábra: Az átkódolás beállításai

A megjelenő ablak bal oldalán található az összes változó. A változók közül a BMI változót mozgattuk át a középen található nyíl segítségével az *OUTPUT VARIABLE* ablakba. Ezt követően rövid változónevet (BMI_kat) és címkét (BMI kategóriák) adunk az új változónknak. Fontos, hogy ezt követően kattintsuk a *CHANGE* gombra, hiszen csak így fog a kódolásunk új neve életbe lépni. Ennek visszaigazolásként az *INPUT VARIABLE*→*OUTPUT VARIABLE* ablakban megjelenik az eredeti és az új változó neve egyaránt.

Ezt követően a régi változók értékeinek megadása és az új kalibrálása történik. A felugró ablakot két részre lehet osztani, hiszen a baloldalon a kiinduló, a jobboldalon az új változó értékei láthatóak.



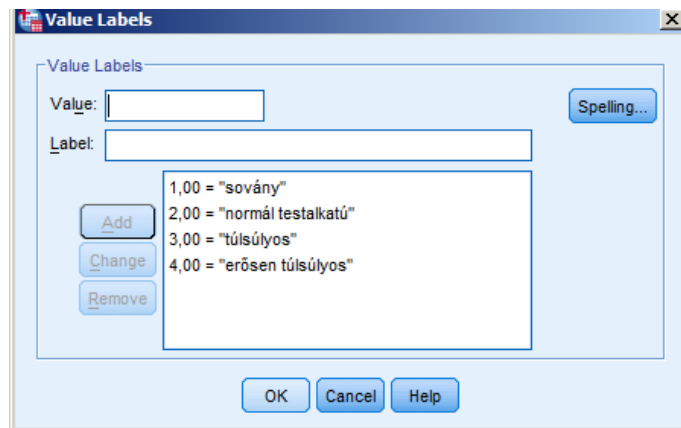
6/7. ábra: Az új diszkrét változó kategóriáinak meghatározása

Amennyiben a *VALUE* opciót jelöljük be, akkor egyesével megadhatjuk a régi értékeket. A *SYSTEM-MISSING* vagy a *SYSTEM -OR USER- MISSING* opciók segítségével azokat a tételeket tudjuk kizárni, melyek a feltételeinknek nem felelnek meg. A *RANGE* opció beállításainak segítségével tudjuk a különböző az osztályközök feltételeit megszabni. Az első alternatíva, amikor a *RANGE... THROUGH...* választásával az osztályközök alsó és felső határát tudjuk az értéktartomány beírásával megadni (pl.: 18.5 és 24,9 közöttiekből 2-es csoportot). A *RANGE LOWEST THROUGH* segítségével az aluról nyitott osztályközt tudjuk megadni, míg a *RANGE THROUGH HIGHEST* opció segítségével a felülről nyitott osztályközt definiálhatjuk (pl.: 30-as érték felett, 4-es kódot kapnak). A jobb oldalon található új értékek megadására, akkor van lehetőségünk, ha a régi értékeket az előzőekben ismertetett módon megadtuk. A régi értékhez tartozó új értéket a *VALUE* szövegdobozba rögzítjük, majd az *ADD* gombra nyomva véglegesítjük. Ezt követően a változó kalibrálása megjelenik az *OLD*→*NEW* ablakban. Ha a kategóriákat módosítani szeretnénk, akkor a *CHANGE*, ha

törölni, akkor a *REMOVE* gombot válasszuk. Amennyiben az új kategóriák szöveges formátumúak, akkor az *OUTPUT VARIABLES ARE STRINGS* jelölőnégyzetet kipipáljuk.

A beállításokat elvégezve nyomjuk a *CONTINUE* gombot, majd az *OK* gomb lenyomását követően a négy kategória létrejön.

Miután a négy kategória, csak egy numerikus értéket tartalmaz, ezért a kategóriák elnevezését a változó nézetben a már leírt módon készítjük.

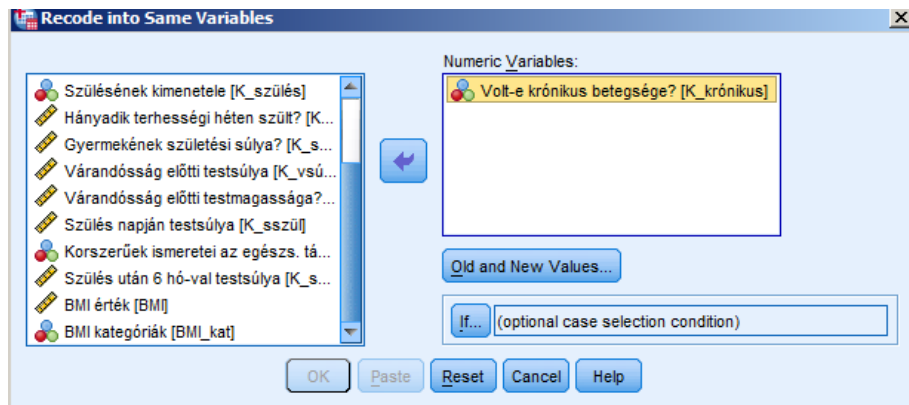


6/8. ábra: A létrejött új csoportok elnevezése, címkézése

Az adatok korrekciójánál gyakran fordul elő, hogy nem kell egy teljesen új változót létrehozni, elegendő a meglévő változót javítani, változtatni (pl. abban az esetben is ezt az opciót javasoljuk használni, amikor a skála irányultságát kívánjuk megfordítani). Ebben az esetben a *TRANSFORM* főmenü *RECODE INTO SAME VARIABLES*-re kattintva tudjuk a szükséges beállításokat elvégezni.

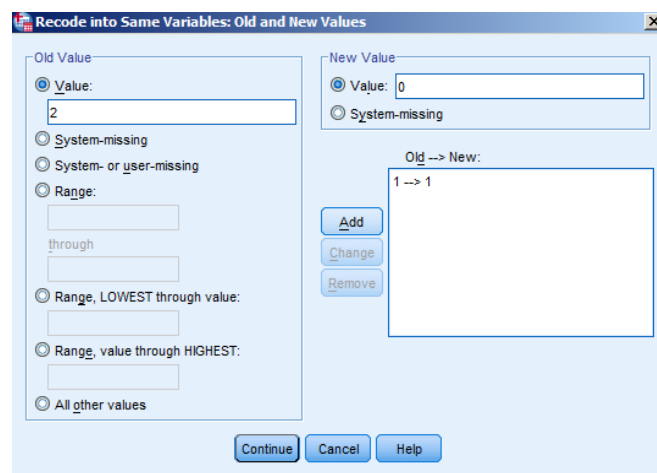
Gyakorló feladatként, az adatbázisunk segítségével kódoljuk át a krónikus betegség meglétére irányuló kérdésben lévő csoportokat. Az adatbázist nézve látható, hogy eredetileg az kismamák, akiknek volt krónikus betegsége a vizsgált időszakban 1-es kódot, akinek nem 2-es kódot kapott. Feladatunk, hogy azokat a kismamákat, akiknek nem volt krónikus betegsége 0-kódra változtassuk, a többi érték (1-es) megtartása mellett.

A *TRANSFORM* főmenü *RECODE INTO SAME VARIABLES*-re kattintva a következő beállításokat tartalmazó ablak jelenik meg.



6/9. ábra: Az átkódolás beállításai

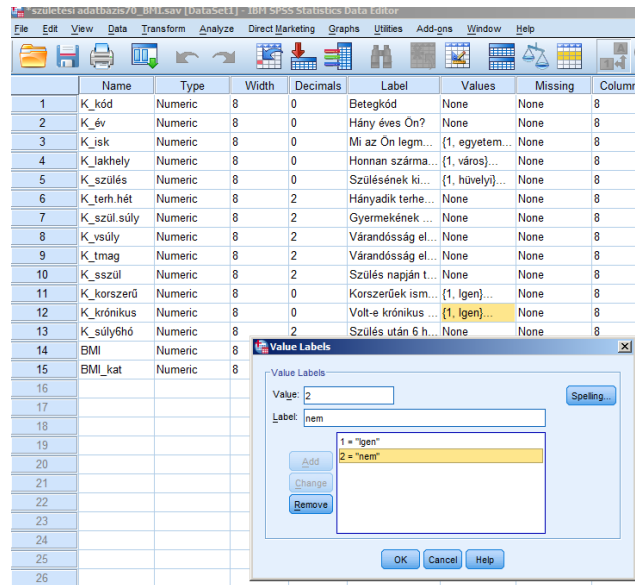
A bal oldali ablakban található a választható változók listája, amelyből a krónikus betegségre vonatkozó ismérvet, kérdést választottuk ki („Volt-e krónikus betegsége?”), melyet a középen található nyíl segítségével a *NUMERIC VARIABLES* ablakba mozgattunk. Ezt követően az ismérvváltozók kalibrálása történik meg.



6/10. ábra: A csoportkódok változtatásának beállításai

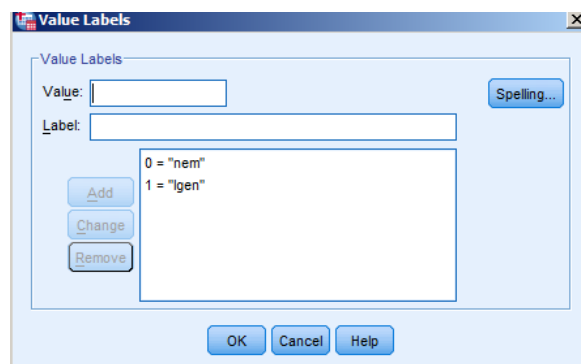
Jelen esetben a *VALUE* opciót jelöljük be, és egyesével megadhatjuk a régi értékeket, majd ezt követően az új érték kategóriákat. Első lépésben az *OLD VALUE* dobozban is és a *NEW VALUE* dobozban is az 1-es értéket gépeljük, melyet a középen lévő *ADD* gomb megnyomását követően a középső dobozban láthatunk. Ezt követően következik a régi (*OLD VALUE*) 2-es értékek kódolása 0-ás új értékévé (*NEW VALUE*). Miután itt is befejeztük a beállítást (*ADD*) a *CONTINUE* és *OK* gomb lenyomását követően az átkódolás megtörténik, melyet az *OUTPUT* táblában is jelez számunkra a program.

Ezt követően az utolsó lépésként a változó nézetben a címkék átnevezését is el kell végezni, hiszen ez nem történik automatikusan, mivel a program nem tudhatja, hogy milyen elnevezést szánunk az új címkének.



6/11. ábra: A régi változókód törlése

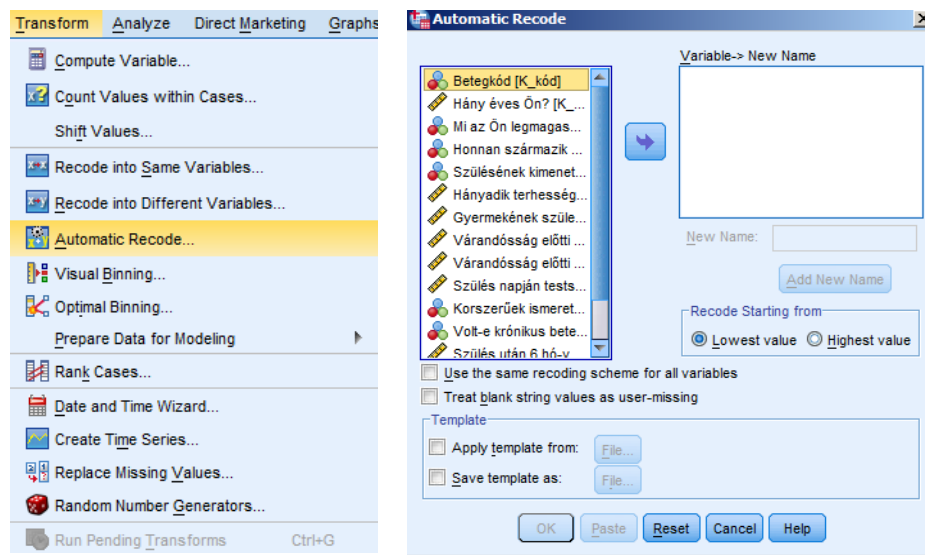
A változó nézetben (*VARIABLE VIEW*) a kérdéshez tartozó *VALUES*-re kattintva tudjuk kijelölni a 2-es kódot, melyet törölni (*REMOVE*) szeretnénk. Az *OK* gomb lenyomását követően a régi kód és hozzá tartozó címke eltűnik. Ezt követően tudjuk az új kódot és kategória nevet megadni.



6/12. ábra: Az új változókód megadása

A *VALUE* értékdobozba megadjuk a 0-s értéket, melyet a *LABEL* ablakban „nem” változó elnevezéssel látunk el. Ezt követően az *OK* gomb lenyomásával az új kategóriakód lép érvénybe.

Az adatok átkódolásának kétségtelenül legegyszerűbb, de nem minden esetben használható módszere az automatikus átkódolás, melyet a *TRANSFORM* menü *AUTOMATIC RECODE* opciójával tudunk megtenni.



6/13. ábra: Az automatikus átkódolás beállítási moduljai

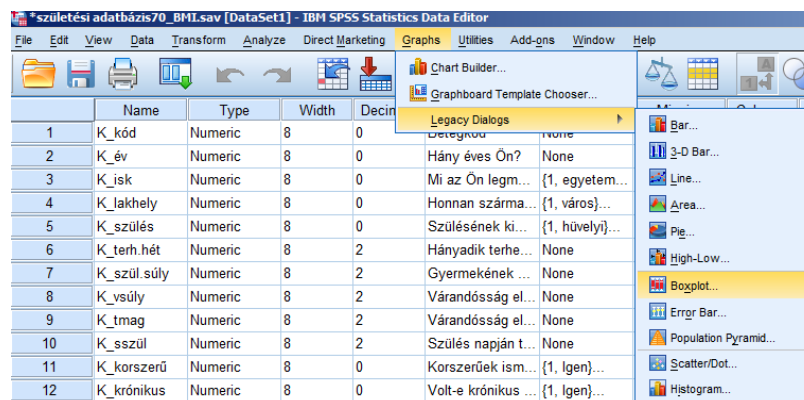
A jobb oldali ábrán először az átkódolandó változót jelöljük ki és a nyíl segítségével átmozgatjuk az üres ablakba, majd elnevezzük a *NEW NAME* dobozban. A beállítások során jelölnünk kell azt is, hogy az újrakódolást a legkisebb (*LOWEST VALUE*) vagy a legnagyobb értéktől (*HIGHEST*) kezdje.

Nagyon gyakran előkerülő probléma, hogy az adatok értékelése során vizsgálnunk kell azok ismérveit és az esetlegesen megjelenő kiugró értékről el kell döntenünk, hogy létezik-e valóságtartalma, vagy adatbeviteli hiba eredményeként keletkeztek. Ezt az adattisztítási módszert összefoglaló néven *kiugró (outlier) értékek vizsgálatának* nevezzük. A módszer eredménye nyomán a kutatónak a detektált kiugró értékekről kell eldöntenie, hogy kizárja, vagy az elemzés számára megtartja. Leginkább a folytonos változók esetében használható és érdemes a statisztikai analízisek előtt az extrém értékeket vizsgálni. A döntés meghozatalában a kutató szakértelme és a témában való jártassága, olvasottsága játszik priori szerepet. A gyakorlatban a kutatók azokat az értékeket szokták kizárni, melyek adatbeviteli, kódolási hibák végett jöttek létre vagy olyan indok nélküli esemény következménye, melyre a kutatónak konkrét, objektív magyarázata nincsen. Fontos, hogy meg kell tudni határozni azt a konkrét értéket, mely felett kiugró, extrém értéknek minősül az adat. Hibát követünk el, abban az esetben, ha a kiugró érték megtartása mentén az torzító hatással bír (pl. normalitásvizsgálat során), illetve akkor is, ha kizárás mellett döntünk, pedig az nem lenne indokolt, hiszen az

érték egy valós megfigyelést tartalmaz és az általánosíthatóságot is segíti. Az egyik leggyakrabban alkalmazott módszertan az *outlier* (kiugró) értékek vizsgálatára az ún. *Boxplot* diagram. Az ábra segítségével egyértelműen kiderülhet, hogy van-e kiugró értékünk és az mely egyednél fordul elő.

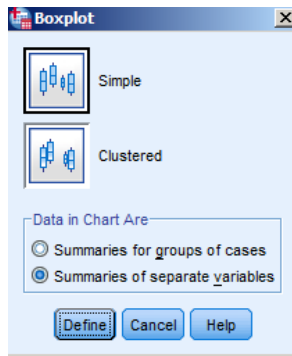
Gyakorló feladatként, vizsgáljuk meg, Boxplot ábra segítségével az adatbázisunkban található kismamák életkorát, valamint BMI értékét (születési adatbázis70_BMI.sav).

A grafikus ábra készítését megtaláljuk a *GRAPHS* menüpont *LEGACY DIALOGS/BOXPLOT* opciójában.



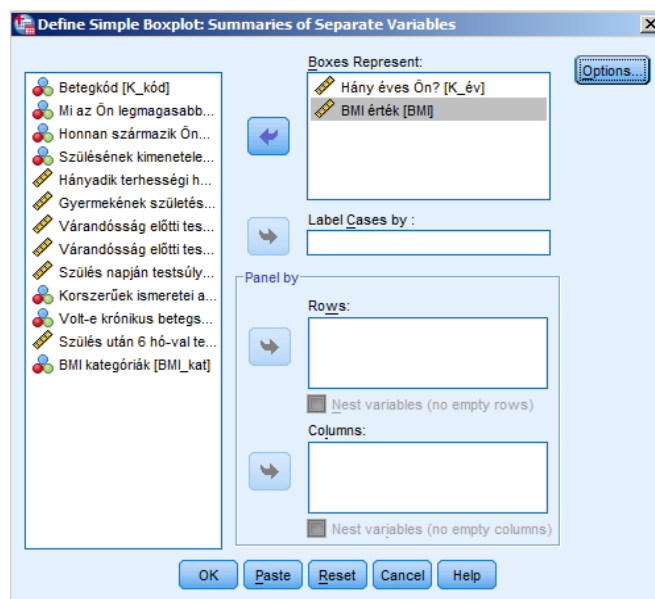
6/14. ábra: A Boxplot ábra elérési útvonala

A *BOXPLOT* menüpont alatt két alternatív megjelenítés áll rendelkezésre: egyszerű (*SIMPLE*) vagy a csoportosított (*CLUSTERED*) lehetőségek. Itt jelöljük az egyszerű megjelenési formát. Ezt követően az adatok típusánál (*DATA IN CHART ARE*) választható az eset (*SUMMARIES FOR GROUPS OF CASES*) vagy a változó szerinti (*SUMMARIES OF SEPARATE VARIABLES*) típus. Itt válasszuk a második lehetőséget. A kalibrálásunk nyomán az egyszerű típusban a változók típusa szerinti megjelenítés lesz a jó megoldás, hiszen így lehetőség van egy vagy több változó eloszlásának bemutatására. Amennyiben az esetek (*CASES*) szerinti ábrázolást választjuk, abban az esetben egy adott változót (pl.: életkor) egy másik változó kategóriái (pl.: lakhely; város, falu) szerint ábrázolhatunk. A *CLUSTERED* opció segítségével, amennyiben a változók típusa szerint ábrázolunk, egyszerre minimum két folytonos változót jelenítünk meg (pl.: kor, BMI) egy másik változó (pl. lakhely) kategóriák szerint.



6/15. ábra: Boxplot típusának kalibrálása

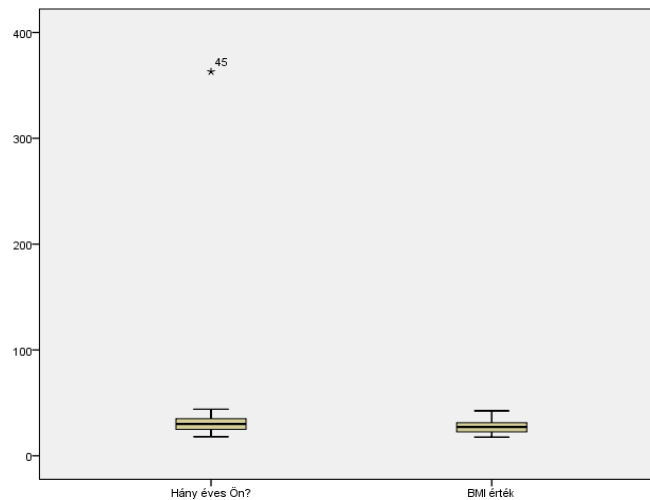
Ezt követően a vizsgálatba bevont változóink kiválasztása következhetik. A *BOXES REPRESENT* ablakban kell az életkor és a BMI érték változókat a középen található nyíl segítségével mozgatni. Ezt követően további beállításokra nincsen szükség, így az *OK* gombra kell kattintani.



6/16. ábra. A vizsgálatba bevont változók kijelölése

Az output ábrán megjelenő téglalapok (dobozok) szélei mutatják az alsó (25) és felső (75) kvartilis közötti távolságot, míg a középen megjelenő vonal a medián (50) értékét. Az interkvartilis (felső és alsó kvartilis különbsége) másfélszerese a dobozból felfelé és lefelé irányuló vonalak hossza. Ideális esetben az értékek ebben a terjedelemben helyezkednek el (normál eloszlás), melyet nyomatékosít a vonalak végén lévő vízszintes jelzés. Amennyiben az érték a doboz szélétől 1,5-3- interkvartilis terjedelemben van kiugró értéként jelöli a

program (jele:O). Ezt a három interkvartilis terjedelmet is meghaladó értéket extrém értéként kezeli és * jelöli.



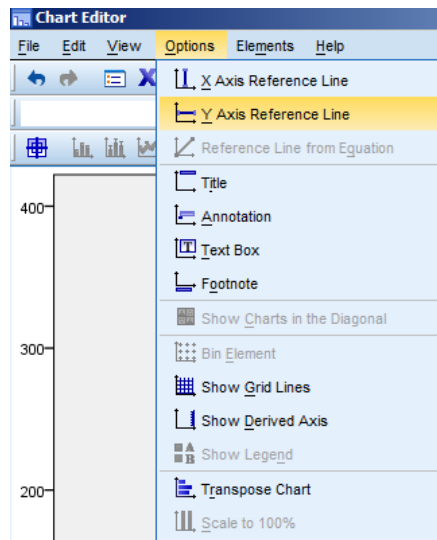
6/17. ábra. Az életkor és BMI érték változók Boxplot ábrája

Az ábrán jól látható, hogy a BMI értéknél sem kiugró, sem extrém érték nem szerepel, így vélelmezzük a változó normális eloszlását. Az életkor változó esetében a 45 sorszámú érték extrém. Az adatszétben megnézhető, hogy a 45 egyedhez mekkora életkor tartozik.

*születési adatbázis70_BMI.sav [DataSet1] - IBM SPSS				
File Edit View Data Transform Analyze Direct Mark				
45 : K_év 363				
	K_év	K_isk	K_lakhely	
25	20	8 általános	falu	c
26	22	8 általános	város	c
27	24	8 általános	város	c
28	25	8 általános	város	c
29	26	8 általános	város	c
30	31	8 általános	város	c
31	32	felsőfokú s...	falu	c
32	31	felsőfokú s...	falu	c
33	33	felsőfokú s...	falu	
34	34	felsőfokú s...	falu	
35	35	felsőfokú s...	falu	
36	36	felsőfokú s...	falu	
37	37	egyetem, f...	város	
38	38	egyetem, f...	város	
39	39	egyetem, f...	város	
40	40	egyetem, f...	város	
41	31	egyetem, f...	város	
42	32	egyetem, f...	város	
43	34	OKJ	város	c
44	35	OKJ	falu	c
45	363	OKJ	falu	c
46	37	OKJ	falu	c

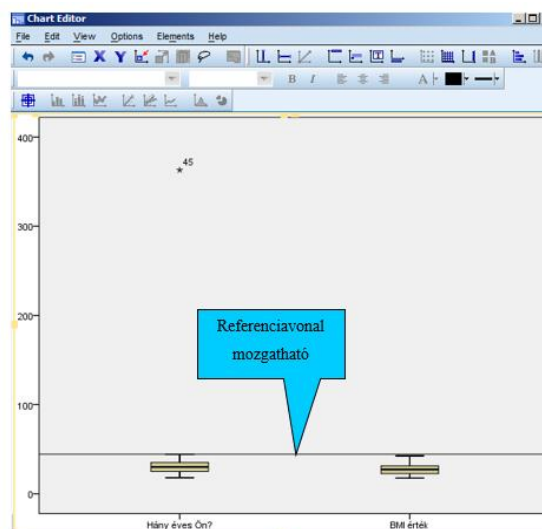
6/18. ábra: Az adatszétben a 45. kismamának az életkor értéke

Látható, hogy a 45 egyedhez tartozó életkor 363, ami az extrém értéket jelenti. Nem kell nagy szakértelem ahhoz, hogy megállapítsuk, hogy ez az életkor nem reális, vagyis adatbeviteli hiba történt. Az ilyen esetekben vagy a *missing* értéktartományt tudjuk definiálni, vagy adatszűrőt tudjuk alkalmazni (*DATA* menüpont *SELECT CASES*). Tegyük most ez utóbbit. Amennyiben az output ábrára kétszer kattintunk, megjelenik a grafikus ábra szerkesztő ablaka (*CHART EDITOR*), ahol az *ELEMENTS* opcióra klikkelve az Y tengelyre elhelyezhetünk egy referencia egyenest (*ADD REFERENCE LINE TO THE Y AXIS*).



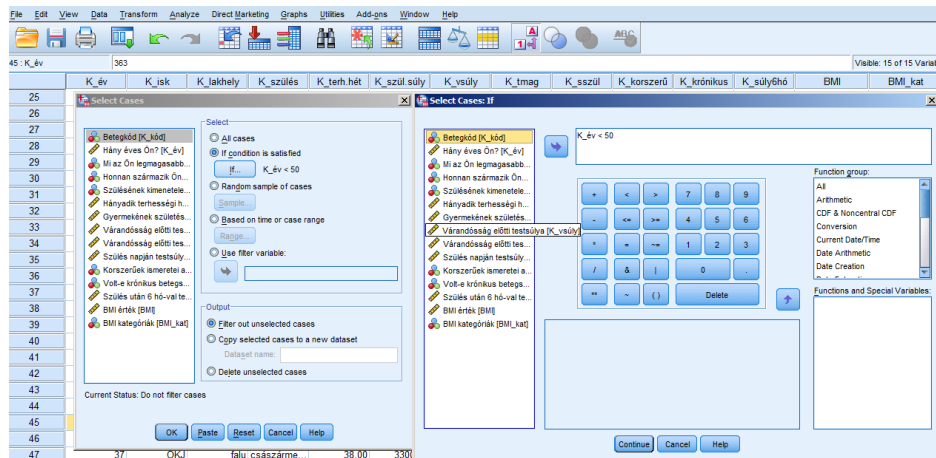
6/19. ábra: A referencia egyenes felvétele

Ezt követően megjelenik a 200-nál a referencia vonal, melyet függőlegesen mozgathatunk, ha ráklickelek a vonalra. Toljuk a referencia vonalunkat a felső vízszintes vonalig, ami 50 évnél lesz, vagyis ezt nevezhetjük küszöbértéknek.



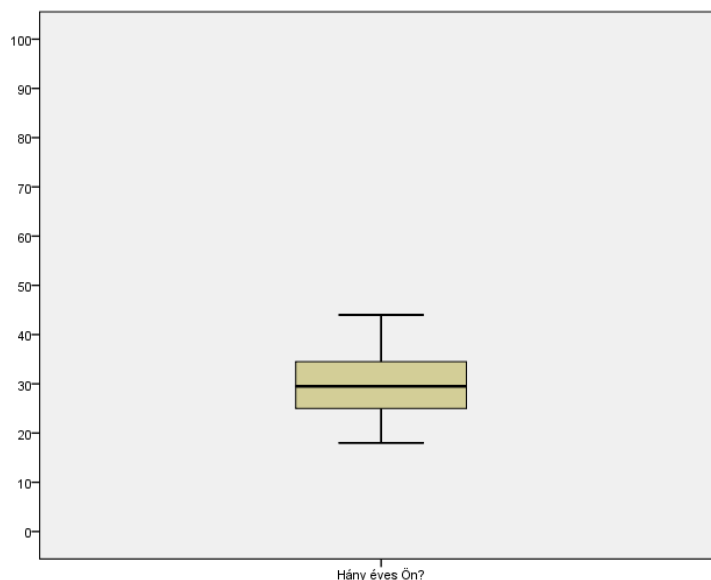
6/20. ábra: A referenciavonal felvétele a grafikus ábrára

Ezt követően a *DATA* menü *SELECT CASES* opciójában végezzük el az adatok szűrését, vagyis az IF ablakban kalibráljuk az életkor változót.



6/21. ábra: A szűrési feltételek beállítása

Az ábra szerinti beállításokat elvégezve nyomjuk meg a *CONTINUE* és *OK* gombokat. Ez azt fogja eredményezni, hogy az életkor változónál a program az 50 év felettieket kiszűri és áthúzza, melyet az adatnézetben láthatunk. Ez azt jelenti, hogy a további elemzés során ezeket az értékeket figyelmen kívül hagyja. Itt fontos megjegyezni, hogy a szűrőfeltétel mindaddig aktív marad, míg vissza nem állítjuk az összes értékre. Amennyiben ezt követően ismét kérünk egy grafikus Boxplot ábrát, akkor a már az életkor változóban sem fognak kiugró és extrém értékek jelentkezni.



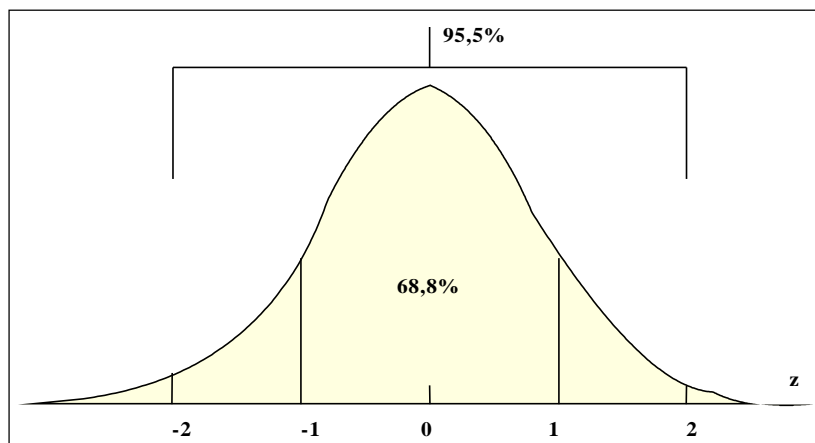
6/22. ábra: A kiugró értéktől tisztított életkor boxplot ábrája

A kiugró értékek vizsgálata során gyakran alkalmazott módszer a **standardizálás** módszere, amikor a vizsgált változónkból standardizált változót képezünk. A standardizálást abban az esetben is használjuk, amikor a különböző skálákon (mértékegységgel) mért változóinkat egységesíteni, egy alapra hozni szeretnénk. Vannak statisztikai módszertanok, melyek megkövetelik, hogy a vizsgálathoz a különböző mértékegységű változóinkat egységesítsük, ami gyakorlatban a standardizálást jelenti.

A **standardizálás** azt jelenti, hogy a várható értéket kivonjuk a valószínűségi változó értékéből, és a különbséget elosztjuk a szórással, így egy **standard normális eloszlású valószínűségi változót** (jele: z) kapunk eredményül. Képletben:

$$z = \frac{x - \mu}{\sigma}$$

A standardizálás eredményeként kapott standard normális eloszlású valószínűségi változó várható értéke zérus, szórása egységnyi, azaz $N(0,1)$. Mindkettő - a normális és a standard normális eloszlású- valószínűségi változó sűrűségfüggvénye ún. harang-görbével, **Gauss-görbével (6/23. ábra)** jellemezhető. Standard normális eloszlás esetén mind a valószínűségi változók, mind a hozzájuk rendelhető valószínűségek táblázatba foglalhatók, melyek segítségével a kapott értékek könnyen és viszonylag gyorsan felhasználhatók gyakorlati problémák megoldására, bár a számítógépek korában az ehhez hasonló táblázatok szép lassan kikopnak a gyakorlati alkalmazások repertoárjából.



6/23. ábra: Fontosabb valószínűségek a z függvényében

A várható értéktől egységnyi szórással eltérő intervallum – és ez nemcsak a standard, hanem az általános normális eloszlás esetére érvényes – és a valószínűségi görbe által bezárt terület 68,8 %-os valószínűséget reprezentál. A kétszeres szórás által meghatározható intervallumhoz tartozó valószínűség 95,5 %; míg a háromszoros szórással lefedhetjük a vízszintes tengely és

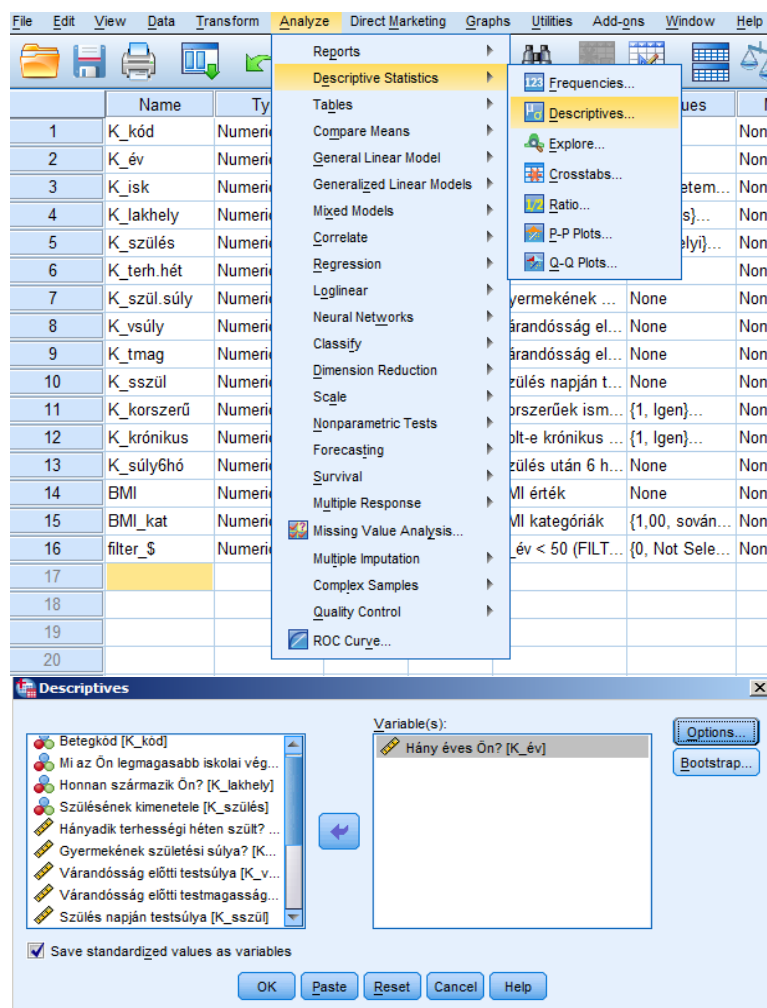
a görbe által meghatározható teljes területet, szinte a teljes valószínűséget (99,9 %). Mivel a standard normális valószínűségi változó sűrűségfüggvénye szimmetrikus, így elegendő a 0 és a pozitív végtelen közé eső számokhoz tartozó valószínűségi értéket meghatározni.

Gyakorlatban ez azt jelentheti, hogy a standardizált értékek -3 és +3 közötti értéket kell felvennie (a normális eloszlás meglétéhez). A standardizált normál eloszlás elmélete alapján, ha a mintanagyság 80 vagy annál kisebb, akkor a 2,5-nél nagyobb standardizált értékekkel rendelkező megfigyelések kiugró (*outlier*) értéknek minősülnek, míg az ennél nagyobb minták esetében a határérték a 3 (Sajtos – Mitev 2007). Ennek nyomán vizsgáljuk.

Gyakorló feladatként, vizsgáljuk meg az életkor változó kiugró értékeit standardizálás módszerével.

Első lépésben a szűrési feltételek - amennyiben nem tettük meg - állítsuk vissza, az alapbeállításra, vagyis ne legyenek szűrési feltételek (*DATA/SELECT CASES/ALL CASES*).

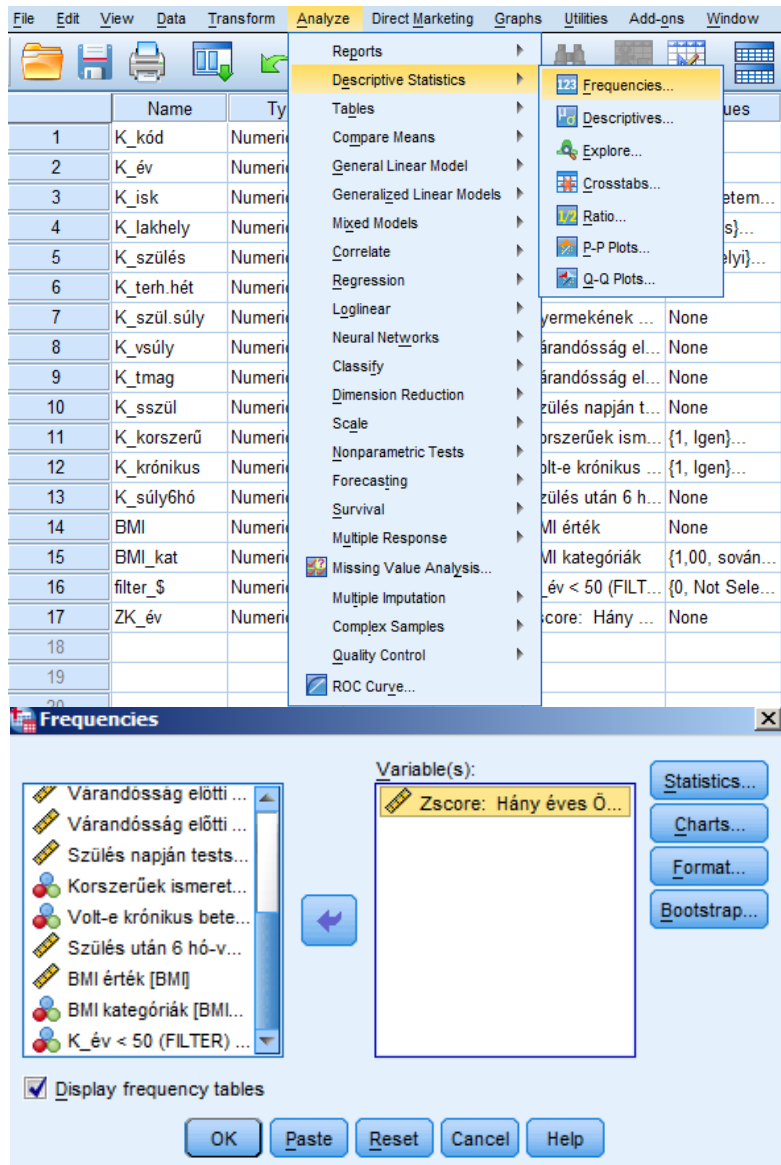
Ezt követően a változó standardizálása következik, melyet az *ANALYZE* menü *DESCRIPTIVES* menüpontjának segítségével végezhetünk el.



6/24. ábra: A standardizálás beállításai

A beállításoknál a bal oldali ablakban a változóinkat láthatjuk, ahol az életkor változót a nyíl segítségével mozgassuk a célváltozókat tartalmazó jobb oldali ablakba. Ezt követően pipáljuk a *SAVE STANDARDIZED VALUES AS VARIABLES* négyzetet, mely által a program létrehoz egy új változót ZK_év néven, mely a standardizált értékeket tartalmazza.

Annak érdekében, hogy a standardizált értékeket megvizsgáljuk, szükséges egy gyakorisági táblát létrehozni. A gyakorisági táblázatot az *ANALYZE* menü *FREQUENCIES* opciójában találjuk.



6/25. ábra: A gyakorisági táblázat beállításai

A beállításoknál nincsen egyéb dolgunk, mint a standardizált változónkat a középső nyíllal a jobb oldali ablakból a bal oldali ablakba mozgassuk. A *DISPLAY FREQUENCY TABLES* négyzet alapbeállításként pipálva van, így a gyakorisági táblázat létrejön, miután az OK gombot aktiváljuk.

Zscore: Hány éves Ön?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid -4,1721	1	1,4	1,4	1,4
-3,9238	1	1,4	1,4	2,9
-3,36754	1	1,4	1,4	4,3
-3,34271	1	1,4	1,4	5,7
-3,31787	3	4,3	4,3	10,0
-2,9304	3	4,3	4,3	14,3
-2,26821	3	4,3	4,3	18,6
-2,24337	7	10,0	10,0	28,6
-2,21854	5	7,1	7,1	35,7
-,19370	3	4,3	4,3	40,0
-,16887	3	4,3	4,3	44,3
-,14404	3	4,3	4,3	48,6
-,11920	2	2,9	2,9	51,4
-,09437	5	7,1	7,1	58,6
-,06954	4	5,7	5,7	64,3
-,04470	4	5,7	5,7	70,0
-,01987	2	2,9	2,9	72,9
,00497	3	4,3	4,3	77,1
,02980	2	2,9	2,9	80,0
,05463	3	4,3	4,3	84,3
,07947	2	2,9	2,9	87,1
,10430	3	4,3	4,3	91,4
,12914	2	2,9	2,9	94,3
,15397	1	1,4	1,4	95,7
,17880	1	1,4	1,4	97,1
,22847	1	1,4	1,4	98,6
8,15051	1	1,4	1,4	100,0
Total	70	100,0	100,0	

Kiugró érték

6/26. táblázat: Az életkor standardizált változójának gyakorisági

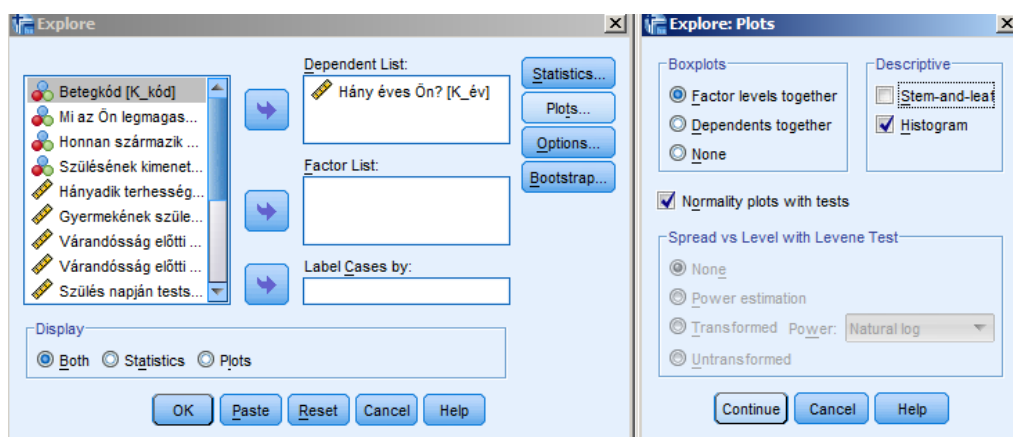
Az adatbázisunk elemszáma összesen 70 fő ($n=70$), így a 2,5-nél nagyobb standardizált értékek kiugró értéknek minősülnek. A táblázatot megvizsgálva látható, hogy egy ilyen létezik (8,15). Ezt követően a kiugró értékünk kizárása következik. A gyakorlatban sokszor előfordul, hogy a kiugró értékeink miatt nem tud megvalósulni a változónk normális eloszlása. A kiugró értékek kiszűrését követően lehetséges, hogy a változó normális eloszlásúvá válik.

A kiugró értékek vizsgálata során is említésre került egy szintén az adatok transzformációjánál is megjeleníthető módszer, melyet egységesen **normalitásvizsgálatnak** nevezhetünk. A normalitásvizsgálat alkalmas arra, hogy két valószínűségi változó eloszlását összehasonlítsuk, vagy ellenőrizzük, hogy egy valószínűségi változónak az eloszlása, az általunk feltételezett eloszlásból (normális eloszlásból) származik-e.

A normalitásvizsgálatának léteznek kvantitatív (numerikus) és grafikus módszerei. A grafikus módszerek közül a leggyakrabban a normál eloszlás görbét tartalmazó hisztogrammal, valamint a kvantilis-kvantilis (Q-Q Plot) ábrázolással ismerkedünk meg.

A grafikus módszerek ábrája alapján akkor tekinthetünk egy változó normális eloszlásúnak, ha a változó eloszlásának alakja jól lefedi a normál eloszlás haranggörbét (Histogram), valamint a Q-Q ábrán a hipotetikus normál eloszlás egyenesére is jól illeszkednek az adatok. A numerikus módszerek közül a Kolmogorov-Szmirnov, valamint Shapiro-Wilk tesztek mutatjuk be. Ez utóbbit akkor érdemes alkalmazni, amikor viszonylag kis minta áll rendelkezésünkre, vagyis a minta elemszáma kevesebb, mint 50. Amennyiben a számított szignifikancia érték magasabb, mint 5 %, akkor elmondható, hogy a változó normál eloszlást követ. Adattranszformáció segítségével gyakran sikerül a normális eloszlástól eltérő változót normális eloszlásúvá alakítani.

Gyakorló feladatként, vizsgáljuk meg, hogy az eredeti életkor változó normális eloszlásból származik-e. A grafikus és numerikus vizsgálatokat megtaláljuk az *ANALYZE* menü *DESCRIPTIVE STATISTICS* menüpontjának *EXPLORE* opciójában.



6/27. ábra: A normalitásvizsgálat numerikus és grafikus beállításai

A függő változó dobozban a felső nyíl segítségével az életkor változó mozgassuk, majd a Plots gombra klikkelve a jobb oldali dobozban (*DESCRIPTIVE*) válasszuk a *HISTOGRAM* ábrát, valamint pipáljuk ki a *NORMALITY PLOTS WITH TESTS* négyzetet. Ezt követően

nyomjuk meg a *CONTINUE* gombot. Miután további beállításokra most az alapbeállításokon túl nincsen szükségünk, nyomjuk meg az *OK* opciót.

Ezt követően az *OUTPUT* nézetben több számítási eredményeket tartalmazó táblázat és ábra is megjelenik, melyek közül, most a számunkra relevánsat ismertetjük.

6/2. táblázat: A normalitás vizsgálat numerikus eredményei

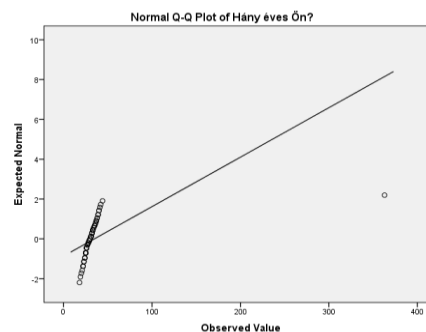
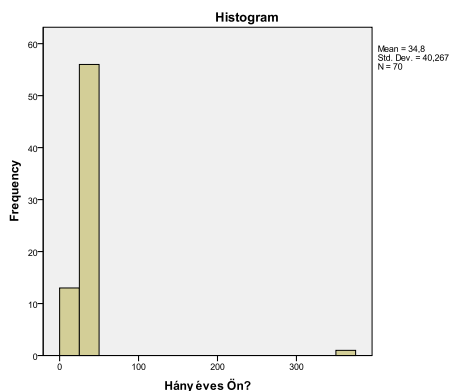
Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Hány éves Ön?	,400	70	,000	,205	70	,000

a. Lilliefors Significance Correction

Forrás: saját szerkesztés

Látható, hogy a mindkét tesztnél (Kolmogorov- Szmirov és a Shapiro- Wilk) a szignifikancia értékek alacsonyabbak, mint 0,05, így a normalitás nem teljesül.



6/28. ábra: Az életkor normalitás vizsgálatának grafikus ábrái (hisztogramm, Q-Q ábra)

A grafikus ábrák megerősítik, hogy a változó eloszlása nem nevezhető normálisnak, hiszen a hisztogram ábrája nem fedi le a haranggörbét, valamint a Q-Q ábra elemei sem illeszkednek az egyenesre. Amennyiben az okokat keressük a korábban bemutatott Boxplot ábra szolgáltatathat megfelelő információt, hiszen látható, hogy kiugró érték szerepel az adatsorban. Ennek kizárását követően (*DATA/SELECT CASES/ IF/ K_év<50*) futassuk le ismét a normalitás vizsgálatot.

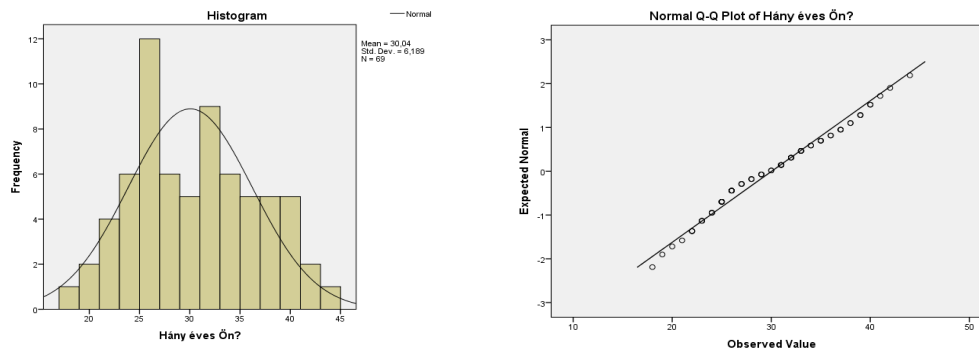
6/3. táblázat: A kiugró érték kiszűrése után lefutott vizsgálat eredményei
Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Hány éves Ön?	,106	69	,055	,977	69	,220

a. Lilliefors Significance Correction

Forrás: saját szerkesztés

Mindkét numerikus tesztnél a szignifikancia érték meghaladja a kritikus értéket (0,05), így a változó normális eloszlásúnak tekinthető, melyet a grafikus ábrák is megerősítenek.



6/29 .ábra: A kiugró érték kiszűrése után lefutott vizsgálat grafikus ábrái

A grafikus módszerek alapján mindegyik változó normál eloszlásúnak minősíthető, az eloszlások alakja jól lefedi a normál eloszlás haranggörbéjét, valamint a Q-Q ábrán (6/29. ábra) a hipotetikus normál eloszlás egyenesére is jól illeszkednek az adatok¹⁰.

Talán ennek az illusztratív példának a segítségével is alátámasztást nyert, hogy a gyakorlati adatelemzés során a legtöbb kutató az *outlier* értékek kiszűrését követően vizsgálja a normalitást.

¹⁰ Tekintve, hogy a változó esetében a ferdeségi és csúcsossági értékek minimális mértékben térnek el nullától abszolút értékben ezért ezek a mutatók alátámasztják a normál eloszlás jelenlétét.

7. LEÍRÓ STATISZTIKA, STATISZTIKAI TÁBLÁZATOK, STATISZTIKAI ÁBRÁK (Ács Pongrác, Raposa László Bence)

7.1. Elméleti alapok, leíró statisztikai mutatószámok

A leíró statisztika mára a gyakorlati adatelemzés során egy kiértékelési eljárást jelent, mely során a mintánk egyes változóit, a kérdőívben szereplő kérdéseket jellemzik. A statisztika nemzetközi irodalma szerinti csoportosítás alapján megkülönböztethetjük: a *leíró statisztikát*, a *következtetési statisztikát* és a *statisztikai döntéelméletet*.

A leíró statisztika alapvetően a numerikus információk összegyűjtését, az információk összegzését, és tömör jellemzését szolgáló módszereket foglalja magában. A leíró statisztika legfontosabb területei:

- az adatgyűjtés,
- az adatok ábrázolása,
- az adatok csoportosítása, osztályozása,
- az adatokkal végzett egyszerűbb aritmetikai műveletek,
- az eredmények megjelenítése.

A leíró statisztikának (*descriptives*) egyik priori célja, hogy helyzetképet, összefoglaló információt adjanak a minta jellemzőiről. Gyakran nevezik alapstatisztikának (*basic statistics*), vagy egyváltozós elemzéseknek is. Szinte minden adatbázis gyakorlati adatelemzésének az első lépése, amikor az ismérveket egymástól függetlenül alapvetően elemszáma, középértékei, adatainak változékonysága mentén vizsgáljuk.

7/1. táblázat: A leggyakrabban alkalmazott leíró statisztikai mutatószámok

Középértékek/helyzetmutató számok	Szóródási mérőszámok	Eloszlási mutatószámok	Egyéb mutatószámok
Számtani átlag (mean)	Terjedelem (range)	Csúcsosság (kurtosis)	Összeg (sum)
Medián (median)	szórás (standard deviation)	Ferdeség (skewness)	Elemek száma (number of cases)
Módusz (mode)	Variancia (variance)		Minimum (minimum) Maximum (maximum)

Forrás: Sajtos – Mitev (2007) nyomán saját szerkesztés

KÖZÉPÉRTÉKEK

A mennyiségi ismérvértékek, számadatok a sokaság lényeges tulajdonságainak hordozói, önmagukban is igen hasznos tájékoztatást adnak a vizsgált jelenségről, folyamatról. Az adatok rendezése, csoportosítása nagymértékben megkönnyíti a megértést. A számszerű információk azt is lehetővé teszik, hogy segítségével általános, tömör jellemzést adjunk, egyetlen számadatba sűrítsük fontos jellemzőjüket. Ezt a számadatot **középértéknek** hívjuk. A középérték **az azonos fajta adatok tömegének számszerű jellemzője**. A középértékek két nagy csoportját szokás megkülönböztetni: számított középértékek (átlag) és a helyzeti középértékek (medián, módusz).

Számtani átlag (MEAN)

A **számtani átlag** az a szám, amelyet az átlagolandó értékek helyébe téve azok összege azonos marad.

A definícióból közvetlenül adódik, hogy a számtani átlag (\bar{x}) a megfigyelt értékek ($x_1, x_2, x_3, \dots, x_n$) összegének, és az elemek számának (n) hányadosa:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

itt fel kell használni ezeket a gyakoriságokat, amit a **súlyozott számtani átlag** formulájával érünk el.

$$\bar{x} = \frac{\sum_{i=1}^k f_i x_i}{\sum_{i=1}^k f_i} = \frac{\sum_{i=1}^k f_i x_i}{n}$$

A számtani átlag a leggyakrabban alkalmazott középérték, azonban nem minden esetben alkalmazható, főleg a kiugró értékekre való érzékenysége miatt. Alkalmazása főleg az arány és intervallum skála típusú adatoknál a legrelevánsabb.

Medián (MEDIAN)

A medián a mennyiségi ismérvek azon értéke, amelynél ugyanannyi kisebb, mint amennyi nagyobb érték fordul elő. Meghatározása egyszerű gyakorisági sor esetén könnyű, hiszen a

rangsorba rendezett középső tag értéke. A medián meghatározásának első lépéseként a számértékeket rangsorba rendezzük és

- ha n páratlan, akkor az $(n+1)/2$. sorszámú egyed ismervváltozatának értéke lesz a medián

$$Me = x_{\left(\frac{n+1}{2}\right)};$$

- ha n páros, akkor az $n/2$. és $(n/2)+1$. egyed ismerv-változatainak egyszerű számtani átlaga

$$Me = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}.$$

lesz a medián:

A mediánt leggyakrabban a sorrendi (ordinális) skálán alkalmazzuk.

Módusz (MODE)

A módusz az ismervértékek tipikus, leginkább jellemző értékét jelöli, legtipikusabb érték. Diszkrét értékekkel rendelkező mennyiségi ismerv módusza a sokaságban leggyakrabban előforduló ismervérték. Folytonos mennyiségi ismerv módusza ott található, ahol az előforduló értékek legjobban sűrűsödnek, ahol a gyakorisági görbe maximuma van. Alkalmazása leginkább a diszkrét, kategorizált változók esetén megfelelő, de alkalmazható a skála típusú változók esetében is.

Látható, hogy a skálatípusoknak megvan a legmegfelelőbb középértéke, vagyis a nominális skáláknál a módusz, az ordinális skáláknál a medián, az arányskálák esetében az összes középérték alkalmazható.

SZÓRÓDÁSI MÉRŐSZÁMOK

Láthatóvá vált, hogy a középértékek egyike sem felel meg tökéletesen minden elvárásnak ezért, kívánatos lehet a további árnyaltabb elemzés is. Szóródásnak nevezzük a statisztikában az adatok (általában a mennyiségi ismervértékek) eltérését egymástól, vagy egy meghatározott, a sokaság egészét jellemző értéktől. A szóródás vizsgálata igen fontos helyet tölt be a statisztika módszertanában, szinte valamennyi módszer kapcsolódik hozzá.

A szóródás jelenségének gyors mérésére természetesen jól alkalmazhatók azok a mutatószámok, amelyek csak néhány, a helyzetüknél fogva jelentős számadatot hasznosítanak. Az egymástól eltérő, tehát szóródó adatok egy-egy középértéktől, így a számtani átlagtól, mint közepes értéktől is eltérnek. Mindez azt is jelenti, hogy a szóródás mérésére szerkesztett mutatószámok egy része ezt a tulajdonságot használja fel.

Valamennyi szóródást mérő mutatószámmal szemben megfogalmazódik az a követelmény, hogy értékük a szóródás hiánya esetén nulla, a szóródás megléte esetén nullától különböző számérték legyen.

A szóródás jelenségét számtalan mérőszám vizsgálja, ezek közül a legelterjedtebbek:

- **szóródás terjedelme (R),**
- interkvartilis terjedelem (TQ),
- **szórás (σ), variancia (σ^2)**
- relatív szórás (V)

A szóródás terjedelme (RANGE): az előforduló legnagyobb és legkisebb érték különbsége:

$$R = x_{\max} - x_{\min}$$

Kizárólag arány és intervallumskálán alkalmazzuk. A szórás terjedelménél szokás még az interkvartilis terjedelmet is említeni. Az **interkvartilis terjedelem** azt az intervallumot jelöli, ahol az összes érték középső 50 %-a helyezkedik el. Az interkvartilis terjedelem képlete:

$$TQ = Q_3 - Q_1$$

A leggyakrabban alkalmazott szóródási mérőszám a **szórás (STANDARD DEVIATION)**, ami az egyes értékek számtani átlagtól való eltéréseinek négyzetes átlagát jelenti.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Ha gyakorisági sorból számítjuk a szórást, a mutatószám súlyozott formáját kell alkalmazni:

$$\sigma = \sqrt{\frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{\sum_{i=1}^k f}} = \sqrt{\frac{1}{n} \sum_{i=1}^k f_i (x_i - \bar{x})^2}$$

A szórás négyzetét **varianciának** (σ^2) hívjuk. Önálló tartalommal nem bír, bizonyos statisztikai eljárásokban kulcsfontosságú mutatószám, arány és intervallumskálán esetén használjuk.

A szóródás eddig megismert mérőszámai a mennyiségi ismerv mértékegységében fejezik ki a szóródás nagyságát. Sok esetben szükség lehet arra, hogy elvonatkoztassunk a mértékegységektől, és ezáltal összehasonlíthatóvá tegyük a különböző jelenségek, különböző mértékegységben kifejezett szóródását, erre való a **relatív szórás**, ami kifejezi, hogy az egyes értékek átlagosan hány %-kal térnek el az átlagtól.

ELOSZLÁSI MUTATÓSZÁMOK

Ferdeség (*SKEWNESS*): egy eloszlás horizontális alakját leíró mutatószám. Jobb oldali aszimmetria esetén (jobbra ferde eloszlás) az érték pozitív, míg bal oldali aszimmetria (balra ferde) esetén a számított mutató negatív értéket vesz fel. A gyakorlati életben többször a jobb oldali aszimmetriával találkozunk.

Csúcsosság (*KURTOSIS*): az eloszlás alakját számszerűsítő mutatószám. Ideális esetben, normális eloszlásnál az értéke nulla, míg pozitív értéknél csúcsos, negatívnál lapult eloszlásról beszélünk.

EGYÉB MUTATÓSZÁMOK

Az egyéb mutatószámok közül kizárólag a SPSS program leíró statisztikai moduljában található mutatószámokat említjük.

Összeg (*SUM*), jelenti az elemek összegét, vagyis kumulált értéküket.

Esetek száma (*NUMBER OF CASES*) a vizsgálatba bevont megfigyelt egyedek száma jelenti.

Minimum (*MINIMUM*): az egyedek értékei közül a legkisebb elem.

Maximum (*MAXIMUM*): az egyedek értékei közül a legnagyobb elem.

Az adatrögzítéskor a figyelmen kívül hagyott adatok szűrésekor jól alkalmazható a minimum és maximum szélsőértékek, hiszen a valós adattartományon kívüli értékek előkerülhetnek.

Az SPSS program segítségével három elérési útvonalon keresztül van lehetőségünk a leíró statisztika elkészítésére az *ANALYZE/DESCRIPTIVE STATISTICS/DESCRIPTIVE* vagy az *ANALYZE/DESCRIPTIVE STATISTICS/FREQUENCIES* vagy az *ANALYZE/DESCRIPTIVE STATISTICS/EXPLORE* modulokkal.

Gyakorló feladatként, elemezzük a leíró statisztikai mutatószámainak segítségével a kisgyermek születési súlya változót.

Annak eldöntésére, hogy melyik módszerrel számítjuk a leíró statisztikai módszereket kötelezően alkalmazandó szabály nincsen. A *DESCRIPTIVE* modult leggyakrabban az intervallum vagy arányskálán (SPSS-ben skála) mért változók esetében használjuk, abban az esetben, ha gyakorisági táblára nincsen szükségünk. A *FREQUENCIES* modul leginkább nominális vagy ordinális skálán mért változók esetében használandó, amikor gyakorisági táblára és grafikus ábrázolásra is szükségünk van. Természetesen ebben az opcióban is vizsgálhatunk arány vagy intervallumskálán mért változókat, azonban talán az

eredményközlés során készített összefoglaló táblázat nem olyan látványos, mint a *DESCRIPTIVE* modulban. Az *EXPLORE* modulban a már megismert módszertanokon túl, leíró statisztikai mutatókat is számoltathatunk, ahol a mintát részsokaságokra bonthatjuk.

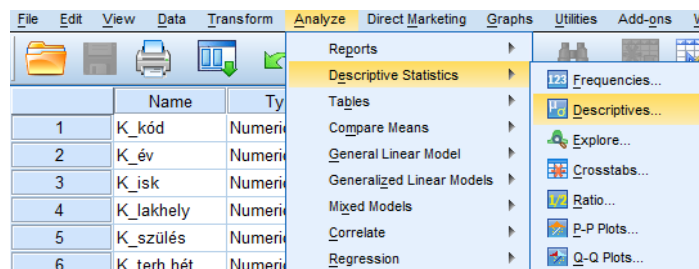
A következő táblázat Kovács (2004) alapján került közlésre Sajtos és Mitev (2007) könyvében, ahol a összefoglalták az SPSS programban található utasítások alapján, a leíró statisztika eszköztárát a megfelelő mérési skálák vonatkozásában (a táblázatban a D jelöli a *DESCRIPTIVES* és az F a *FREQUENCIES* modult).

7/2. tábla: A leíró statisztikában használt mutatók és skálatípusok összefüggése, az SPSS-ben használt modulok Sajtos- Mitev (2007) nyomán

Mutatók	Nominális skála	Ordinális Skála	Intervallum vagy arányskála
Középértékek	Módusz FREQUENCIES	Medián, (Módusz) FREQUENCIES	Átlagok (MEDIÁN, MÓDUSZ)
Szóródási mutatószámok		Minimum, Maximum DESCRIPTIVES, FREQUENCIES	Minimum, Maximum DESCRIPTIVES, FREQUENCIES
Eloszlási mutatószámok	Gyakoriság, relatív gyakoriság FREQUENCIES	Terjedelem	Szórás, variancia DESCRIPTIVES, FREQUENCIES
Egyéb mutatók			Ferdeség, csúcsosság
Grafikus megjelenítés	Oszlop és kördiagram (gyakoriságra) FREQUENCIES		Hisztogram FREQUENCIES

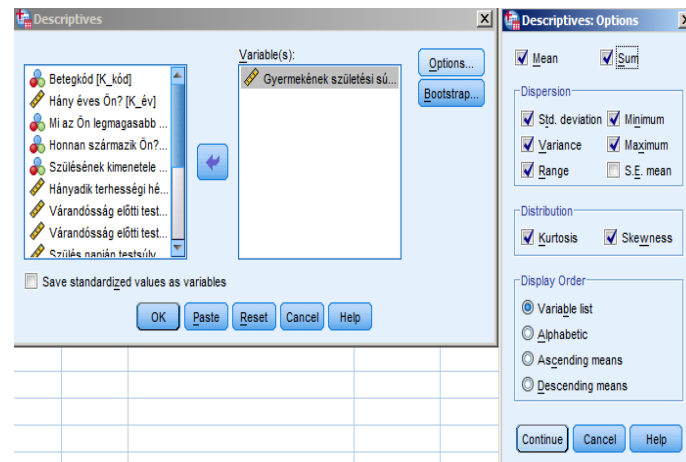
Forrás: Sajtos- Mitev (2007), Kovács (2004) nyomán saját szerkesztés

Első esetben a *ANALYZE/DESCRIPTIVE STATISTICS/DESCRIPTIVE* modult választjuk.



7/1. ábra: A DESCRIPTIVE elérési beállítása

Ezt követően jutunk el a tényleges beállításokhoz. A bal oldali dobozból jelöljük ki a célváltozónkat, mely jelen esetben a gyermek születési súlya lesz. Lehetőségünk van egyszerre több változót is vizsgálni, de amennyiben többet kívánunk egyszerre elemezni, akkor törekedjünk az azonos mérési skálán mért változókra.



7/2. ábra: A DESCRIPTIVE beállításai

A változó kijelölését követően nyomjuk meg az *OPTIONS* gombot, melynek segítségével a szükséges változók kijelölésére van lehetőségünk. Az átlag hibáján kívül (*STANDARD ERROR MEAN*), minden más mutatószám melletti négyzetet pipáljuk ki. Ezt követően *CONTINUE* és *OK* gombok következnek.

7/3. táblázat: A gyermekek születési súlyának leíró (összesítő) statisztikája

	Descriptive Statistics											
	N	Range	Minimum	Maximum	Sum	Mean	Std. Deviation	Variance	Skewness	Kurtosis		
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Gyermekek születési súlya?	70	2050,00	2150,00	4200,00	233080,00	3329,7143	547,43929	299689,772	-.294	,287	-.856	,566
Valid N (listwise)	70											

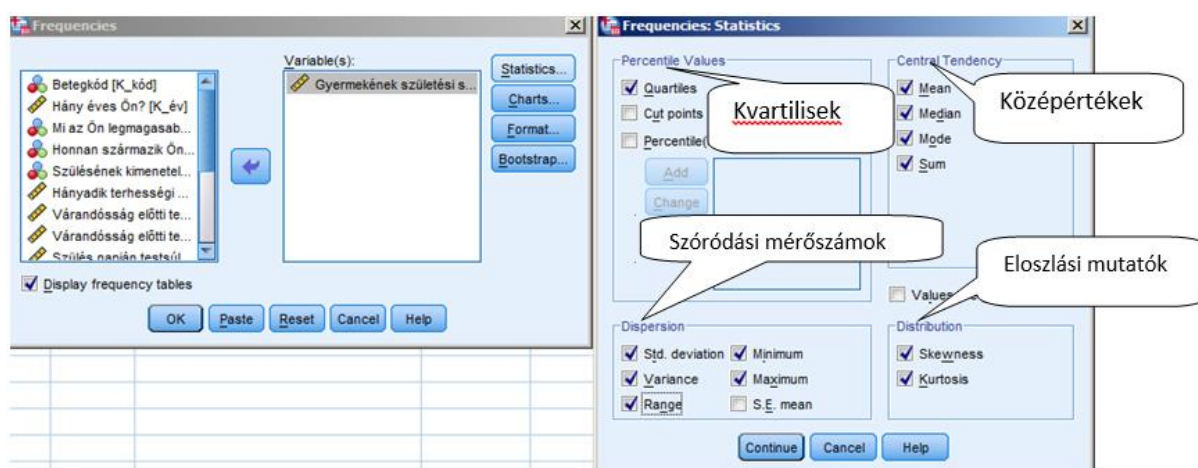
Forrás: saját szerkesztés

A táblázat a következő eredményeket tartalmazza:

- 70 kismama adata szerepel az adatbázisunkban (n=70).
- A változó terjedelme (*RANGE*) 2050 gramm, melyet a maximális és minimális érték különbségéből kapunk.
- *MINIMUM* érték, a legkisebb születési súllyal rendelkező kisbaba 2150 gramm, míg a legnagyobb (*MAXIMUM*) súlyú kisbaba a vizsgált egyedek között 4200 gramm volt.
- Az összeg 233 080 gramm, ami jelenti, hogy a vizsgálatba bevont újszülöttek összesen ekkora súllyal rendelkeztek.

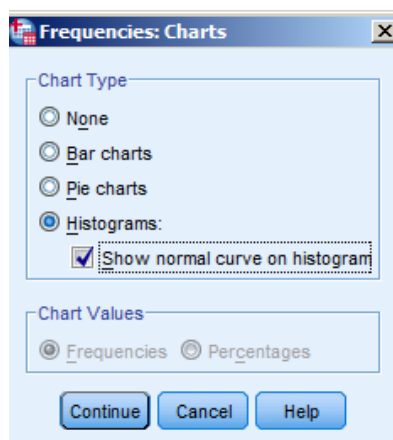
- A kisbabák átlagos súlya (*MEAN*) 3329,71 gramm volt.
- 547,44 gramm volt a babák születési testsúlyának a szórása.
- 299 689,77 gramm a testsúly varianciája.
- A ferdeség (*SKEWNESS*) értéke: -0,29, ami enyhe bal oldali aszimmetriát mutat.
- Csúcsosság (*KURTOSIS*): -0,86 , vagyis lapult eloszlásról beszélhetünk.

Ezt követően vizsgáljuk meg a gyakoriságok segítségével ugyanezt a változót (*ANALYZE/DESCRIPTIVE STATISTICS/FREQUENCIES*).



7/3. ábra: A FREQUENCIES beállításai

A beállítások nyomán válasszuk az előbbieken bemutatott mutatószámokat, illetve jelöljük a kvartiliseket. Ezt követően nyomjuk meg a *CONTINUE* gombot és most válasszuk a beépített grafikus ábra (*CHARTS*) modult.



7/4. ábra: A FREQUENCIES beállítás grafikus ábra (CHARTS) modulja

Az egyszerű ábratípusok közül (oszlop, kördiagram, hisztogram) folytonos változóhoz illően válasszuk a Hisztogramot, melyre egy Gaus-görbét kérünk illeszteni. A *CONTINUE* gomb megnyomását követően, további beállításokat nem eszközölve, nyomjuk meg az *OK*-ot.

7/4. táblázat: A FREQUENCIES eredményei

Statistics

Gyermekekének születési súlya?

N	Valid	70
	Missing	0
Mean		3329,7143
Median		3400,0000
Mode		3000,00 ^a
Std. Deviation		547,43929
Variance		299689,772
Skewness		-,294
Std. Error of Skewness		,287
Kurtosis		-,856
Std. Error of Kurtosis		,566
Range		2050,00
Minimum		2150,00
Maximum		4200,00
Sum		233080,00
Percentiles	25	2907,5000
	50	3400,0000
	75	3812,5000

a. Multiple modes exist. The smallest value is shown

Forrás: Saját szerkesztés

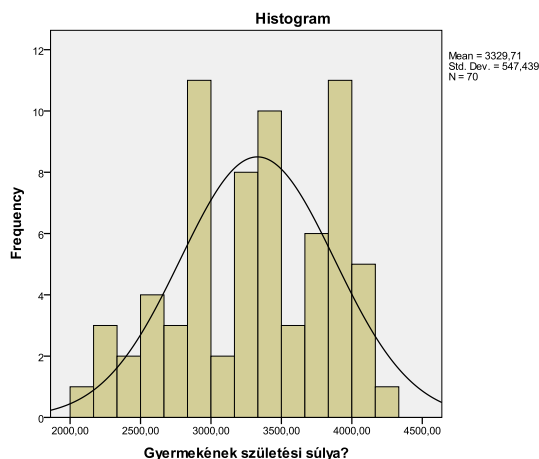
A táblázat adatai a fent ismertetett adatokat tartalmazzák, új adatként a kvartilisek jelentek meg. Az első kvartilis (25 percentilis) a kismamák 25 százaléka kevesebb, mint 2907,5 gramm súlyú kisbabát hozott világra. A második kvartilis a medián, vagyis az esetek 50%-ban a gyermekek súlya kevesebb, mint 3400 gramm, míg az esetek másik felében ennél több. A harmadik kvartilis, vagyis a kisbabák 75%-a kevesebb, mint 3812,5 gramm súllyal született. Új adatként a módusz jelent meg, vagyis a leggyakrabban a gyermekek születési súlya 3000 gramm volt, azonban az érték mellett megjelenő jel, jelzi számunkra, hogy több módusza is van a változónak. Ezt követően a gyakorisági táblázat is megjelenik.

7/5. táblázat: A születési súly gyakorisági táblázata (részlet)

Gyermekének születési súlya?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2150,00	1	1,4	1,4	1,4
	2200,00	1	1,4	1,4	2,9
	2300,00	2	2,9	2,9	5,7
	2400,00	1	1,4	1,4	7,1
	2500,00	1	1,4	1,4	8,6
	2550,00	1	1,4	1,4	10,0
	2600,00	1	1,4	1,4	11,4
	2640,00	1	1,4	1,4	12,9
	2650,00	1	1,4	1,4	14,3
	2700,00	1	1,4	1,4	15,7
	2770,00	1	1,4	1,4	17,1
	2800,00	1	1,4	1,4	18,6
	2850,00	1	1,4	1,4	20,0
	2900,00	3	4,3	4,3	24,3
	2910,00	1	1,4	1,4	25,7
	2980,00	1	1,4	1,4	27,1
	3000,00	5	7,1	7,1	34,3
	3100,00	1	1,4	1,4	35,7
	3120,00	1	1,4	1,4	37,1
	3200,00	4	5,7	5,7	42,9
	3210,00	1	1,4	1,4	44,3
	3300,00	3	4,3	4,3	48,6
	3400,00	4	5,7	5,7	54,3
	3450,00	1	1,4	1,4	55,7

Forrás: saját szerkesztés

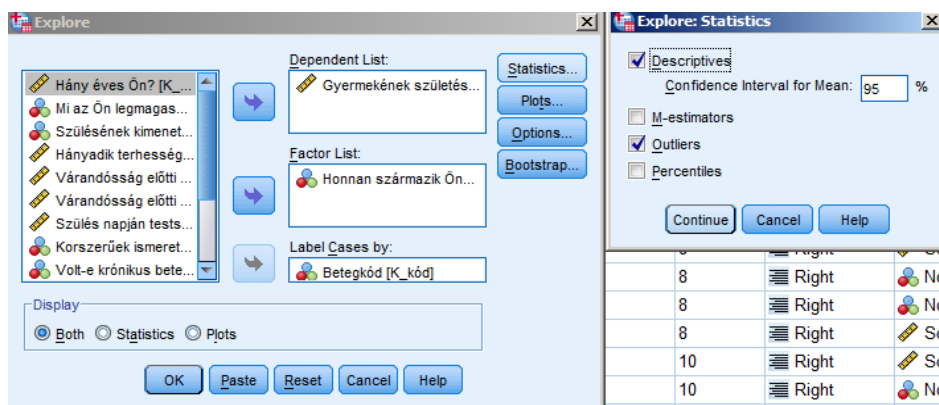
A táblázat első oszlopában (*VALID*) a folytonos változónk ténylegesen előforduló ismérvértékei találhatók. A második oszlop az ismérvértékhez előforduló abszolút gyakoriságokat mutatja (*FREQUENCY*). A harmadik oszlop (*PERCENT*) az ismérvváltozók százalékos megoszlását tartalmazza. A negyedik oszlop (*VALID PERCENT*) az érvényes, felhasználható értékek százalékban kifejezett megoszlását (relatív gyakoriságát) mutatja. Az ötödik (*CUMULATIVE PERCENT*) oszlop a kumulált gyakoriságot mutatja. A táblázat alapján látható, hogy összesen 5 kisbaba volt, aki 3000 grammal született, ami az összes baba 7,1%-a. a 34,3%-os kumulált gyakoriság azt mutatja, hogy a babák 34,3 %-a maximum 3000 grammal született.



7/5. ábra: A kisbabák születési súlyának histogramja

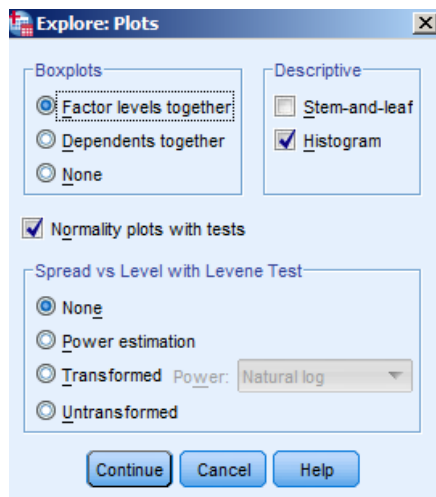
A beállításoknál kért histogram alkalmas arra, hogy a változó normális eloszlását megjelenítse. A histogram nem más, mint egy hézag nélküli oszlopdiagram, amelyen leginkább a folyamatos és diszkrét változóinkat szoktuk jelölni.

Ezt követően vizsgáljuk meg a gyakoriságok segítségével ugyanezt a változót *ANALYZE/DESCRIPTIVE STATISTICS/EXPLORE* modul felhasználásával.



7/6. ábra: Az EXPLORE modul beállításai

A függő változónak jelöljük a gyermekek születési súlyát, míg a csoportosító vagy faktorváltozóként a származási helyet választjuk. Az eseteket a sorszám azonosító alapján vizsgáljuk. Ezekkel a beállításokkal olyan plusz információkhoz is jutunk, hogy az alapsokaságot részekre bontva, a kategóriákra külön-külön kapunk leíró statisztikát, ami pedig az előbbieken megismert adatokat tartalmazza a választott kategória szerint. Természetesen a korábban ismert normalitás vizsgáltra, mind grafikus, mind numerikus változatban lehetőségünk van.



7/7. ábra: Az EXPLORE menü grafikus beállításai

7/6. táblázat: A FREQUENCIES modul eredményei

Descriptives				Statistic	Std. Error
Gyermekének születési súlya?	város	Honnan származik Ön?			
		Mean		3260,8571	83,32822
		95% Confidence Interval for Mean	Lower Bound	3091,5138	
			Upper Bound	3430,2005	
		5% Trimmed Mean		3266,4286	
		Median		3300,0000	
		Variance		243025,714	
		Std. Deviation		492,97638	
		Minimum		2200,00	
		Maximum		4200,00	
		Range		2000,00	
		Interquartile Range		750,00	
		Skewness		-,106	,398
		Kurtosis		-,569	,778
	falu	Mean		3398,5714	100,76943
		95% Confidence Interval for Mean	Lower Bound	3193,7833	
			Upper Bound	3603,3595	
		5% Trimmed Mean		3425,3968	
		Median		3470,0000	
		Variance		355406,723	
Std. Deviation		596,15998			
Minimum		2150,00			
Maximum		4100,00			
Range		1950,00			
Interquartile Range		980,00			
Skewness		-,532	,398		
Kurtosis		-,872	,778		

Forrás: saját szerkesztés

A táblázaton jól látszik, hogy a sokaságot a születési súly alapján két részsokaságra bontotta a származási hely szerint, vagyis az eredmények a városi és falusi kismamák szerint van megjelenítve. A leíró statisztikai adatok között találunk olyan adatot, mely eddig még nem jelent meg, viszont kiemelt fontosságú információkat szolgáltat. A 95% *CONFIDENCE*

INTERVAL FOR MEAN LOWER BOUND/UPPER BOUND statisztikai becslés eredményeit tartalmazza.

A *statisztikai becslés* az ismeretlen alapsokaság valamely konstans paraméterének közelítő jellegű meghatározása. Ilyen paraméterek: várható érték (véges alapsokaságnál, átlag), szórás és az arány. A statisztikai becslések elméleti és gyakorlati bemutatása a következtetési vizsgálatok című fejezetben részletesen megtörténik.

A konkrét példánál maradva a *95% CONFIDENCE INTERVAL FOR MEAN LOWER BOUND/UPPER BOUND* jelenti azt, hogy 95%-os megbízhatóság mellett megállapíthatjuk, hogy a városi kismamák gyermekeinek a születési súlya minimum 3091, 51 gramm és 3430, 20 gramm között lesz.

Gyakran kell ezeket a becsléseket tartalmazó eredményeket grafikusán is megjeleníteni, melynek bemutatását a grafikus ábrázolások során részletesen bemutatunk.

Az *EXPLORE* modul, mint korábban bemutattuk alkalmas a kiugró értékek és a normalitás grafikus vizsgálatára is. A kiugró értékek vizsgálatára a grafikus elemzésen kívül egy táblázat is segítségünkre lehet, ahol az egyes kategóriák öt legnagyobb és öt legkisebb értékét mutatják, és rámutatva a konkrét (*LABELS*, betegkód) esetekre.

7/7. táblázat: Az öt legnagyobb és legkisebb érték táblázata

Extreme Values						
		Honnan származik Ön?		Case Number	Betegkód	Value
Gyermekeinek születési súlya?	város	Highest	1	24	24	4200,00
			2	56	56	4050,00
			3	8	8	4000,00
			4	55	55	3900,00
			5	54	54	3850,00
	falv	Lowest	1	14	14	2200,00
			2	60	60	2400,00
			3	1	1	2500,00
			4	2	2	2600,00
			5	37	37	2700,00
falv	Highest	1	22	22	4100,00	
		2	33	33	4100,00	
		3	67	67	4100,00	
		4	34	34	4080,00	
		5	21	21	4000,00 ^a	
	falv	Lowest	1	15	15	2150,00
			2	61	61	2300,00
			3	31	31	2300,00
			4	62	62	2550,00
			5	36	36	2640,00

Only a partial list of cases with the value 4000,00 are shown in the table of upper extremes.

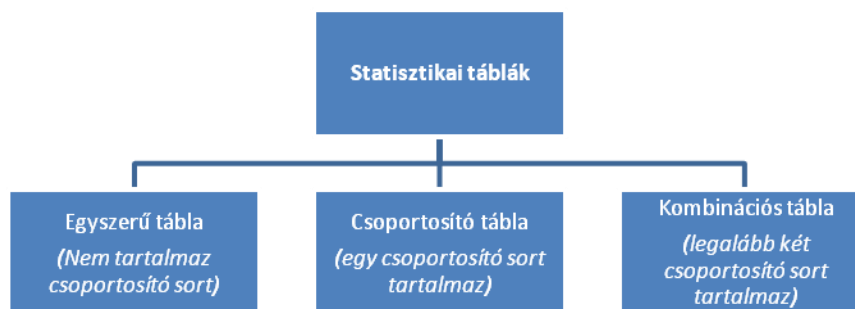
Forrás: saját szerkesztés

Természetesen ez a táblázat segíthet nekünk a szélsőséges, illetve kiugró (*outlier*) értékek meghatározásában. Az ezt követően számított eredmények értelmezését már az adatok

transzformációja című fejezetben részletesen tárgyaltuk, így most ennek ismételt értelmezésétől (normalitás vizsgálat, kiugró értékek vizsgálata) eltekintünk.

7.2. Statisztikai táblák, táblázatok

A statisztikai tábla szintén gyakran előforduló adatprezentációs eszköz, hiszen a célja az adatok rendezése, tömörítése. Ezeknek a tábláknak a fő céljuk, hogy a bennük szereplő adatok egyfajta összefüggésrendszert képezzenek és a vizsgált jelenségről összetett képet tükrözzenek. Statisztikai tábla a statisztikai sorok rendszere, melyben az adatok egy, illetve több ismérv szerint lehetnek felsorolva. A statisztikai táblák statisztikai sorokat (idő-, területi, minőségi, mennyiségi sor) tartalmaznak. A táblákat általában két szempont szerint szokás tipizálni. A dimenziószám szerint leginkább két- vagy háromdimenziós táblákkal találkozhatunk. Ennek eldöntése a táblában található ismérvek (változók) számától függ. A másik tipizálást az ismérvek felsorolásának a célja határozza meg, hiszen történhet összehasonlítás vagy csoportosítás kedvéért. Ennek megfelelően típus szerint a következő táblákat különböztetjük meg:



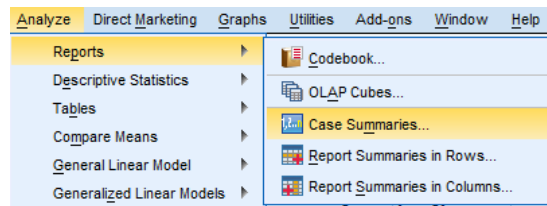
7/8. ábra: A statisztikai táblák csoportosítása

Forrás: Ács (2009)

A statisztikai táblák többsége kombinációs tábla. Abban az esetben, ha a táblában gyakorisági sorok szerepelnek, vagyis a felsorolt adatok gyakoriságok, kontingencia tábláról beszélünk. A statisztikai táblákkal és a grafikus ábrákkal is szemben fontos formai meghatározások léteznek, melynek hiánya csökkentheti a kutatások (diplomamunkák, szakdolgozatok) megítélését. Ezek a formai követelmények: a cím, a forrás és a magyarázó szövegek feltüntetése. Tartalmi követelmény (teljes körűség, besorolhatóság), hogy minden egyednek kell találni kizárólag egy helyet, ahová el tudjuk a rá vonatkozó adatok alapján helyezni.

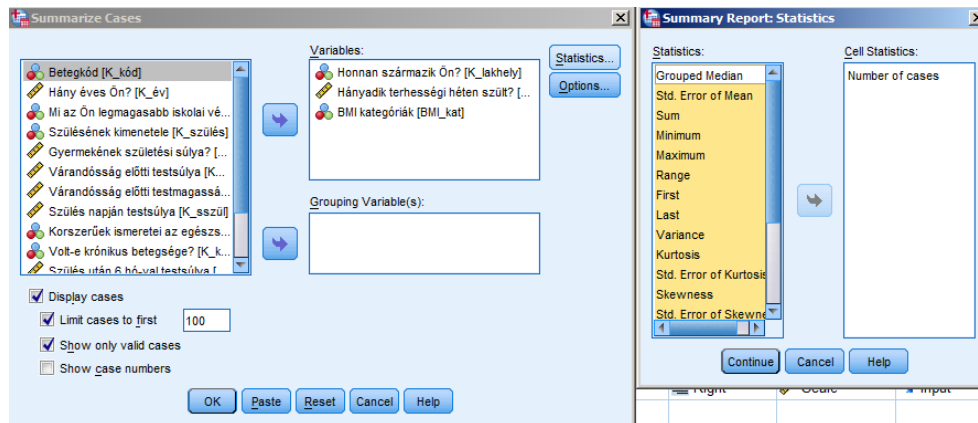
Az SPSS program segítségével több megoldás is kínálkozik a statisztikai táblák elkészítésére.

Az egyik egyszerű módja a táblázat készítésnek, amikor az adatokat (eseteket) rendezett formában kívánjuk bemutatni. Ezt az *ANALYZE* menü *REPORTS* opciójának *CASE SUMMARIES* paneljéből érjük el. Ezt követően a változók beállítása és kijelölése következik.



7/9. ábra: Az egyedek változók szerinti bemutatása

A következő gyakorlati példa segítségével mutassuk be az egyes egyedeket származási hely, születési hét, illetve BMI kategóriák szempontjából.



7/10. ábra: A táblázat beállításai

A bal oldali ábrán tudjuk a felsorolásba bevont változókat kijelölni. A változók megadását követően lehetőségünk van csoportosító változó választására (*GROUPING VARIABLES*) is. A *STATISTICS* gombot megnyomva van lehetőségünk különböző statisztikai mutatószámok listázására is. Az *OPTIONS* panelt választva adhatunk nevet a táblázatunknak (*TITLE*), illetve szöveges megjegyzést (*CAPTION*). A *OK* gomb megnyomását követően a táblázat létrejön.

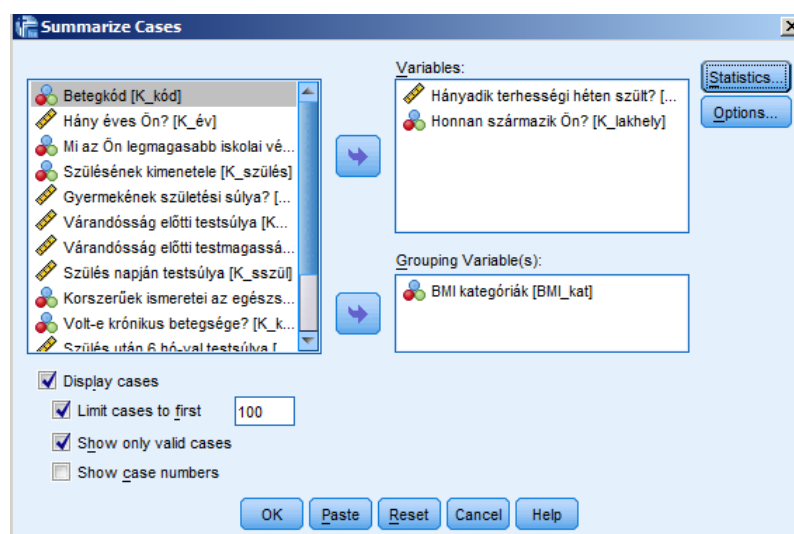
7/8. táblázat: Az egyes egyedek megadott változónkénti felsorolásának táblázata (részlet)

Case Summaries^a

	Honnan származik Ön?	Hányadik terhességi héten szült?	BMI kategóriák
1	város	38,00	normál testalkatú
2	város	35,00	normál testalkatú
3	város	36,00	normál testalkatú
4	város	38,00	normál testalkatú
5	város	37,00	normál testalkatú
6	város	35,00	normál testalkatú
7	város	34,00	normál testalkatú
8	város	36,00	normál testalkatú
9	város	37,00	normál testalkatú
10	város	38,00	normál testalkatú
11	város	39,00	normál testalkatú
12	város	40,00	normál testalkatú
13	város	41,00	normál testalkatú
14	város	38,00	normál testalkatú

Forrás: saját szerkesztés

Gyakorlásként a BMI kategóriák szerinti csoportosításban jelenítsük meg a táblázatunkat.



7/11. ábra: A csoportosító változó bevonásának beállításai

A beállítások ablakban a csoportosító változónál jelöljük ki a BMI kategória változót. Ezt követően az OK gomb megnyomásával a végeredményhez jutunk.

7/9. táblázat: A BMI kategóriák szerinti csoportosítást követően felsorolt esetek táblázata (részlet)
Case Summaries^a

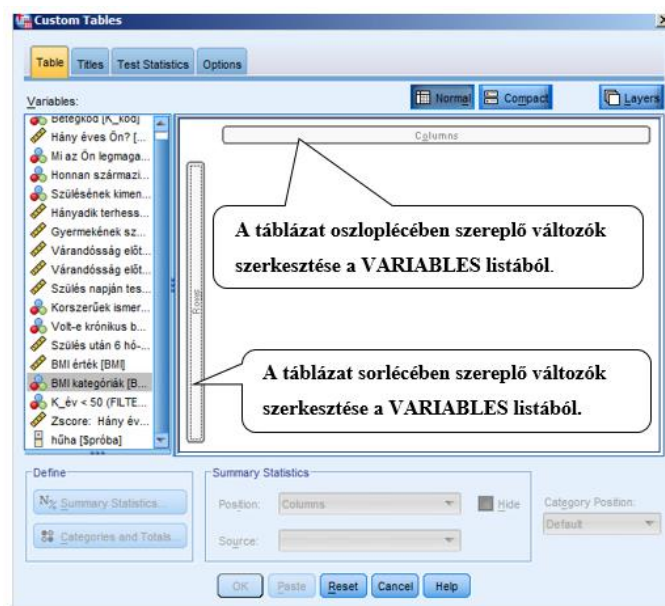
			Hányadik terhességi héten szült?	Honnan származik Ön?
BMI kategóriák	sovány	1	39,00	város
		2	39,00	város
		3	37,00	város
		4	38,00	falu
		5	39,00	falu
	Total	N	5	5
normál testalkatú	1	38,00	város	
	2	35,00	város	
	3	36,00	város	
	4	38,00	város	
	5	37,00	város	

Forrás: saját szerkesztés

A táblázatban jól látható, hogy a BMI kategóriák szerinti felsorolásban jelennek meg a kért változók.

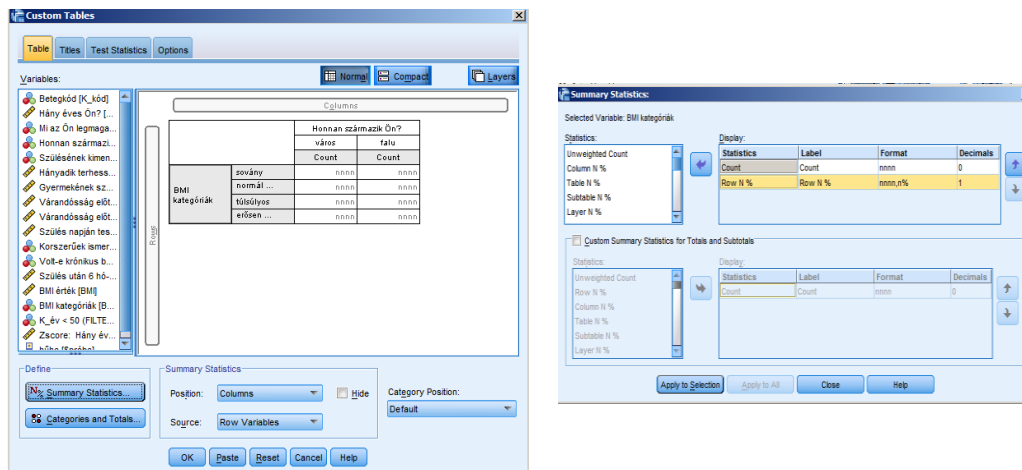
A tipikus táblázatkészítő modult az *ANALYZE* menü *TABLES*, *CUSTOM TABLES* modulján keresztül érhető el. A következő felugró panelnél az *OK* gomb megnyomásával jutunk a beállítások ablakhoz.

A beállításokat tartalmazó ablakban látható alapstruktúra egy kétdimenziós kereszttáblázat, melyben a cellák gyakorisága (*COUNT*) jelenik meg alapbeállításként.



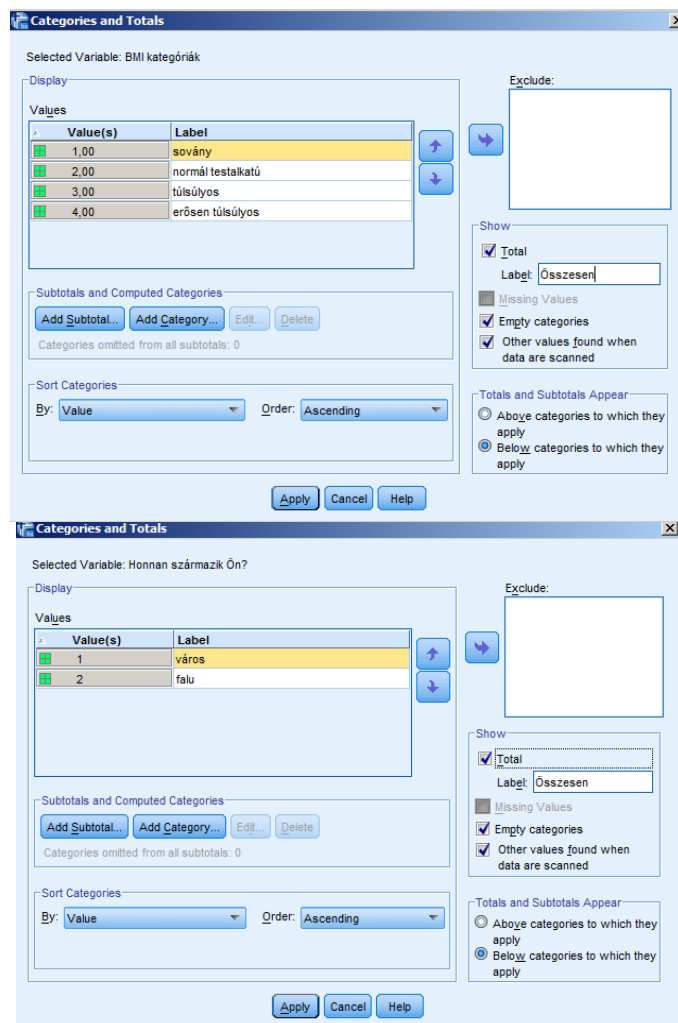
7/12. ábra: Statisztika táblázat beállítás modulja

Gyakorló feladatként, ábrázoljuk egy kombinációs (kontingencia) statisztikai táblázatban a BMI kategóriákban szereplő egyedeket származási helyük bontásában. Tegyük ezt úgy, hogy a BMI kategóriákhoz tartozó egyedek darabszáma mellett jelöljük a %-os megoszlásukat is. Az első beállításként a BMI kategória változót mozgassuk a sor fejlécbe, majd a származási hely változót megfogva az oszlop fejlécre irányítsuk.



7/13. ábra: A változók és mutatószámok kalibrálásának ábrája

Miután a változóinkat a sor és oszlop lécben megjelenítettük állítsuk be a százalékos megjelenítést, melyet a *DEFINE* ablakban a *SUMMARY STATISTICS* opcióban tehetünk meg. Itt a megjelenő panelből a mutatószámok (*STATISTICS*) közül válasszuk a középső nyíl segítségével a *ROW N%* opciót, mely által sor szerinti százalékos arányok fognak megjelenni. Az újabb statisztikai mutatószám táblázatba való integrálását az *APPLY to SELECTION* gombbal végezzük el. Ezt követően szerkesszük kombinációs táblázattá a *CATEGORIES* and *TOTALS* opció segítségével.



7/14. ábra: Az összesen sorok szerkesztésének beállításai

Mivel a kombinációs táblázat, minimum két összesen sor tartalmaz, ezért mindkét változónknál a *TOTAL* dobozban a *LABEL* címkét írjuk át összesen feliratra. Az *APPLY* gomb lenyomását követően a főpanelen az *OK* gomb segítségével kapjuk meg az *OUTPUT* nézetben az általunk szerkesztett táblázatot.

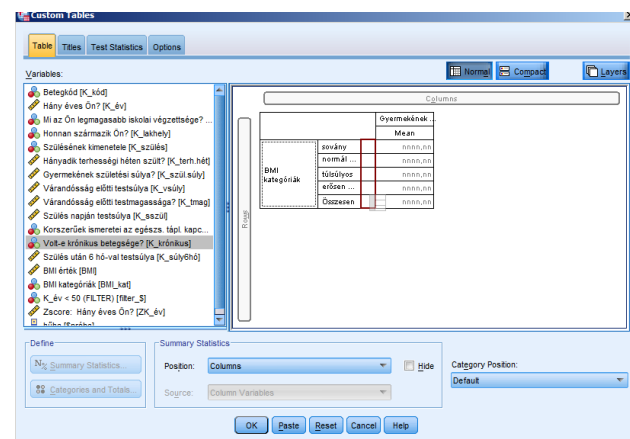
7/10. táblázat: A beállítások nyomán létrejött kombinációs táblázat

		Honnan származik Ön?					
		város		falu		Összesen	
		Count	Row N %	Count	Row N %	Count	Row N %
BMI kategóriák	sovány	3	60,0%	2	40,0%	5	100,0%
	normál testalkatú	19	82,6%	4	17,4%	23	100,0%
	túlsúlyos	8	38,1%	13	61,9%	21	100,0%
	erősen túlsúlyos	5	23,8%	16	76,2%	21	100,0%
	Összesen	35	50,0%	35	50,0%	70	100,0%

Forrás: saját szerkesztés

A táblázatból leolvasható, hogy 5 erősen túlsúlyos kismama volt csupán, aki városból származott (21), ami a városban élők a 23,8%-a. Természetesen a táblázatunkat tovább formázhatjuk, szépíthetjük, amennyiben a táblázatra állva először a baloldali egér billentyűt kétszer, illetve a jobb oldali billentyűt egyszer megnyomjuk. Természetesen van lehetőségünk a táblázat dimenziószámának növelésére, a kategóriák bővítésére.

A következő táblázatban nézzük meg a születési testsúly átlagát a BMI kategóriák és a krónikus betegség előfordulása rétegeképző ismérv tekintetében. A beállítást úgy tudjuk megtenni, hogy a születési testsúlyt az oszlopváltozóba mozgatjuk, majd a BMI kategória változó jobb szélére húzzuk a krónikus betegségre vonatkozó változót, ahol azt elengedve megjelenik az új dimenzió.



		Gyermekeknek ...		
		Mean		
BMI kategóriák	sovány	Volt-e krónikus betegsége?	Igen	nnnn,nn
			nem	nnnn,nn
	normál testalkatú	Volt-e krónikus betegsége?	Igen	nnnn,nn
			nem	nnnn,nn
	túlsúlyos	Volt-e krónikus betegsége?	Igen	nnnn,nn
			nem	nnnn,nn
	erősen túlsúlyos	Volt-e krónikus betegsége?	Igen	nnnn,nn
		nem	nnnn,nn	
	Összesen	Volt-e krónikus betegsége?	Igen	nnnn,nn
		nem	nnnn,nn	

7/15. ábra BMI és krónikus betegségek

A folytonos változók alapbeállításként átlagokat jelenítenek meg, melytől természetesen el tudunk térni a *SUMMARY STATISTICS* opcióban található mutatószámok integrálásával.

Az *OK* gomb lenyomását követően eljutunk a statisztikai táblázathoz.

7/11. táblázat: A bővített statisztikai táblázat

				Gyermekének születési súlya?
				Mean
BMI kategóriák	sovány	Volt-e krónikus betegsége?	Igen	.
			nem	2840,00
	normál testalkatú	Volt-e krónikus betegsége?	Igen	3630,00
			nem	3202,63
	túlsúlyos	Volt-e krónikus betegsége?	Igen	3453,33
			nem	3128,67
	erősen túlsúlyos	Volt-e krónikus betegsége?	Igen	3680,00
			nem	3585,33
Összesen		Volt-e krónikus betegsége?	Igen	3582,50
			nem	3254,81

Forrás: saját szerkesztés

7.3. Statisztikai ábrák, diagramok

A statisztikai valamint az elméleti munka egyik leglényegesebb része a közlés, szemléltetés. Ugyanakkor egy kevésbé hozzáértő, esetleg laikus számára a „végeláthatatlan” adatok sora kevésbé értelmezhető, így fontos hogy eredményeinket értelmezhetően, meggyőzően prezentálhassuk. Ezen adatprezentáció formája és talán legszemléltetőbb eszköze a grafikus ábrázolás. A grafikus ábrázolás mindamelllett, hogy szemléletessé teszi munkánkat, egyben segítheti az egyes összefüggések felismerését, hiszen fontos tulajdonsága, hogy nem csak egyes adatok abszolút nagyságáról, mért adatok szintjeiről ad tájékoztatást, hanem viszonyítási alapot is nyújt és arányokat érzékeltet. Az adatok grafikonokkal történő bemutatása lehetséges koordináta rendszeren belül és kívül egyaránt. A koordináta rendszer számos további információt adhat, ugyanakkor vannak olyan esetek, amikor használata nem indokolt, így nem feltétlenül van rá szükség.

Fontos, hogy a grafikus ábrázolás módját és eszközét mindig úgy válasszuk meg, hogy az, az elérni kívánt célt szolgálja. Az ábrázolás fajtáját mindig ahhoz kell igazítanunk, hogy alkalmas legyen azon jelenség bemutatására, melyet szemléltetni kívánunk. A grafikonok ábrázolási eszközei alapvető mértani elemek, mint a pont, a vonal, a négyzet, a téglalap és a kör valamint ezek variációi (Székelyi – Barna 2005).

A grafikonok segítségével változásokat, megoszlásbeli különbségeket, nagyságok összehasonlítását, időbeli eltérést stb. egyaránt szemléltethetünk (Jánosa 2011; Sajtos – Mitev 2007).

A grafikonok szerkesztésénél a következő tényezőkre szükséges odafigyelni:

- az elnevezés, grafikon cím egyértelmű legyen és tudni lehessen, hogy mi az ábrázolt jelenség, tényező
- a grafikon pontosan milyen időszakra, időtartományra vonatkozik
- a feltüntetett és szemléltetett adatok mértékegységének feltüntetése elengedhetetlen, nélküle nem értelmezhető az ábra
- az értelmezhetőség tekintetében fontos még, hogy amennyiben szükséges grafikonunkat/ábránkat mindig lássuk el jelmagyarázattal

Mielőtt azonban kitérnénk az SPSS programban a grafikus ábrázolás elérésére és gyakorlatára, fontos meghatároznunk, milyen grafikus ábrázolási eszközök állnak rendelkezésünkre, és mit tudunk illetve érdemes velük szemléltetni (bemutatásunkban az egyszerű grafikus eszközöké a főszerep, így az analízis részét képező típusok csak említés szintjén szerepelnek) (Sajtos – Mitev 2007).

„Bar Chart / Oszlopdiaqram”: ezen diagram típus egy kategorizált változó tekintetében ábrázolja a sokaság szerkezetét, struktúráját.

„3D Bar Chart / 3D-s oszlopdiaqram”: ezen diagram alkalmazhatósága ugyanaz mint az oszlopdiaqramé hiszen, gyakorlatilag kinézetükben van csak különbség, mely látványosabb megjelenést ad adatainknak.

„Line Chart / Vonaldiaqram”: elsősorban idősorok ábrázolására alkalmas diagram típus, melyben az adatpontokat egy folytonos vonallal összekötve szemléltethetjük a kívánt jelenséget.

„Area Chart / Sávdiaqram”: az adat görbék által lefedett tartomány, terület megjelenítését szolgálja.

„Pie Chart / Kördiaqram”: adatsokaság összetételére, szerkezetének meghatározására használatos grafikus módszer.

„High Low Chart”: egymáshoz tartozó értékpárok összehasonlítására alkalmas diagram típus.

„Pareto Chart”: ezen diagram típus esetében, egy kombinált grafikus eszközzel beszélhetünk, hiszen a már előzőekben bemutatott oszlopdiaqram és a kumulált összeget bemutató vonaldiaqram együttese alkotja.

„Boxplot Chart”: más néven „dobozábra”, e grafikonról leolvasható a legtöbb deskriptív jellemző, mint a medián, a minimum és maximum értékek valamint a kvartilisek is.

„**Control Chart**”: folyamatirányító grafikon, mellyel ellenőrizhetjük, hogy az adott változó az általunk meghatározott határok között változik, mozog.

„**Scatter Chart**”: mennyiségi ismérvek közötti kapcsolat vizsgálatára alkalmas, annyiban hasonlít a „High Low Chart” eszközre, hogy ez a diagram típus is értékpárokkal értékhármasokkal foglalkozik (itt pontfelhőszerű a megjelenítés).

„**Error Bar Chart**”: mint ahogyan az eszköz nevéből is sejthető, ezen grafikonról leolvasható egy adott változó standard hibája valamint szórása, konfidencia intervallumtól történő eltérése is.

„**Hisztogram**”: egy változó eloszlását mutatja, fontos ugyanakkor, hogy metrikus skálák esetén alkalmazható csak.

„**P-P Plot Chart**”: eloszlásfüggvény összehasonlításra alkalmas diagram típus, melynél egy adott változó függvényét hasonlítja a normális függvényhez.

„**Q-Q Plot Chart**”: egy változó esetén kvartiliseket lehet vizsgálni a normális eloszlás elméleti kvartilisei között.

„**Sequence Chart**”: több esemény időrendi ábrázolása sorban. Alkalmazása idősorok esetén ajánlott.

„**Autocorrelations Chart**”: autókorrelációs vizsgálatra szolgál különböző idősorok között.

„**Crosscorrelation Chart**”: két idősor közötti korreláció vizsgálatára alkalmazható.

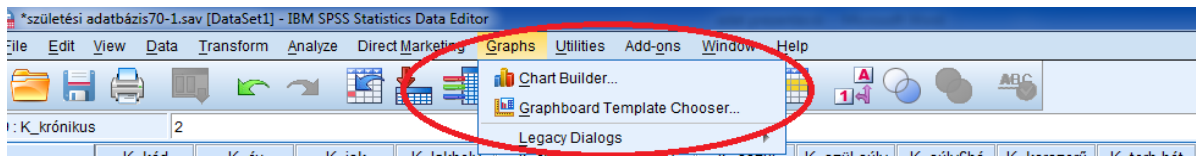
„**Spectral Chart**”: spektrumbecslés esetén alkalmazható, összetett vonaldiagram típus.

„**Population pyramid Chart**”: az epidemiológiában „korfa”-ként is emlegetett népesség piramis ábrázolható segítségével, olyan módon, hogy két egymással ellentétes oszlopdiagramot vagy hisztogramot ábrázol.

„**Mixed Chart**”: kevert diagram típus, melyben két egymástól különböző diagram típust kombinálhatunk egy ábrán, grafikonon belül (Jánosa 2011, Sajtos – Mitev 2007).

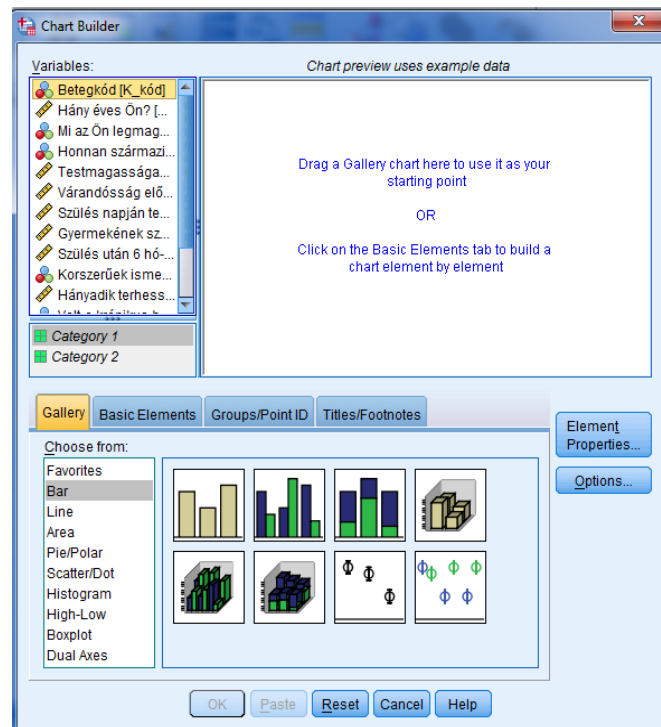
Grafikus ábrázolás a gyakorlatban

Az SPSS-ben a grafikus ábrázolások nagy részét a **GRAPHS** menüpont alatt találjuk, rákattintást követően, ki kell választanunk a számunkra leginkább használható ábrázolási módot, mely a jelenség szemléltetésére is alkalmas. A **GRAPHS** menüponton belül kétféleképpen is készíthetünk diagramot ez a **CHART BUILDER** és a **LEGACY DIALOGS** parancsokkal lehetséges.



7/15. ábra: A GRAPHS menüpont diagram szerkesztési panelje

A *CHART BUILDER* egy rugalmas és elég újszerű szerkesztési módot biztosít számunkra. Amikor először indítjuk el az eljárást, a program megjelenít egy párbeszéd panelt, melyben a mérési skálák beállításaira hívja fel a figyelmünket az egyes változók esetében. Ekkor az *OK* gombra kattintva léphetünk tovább, de még mielőtt ezt megtennénk, bejelöljük a „*Dont show me again*” feliratú rubrikát hisz ebben az esetben nem jelenik meg többet számunkra ez az üzenet.



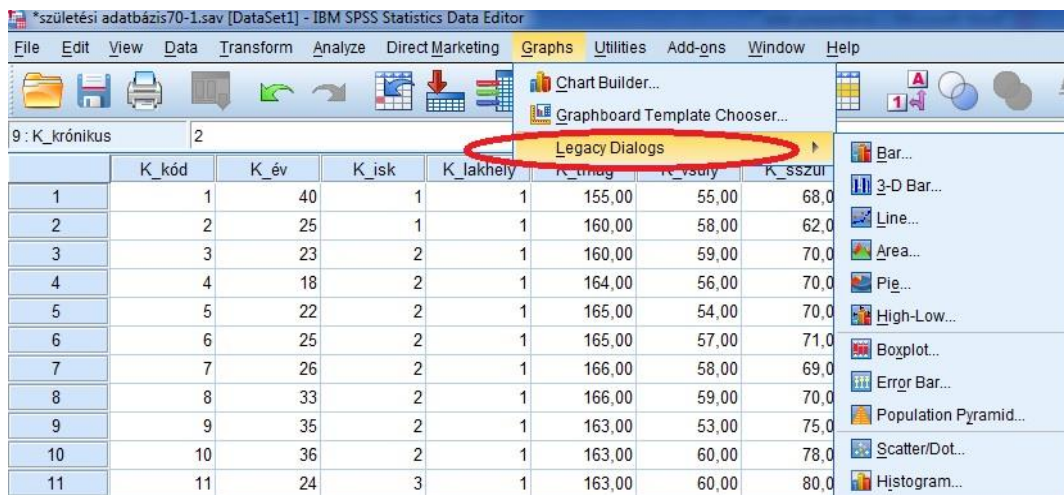
7/16. ábra: CHART BUILDER diagram szerkesztő panel

A bal felső *VARIABLES* mezőben az adatbázisunkba felvitt változókat látjuk. Míg az alul található *GALLERY* fül alatt találjuk az alapvető diagram típusokat, melyek közül kiválaszthatjuk a nekünk legmegfelelőbbet. A diagram felépítésére két lehetőségünk van egyrészt a kívánt változót behúzzuk a „*Chart preview uses example data*” mezőbe majd két kattintással kiválasztjuk a kívánt típust ezután megadjuk a függő és független változókat, melyek függvényében azt vizsgálni szeretnénk. Az adott változót a koordináta rendszer X

(független) és Y (függő) tengelyén helyezhetjük el vagy a „Basic elements” fülre kattintva is felépíthetjük grafikonunkat a legelemibb építőkövekből kiindulva.

A *TITLESS/FOOTNOTES* fülön adhatjuk meg a diagram címét illetve adhatunk hozzá lábjegyzetet ábránkhoz.

Másik diagram beszurási lehetőségünk a *LEGACY DIALOGS* melyre rákattintva egyből kiválaszthatjuk a munkánkhoz szükséges diagram típust. Ezt követően a változó valamint a tengelyek meghatározásával a program elkészíti számunkra a kívánt diagramot (Jánosa 2011).



7/17. ábra: A *LEGACY DIALOGS* diagramszerkesztési panel

A továbbiakban a leggyakrabban használt típusokat (vonalt-, oszlop-, kördiagram illetve a hisztogram) valamint azok gyakorlati kivitelezését egy-egy példával mutatjuk be.

„Vonaldiagram / Line Chart”

Amennyiben egyes adataink között értelmezhető az átmenet, abban az esetben értékpontjainkat egymással összeköthetjük, mely következtében vonaldiagramot kapunk. Ezen esetekben jól látható az egyes időegységek alatt létrejövő változások mértéke, egymáshoz viszonyított aránya is.

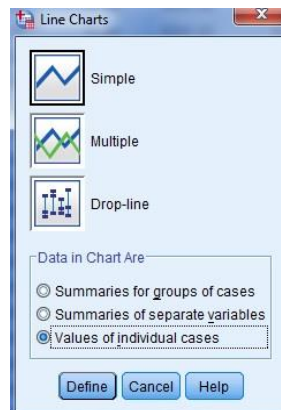
Adatbázisunk a várandósság alatti táplálkozási szokások ismeretét valamint ennek hatását dolgozza fel a testsúlyváltozások tekintetében. Az adatbázis alapját képező kérdőív többféle változót mért fel, a kérdések a következők voltak:

- Betegkód
- Hány éves Ön?
- Mi az Ön legmagasabb iskolai végzettsége?
- Honnan származik Ön?
- Testmagassága?
- Szülés előtti testsúlya?

- Szülés napján mért testsúlya?
- Gyermekének születési súlya?
- Szülés után 6 hónappal mért testsúlya?
- Korszerűnek ítéli meg az egészséges táplálkozással kapcsolatos ismereteit?
- Hányadik terhességi héten szült?
- Volt e krónikus betegsége?
- Szülésének kimenetele?
- Szülés közben érzett fájdalom (1-5 skálán)?

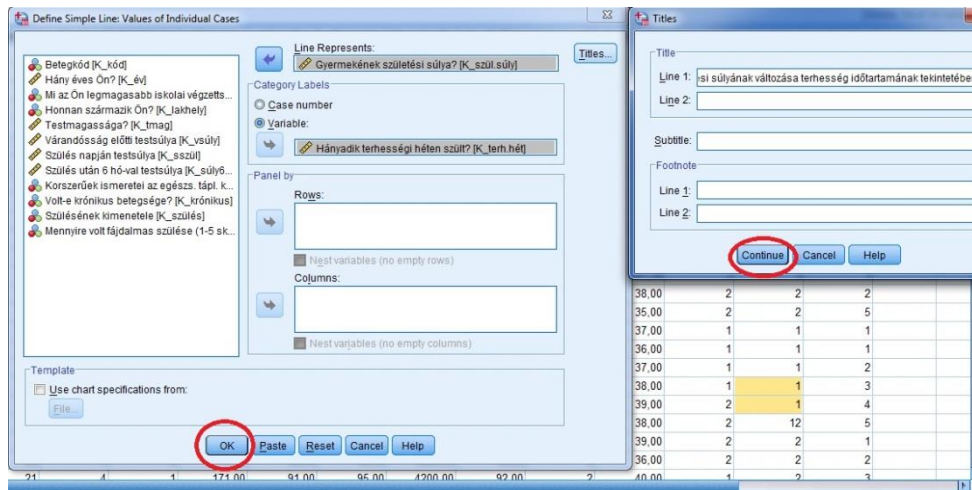
Mivel az előbb említettük, hogy egyes időbeli változások szemléltetésére alkalmas a vonaldiagram így első példánkban, a „gyermek születési súlyát” valamint, a „hányadik terhességi héten szült” adatainkat használjuk fel, ez a két változó oszlopunk van.

Az SPSS-ben több lehetőségünk is van vonaldiagramot készíteni. Az egyik mód, ha a *GRAPH* → *LEGACY DIALOGS* → *LINE CHART* opciót választjuk. Ekkor a vonaldiagram típus kiválasztásának panelét kapjuk, melynél a *SIMPLE* típust választjuk valamint az alsó opciók közül a *VALUES OF INDIVIDUAL CASES* rubrikát jelöljük be, ezt követően a *DEFINE* gombra kattintunk. Fontos megjegyezni, hogy a *SIMPLE* verziót egy adatsor ábrázolásra, a *MULTIPLE* verziót két vagy több adatsor szemléltetésére alkalmazhatjuk. Ha több adatsor összetartozó elemei közötti különbségeket kívánunk vizsgálni, megjeleníteni akkor a *DROP-LINE* típus alkalmazása ajánlott (Jánosa 2011, Sajtos – Mitev 2007).



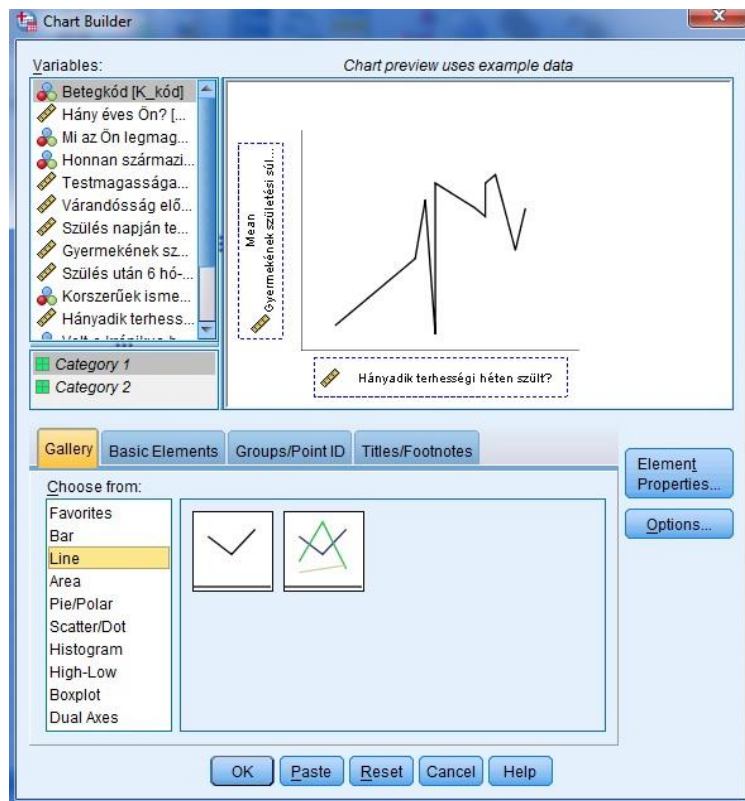
7/18. ábra: Vonaldiagram típus kiválasztási panel

A típus kiválasztását követően azt a szerkesztési panelt kapjuk, melyben az alábbi ábra alapján megadjuk a változókat valamint a *TITLES* gyorsgombbal megadjuk az ábra címét ezt a *CONTINUE* gombbal véglegesítjük, majd végül az *OK* gomb lenyomásával elkészül ábránk.



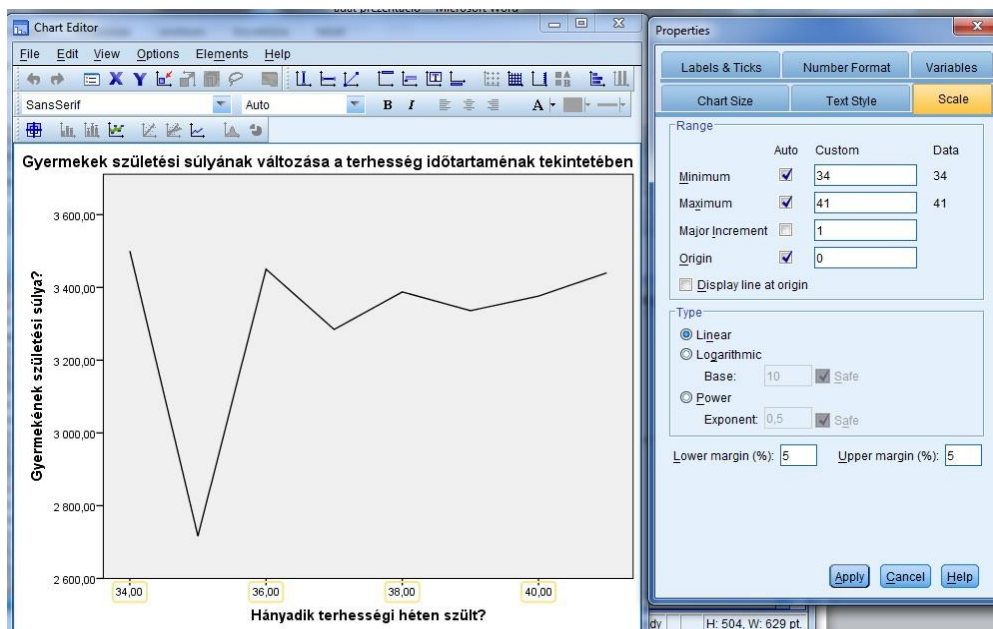
7/19. ábra: Vonaldiagram szerkesztési panel (*LEGACY DIALOGS* menüopcióval)

A másik lehetőség a *CHART BUILDER* eszközzel lehetséges. Először a *GRAPH* → *CHART BUILDER* elérési úton a *GALLERY*-ben kiválasztjuk a *LINE* típust, majd a következő ábrán látható beállítással, változók és tengelyek megadásával létrehozuk a kívánt vonaldiagramot.



7/20. ábra: Vonaldiagram szerkesztési panel a *CHART BUILDER* segítségével

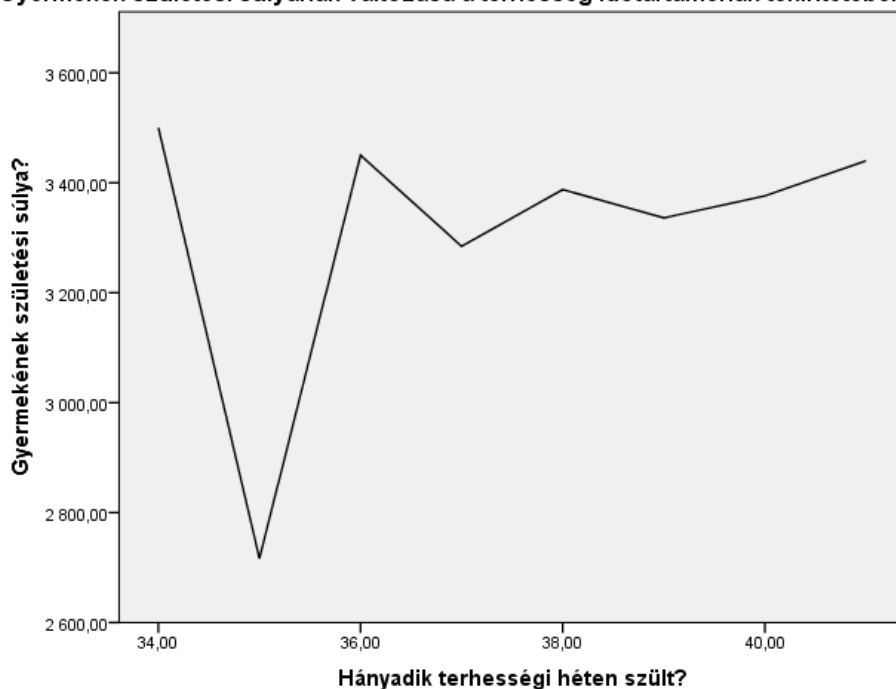
A *TITLES/FOOTNOTES* fülön adhatjuk meg a grafikon címét és amennyiben szükséges az esetleges lábjegyzetet is. A kiválasztott tartalomra rákattintva a *CONTENT* rubrikába beírhatjuk a közölni kívánt információt, majd az *APPLY* gombra kattintva elfogadhatjuk azt. Legvégül az *OK* gombra kattintva véglegesíthetjük a beállításokat és kapjuk meg ábránkat. Amennyiben az ábrán valami nem igényeinknek megfelelő pl.: a terhességi időtartam skálájának beosztása, akkor a módosítások végrehajtásához kétszer rá kell kattintanunk a diagramra az *OUTPUT VIEW* nézetben, ezt követően megjelenik a diagram a szerkesztő ablakban. Ezt követően az *X* tengelyre kattintva (terhességi időtartam) kettőt, megnyitja a program a *PROPERTIES* panelt. Ennél a panelnél a *SCALE* fület kiválasztva beállíthatjuk az időtartam minimum és maximum ábrázolását is, de példánknál maradva a *MAJOR INCREMENT* rubrikában igényeinknek megfelelően állíthatjuk be a skála léptékeit (Jánosa 2011).



7/21. ábra: Vonaldiagram ábrázolásának utólagos módosítása

Akarmelyik verziót is választjuk, végül megkapjuk az alábbi vonaldiagramot (skála változtatás nélkül).

Gyermekek születési súlyának változása a terhesség időtartamának tekintetében



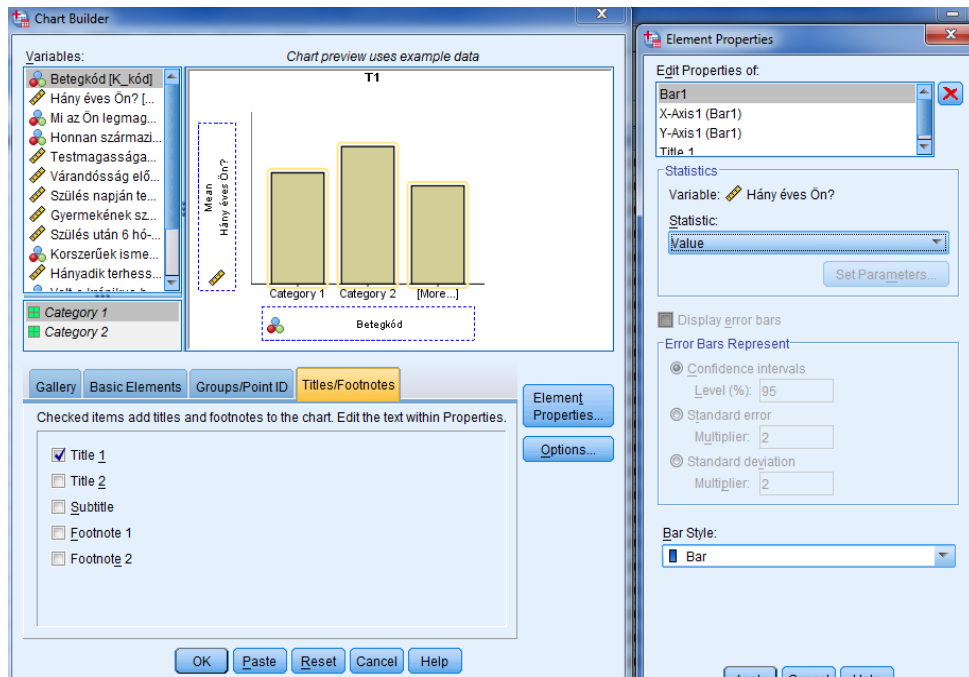
7/22. ábra: Gyermekek születési súly változásának átlaga a terhesség időtartamának tekintetében

„Oszlopdiaagram / Bar Chart”

Az oszlopdiaagram, kategóriákba tartozó gyakoriságok vizsgálatában, összehasonlításában és arányainak meghatározásában leginkább alkalmazott eszköz. Többek között a nagyságok összehasonlítását teszi lehetővé számunkra. Gyakran alkalmazott és közkedvelt grafikus ábrázolási módszer.

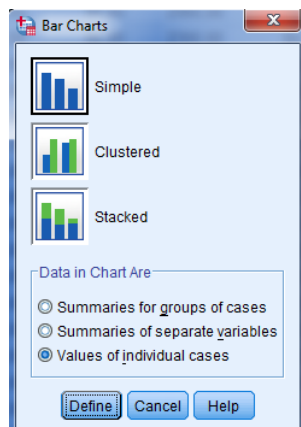
Példánkban az édesanyák életkorát kívánjuk szemléltetni és összevetni, ebben az esetben az x tengely változójául a betegkódot, míg az y tengely változójának az életkorokat választjuk.

Ezen diagram típus grafikus ábrázolásra is két lehetőségünk van az SPSS-ben. Az előbb ismertetett módon belépünk a *CHART BUILDER* menüpontba, itt a *GALLERY*-ben a *BAR* opciót választjuk illetve azon belül a számunkra leginkább megfelelő oszlopdiaagram típust. A változókat a már előbbieken bemutatott módon behúzzuk a megfelelő tengelyekre és a címet megadva, az *OK*-ra kattintva megkapjuk diagramunkat.



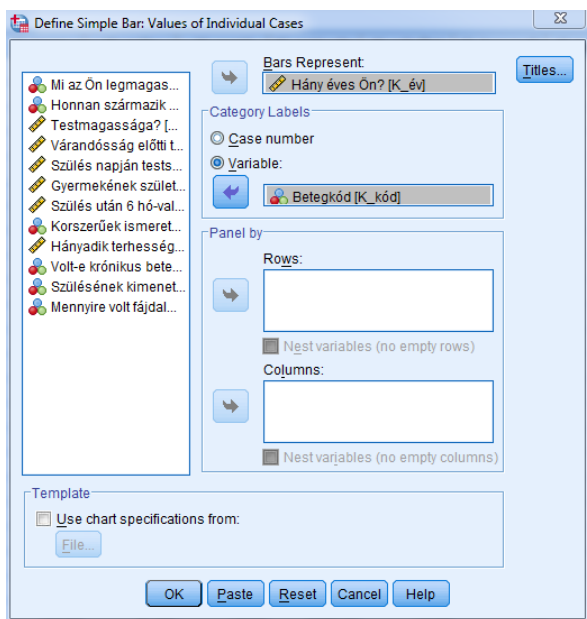
7/23. ábra: Oszlopdiaagram szerkesztési panel (*CHART BUILDER*)

Természetesen itt is lehetőségünk nyílik a *LEGACY DIALOGS* segítségével is az ábra szerkesztésre. Ebben az esetben a *GRAPHS* → *LEGACY DIALOGS* → *BAR* elérési utat követően, ismét a típus választási panelt kapjuk, melyen a *SIMPLE* típust választjuk, úgy hogy az alsó opció ismét a *VALUES OF INDIVIDUAL CASES* lesz, majd a *DEFINE* gombbal léphetünk tovább a szerkesztési panelre. Megjegyezendő, hogy a *SIMPLE* típust alkalmazhatjuk, egy szempont szerint csoportosítva vagy több adatsort alkalmazva, úgy, hogy a változókon belül csoportképző ismerv megadása nem lehetséges. A *CLUSTERED* típusnál egy adott adatsort ábrázolhatunk egy elsődleges és egy másodlagos szempont szerint kategorizálva. A *STACKED* típus esetén több adatsort ábrázolhatunk, melyek egy oszlopon belül egymásra halmozva kerülnek megjelenítésre. Az alsó opciók közül, melyek a megjeleníteni kívánt adatsoportokra vonatkoznak a *Summarise for group of cases* a csoportok szerinti, a *Summarise of seperate variables* a kiválasztott változók szerinti, míg a *Values of individual cases* pedig minden értéket külön megjelenítő ábrázolást jelenti (Jánosa 2011, Sajtos – Mitev 2007).



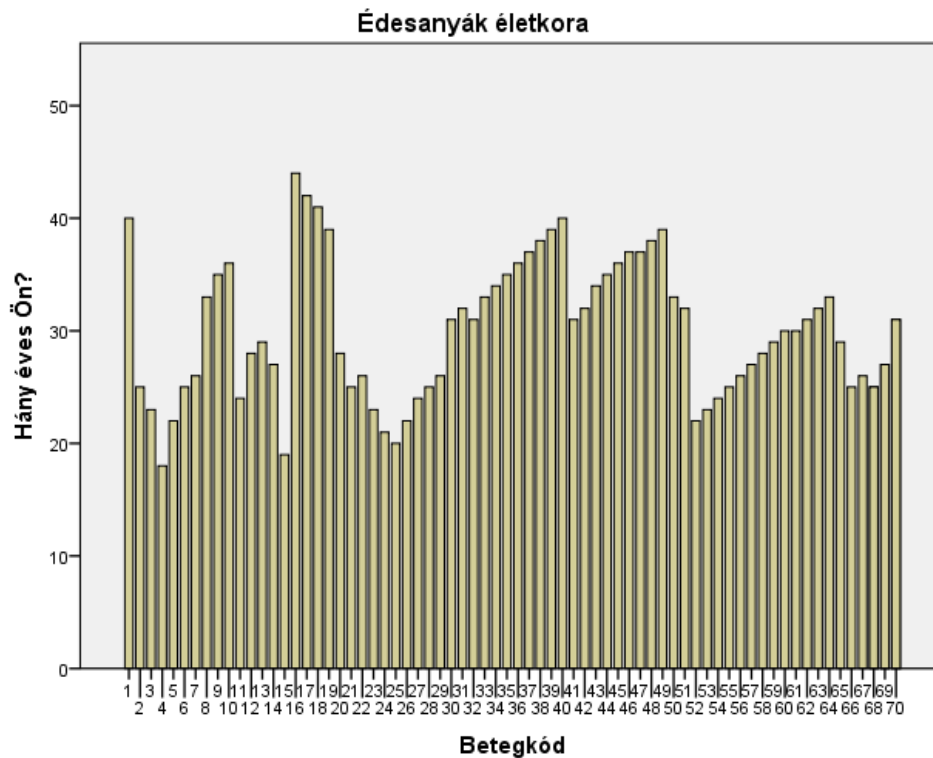
7/24. ábra: Oszlopdiagram típus kiválasztási panel

A szerkesztési panelen az előzőekben ismertettek értelmében a *BARS REPRESENT* sávba a „Hány éves Ön?” változó kerül, míg a *CATEGORY LABELS* menüpontban a *VARIABLE* opciót választva a „Betegkód” változót adjuk meg. A *TITLES* menüpontban megadhatjuk oszlopdiagramunk címét és az esetleges lábjegyzeteket, majd az *OK*-ra kattintva elkészül grafikonunk.



7/25. ábra: Oszlopdiagram szerkesztési panel (*LEGACY DIALOGS* menüopcióval)

Mindkét opció elvégzésekor ugyanazt az oszlopdiagramot kapjuk, mely az édesanyák életkorát mutatja, egymáshoz viszonyítva, betegkód alapján.

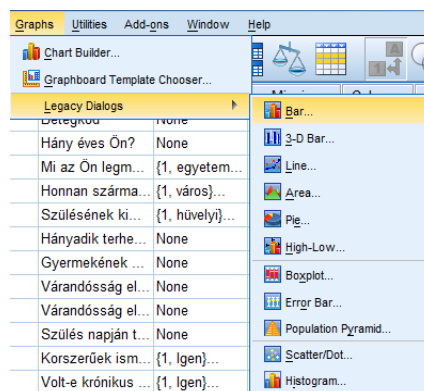


7/26. ábra: Édesanyák életkora betegkód alapján

Az oszlopdigramok során kiemelt helyen szerepel a már a statisztikai becslések során számított konfidencia intervallum (CI) megjelenítése.

Gyakorló feladatként, ábrázoljuk lakhely szerinti bontásban 95 %-os megbízhatóság mellett a kisgyermek várható átlagos születési súlyát.

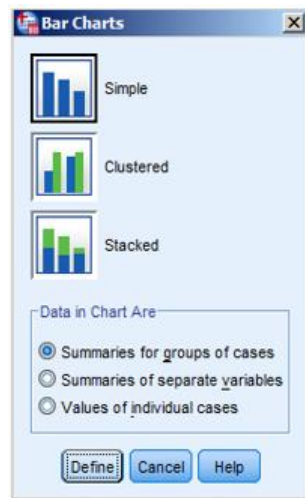
A *GRAPHS* menü *BAR* opciójából tudjuk a beállításokat elindítani.



7/27. ábra: Az oszlopdigram elérési helye

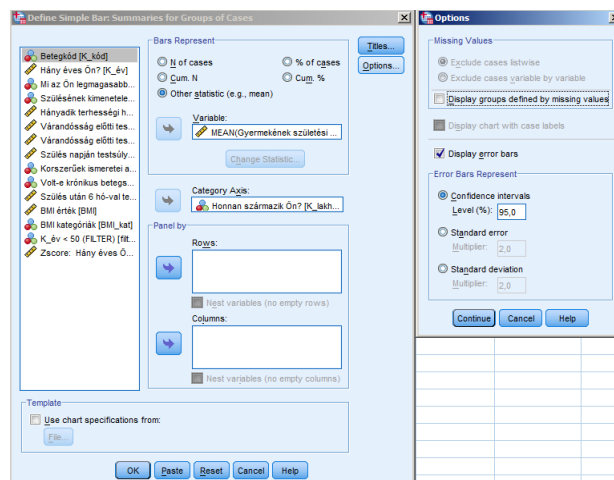
Ezt követően válasszuk az alapbeállításokat tartalmazó lehetőséget, a *SIMPLE* (egy változót esetlegesen csoportképző ismérveket megadva) és a *SUMMARIES FOR GROUP OF CASES*

(összesített adatszoportokat megadott csoportok szerint) beállításokat, melyet a *DEFINE* nevű gombbal tudunk választani.



7/28. ábra: Az oszlopdiaagram alaptípusának definiálása

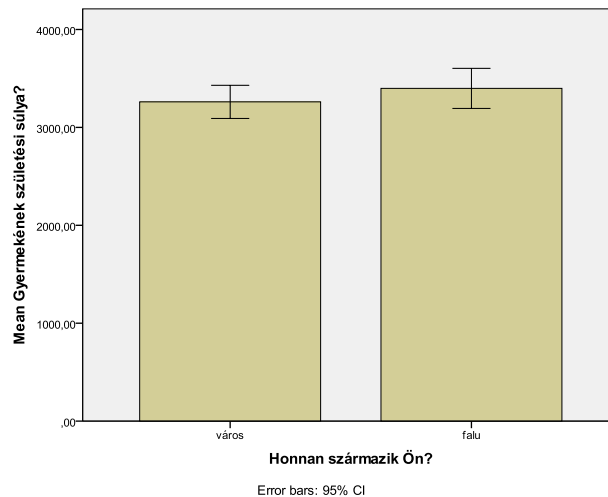
Az alaptípus meghatározását követően a vizsgált változónkat és annak mutatószámát állíthatjuk be. Jelen esetben a gyermekek születési súlyának átlagával kívánunk dolgozni. A csoportképző ismérvnél a származási hely változót kell jelölnünk.



7/29. ábra: Az oszlopdiaagram és a megbízhatósági tartomány meghatározása

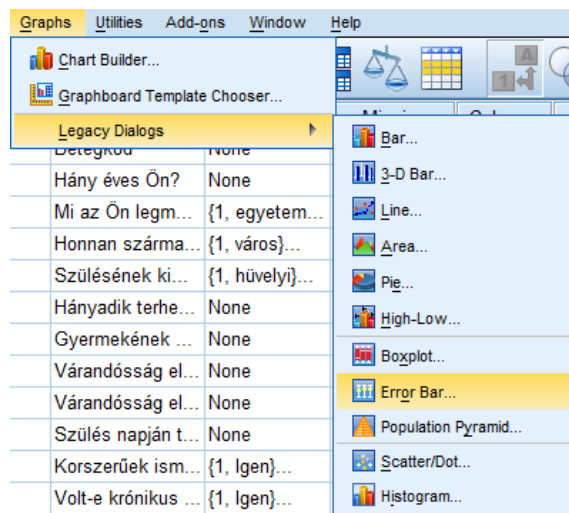
Ezt követően az *OPTIONS* modullal tudjuk az oszlopdiaagramra a megbízhatósági tartományt ráilleszteni, jelöljük a *DISPLAY ERROR BARS* négyzetet, majd a *CONFIDENCE INTERVALS LEVEL(%)*: 95-t lehetőséget.

Ezt követően létrejön az általunk kalibrált diagram.

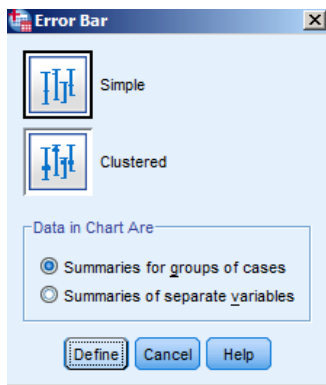


7/30. ábra: Az oszlopdiagram és a megbízhatósági tartományok ábrája (statisztikai becslés)

A megbízhatósági tartománynak a másik megjelenítési formája az *ERROR BAR*, ami beállítástól függően egy pontbecslés (átlagbecslés) megbízhatósági tartományát, standard hibáját vagy szórását ábrázolja.



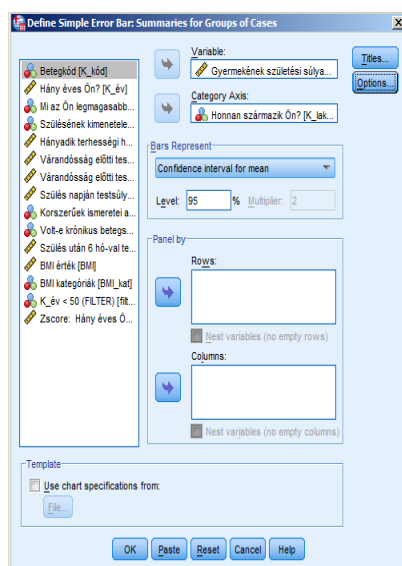
7/31. ábra: A megbízhatósági tartomány, és a hibahatárok megjelenítésnek elérési útja



7/32. ábra: A diagram alaptípusának definiálása

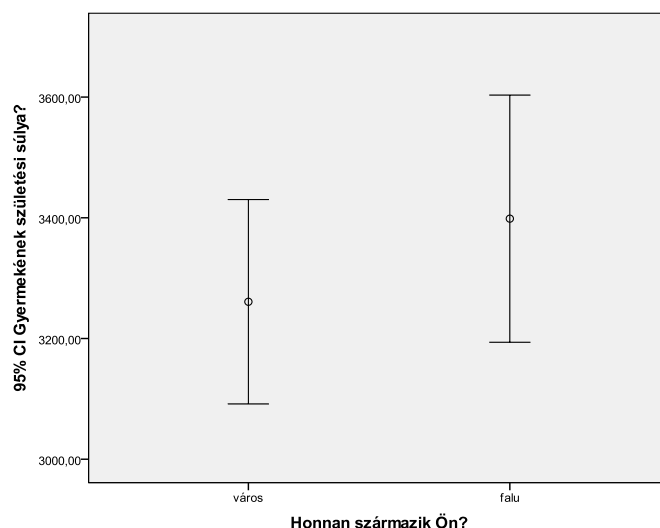
Ezt követően válasszuk az alapbeállításokat tartalmazó lehetőséget, a *SIMPLE* (egy változót esetlegesen csoportképző ismérveket megadva) és a *SUMMARIES FOR GROUP OF CASES* (összesített adatsoportokat megadott csoportok szerint) beállításokat, melyet a *DEFINE* nevű gombbal tudunk véglegesíteni.

Az alaptípus meghatározását követően a vizsgált változónkat és a csoportképző ismérvet (származási hely) állíthatjuk be. A *BARS REPRESENTS* beállításnál van lehetőségünk választani a megjelenítési formát, ahol jelen esetben az átlaghoz tartozó megbízhatósági tartományt választjuk.



7/33. ábra: A diagram beállításai

Ezt követően az *OK* gomb megnyomása után létrejövő statisztikai ábra a megbízhatósági tartományokat jeleníti meg a csoportképző ismérv vonatkozásában.



7/34. ábra: A megbízhatósági tartományok ábrája (ERROR BAR)

„Kördiagram / Pie Chart”

A kördiagram segítségével egy sokaság összetételét jellemezhetjük, illetve részekből összetevődő egész elemeinek valamint azok egymáshoz való viszonyának ábrázolására alkalmazható grafikus módszer. Alapvetően két típust különböztetünk meg, így beszélhetünk egy adatsoros és több adatsoros grafikonról is.

Egy-adatsoros grafikon

Amennyiben egy időegységet, területi vagy szervezeti egységet tekintünk egésznek és annak részeit valamint egymáshoz való viszonyát, arányát, egészből való részesedését ábrázolhatjuk, jellemezhetjük.

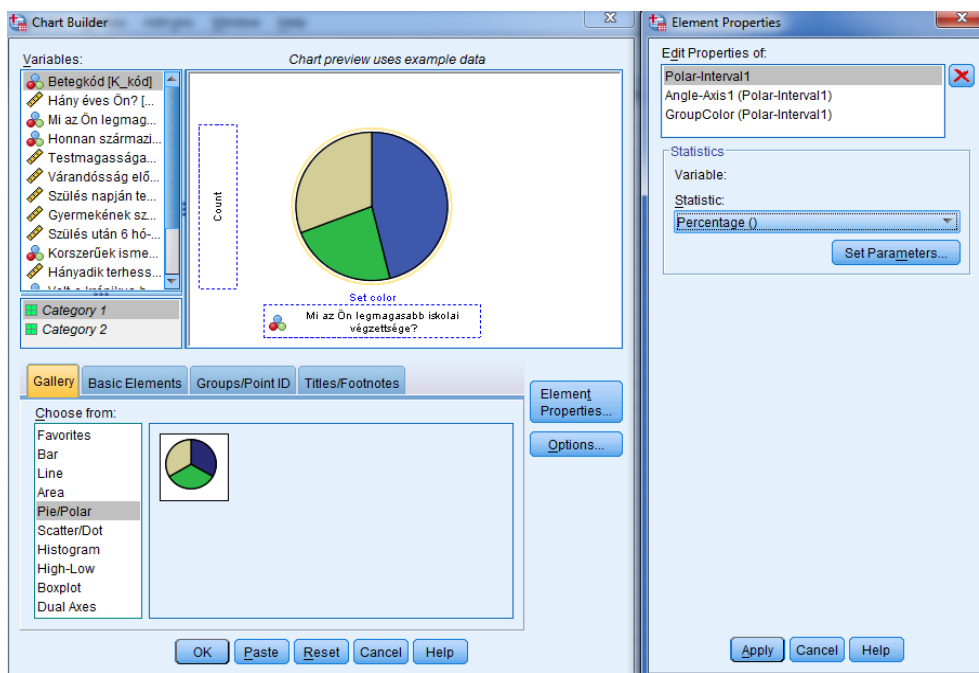
Több-adatsoros grafikon

A több-adatsoros grafikon esetén egyszerre több változást kell ábrázolni, jellemezni (változások szemléltetésére alkalmas típus). Egyrésztől változik a megfigyelt egész nagysága, terjedelme másrésztől a részek, összetevők mérete és egymáshoz viszonyított aránya is.

Az egy-adatsoros grafikon talán a leggyakrabban előforduló megoszlási diagram típus, ennek oka szemléltetése, érthetősége és esztétikai megjelenése egyaránt.

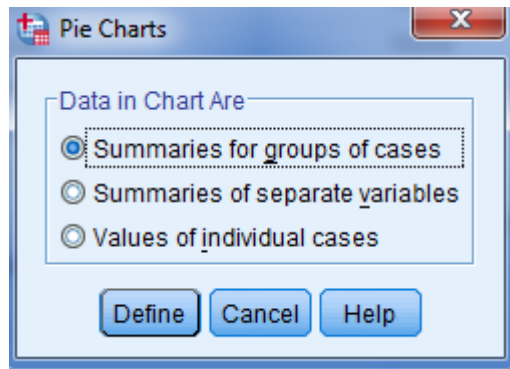
Példánkban az egy-adatsoros grafikon, kördiagram szerkesztését illetve annak metodikáját mutatjuk be. Eddig használt adatbázisunkból most az édesanyák legmagasabb iskolai végzettségének megoszlást mutatjuk be, az összes felmért személyhez, mint egészhez-sokasághoz viszonyítva, más változót most nem vonunk be vizsgálatunkba.

A *CHART BUILDER* segítségével az előzőekben ismertetett módon belépünk a menüpontba, itt a *GALLERY*-ben a *PIE/POLAR* opciót választjuk. A változót („Az Ön legmagasabb iskolai végzettsége”) a már előbbiekben bemutatott módon behúzzuk a megfelelő tengelyre, ugyanakkor ennél a diagramnál csak ezt a változót használjuk, hiszen a viszonyítási alap maga az egész sokaság. A *TITLES / FOOTNOTES* menüpontban megadjuk a grafikon címét, majd az *ELEMENT PROPERTIES* menüpontban, a *STATISTIC* opciónál a legördülő listából a *PERCENTAGE (%)*, azaz százalékos megoszlást választjuk. Majd az *APPLY* gombbal véglegesítjük beállításainkat és az *OK* gombbal létrehozuk ábránkat. (Jánosa 2011, Sajtos – Mitev 2007).



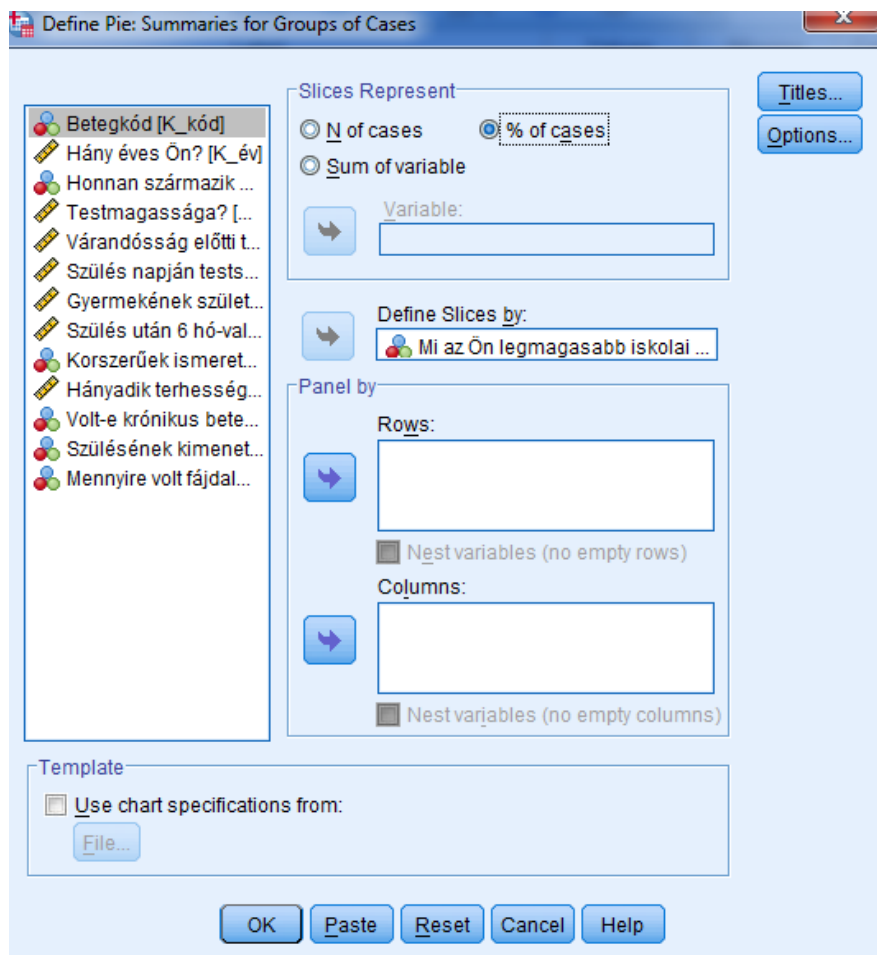
7/35. ábra: Kördiagram szerkesztési panel (*CHART BUILDER*)

A másik szerkesztési mód a már jól ismert *LEGACY DIALOGS* menüponttal lehetséges, itt a *GRAPHS* → *LEGACY DIALOGS* → *PIE* elérési utat követően, ismét a típus választási panelt kapjuk, melyen egy változó lévén a *SUMMARIES FOR GROUPS OF CASES* opciót választjuk, majd a *DEFINE* gomb lenyomásával a szerkesztő felületre lépünk tovább. Ez esetben a *SUMMARIES FOR GROUPS OF CASES* egy változó értékein belüli kategorizálást jelent, melyben az adatok összesítése után ezek megjelenítése történik. A *SUMMARISE OF SEPERATE VARIABLES* esetén a kiválasztott változók összesítését követően azok megoszlásának ábrázolását láthatjuk. A *VALUES OF INDIVIDUAL CASES* esetén minden értéket külön jelenít meg ábránk.



7/36. ábra: Kördiagram típus kiválasztási panel

A szerkesztő felületen annyi a teendő, hogy mivel százalékos megoszlást kívánunk látni ezért, a *% OF CASES* opciót választjuk, majd a *DEFINE SLICES BY* mezőbe helyezzük a vizsgálni kívánt változót, ezt követően az *OK* gombra kattintva elkészül grafikonunk.



7/38. ábra: Kördiagram szerkesztési panel (*LEGACY DIALOGS* menüopcióval)

Bármelyik verziót is választjuk, végeredményként az alábbi kördiagramot kapjuk eredményül.



7/39. ábra: Édesanyák iskolai végzettségének megoszlása

Hisztogram / „Histogram”

A hisztogram hasznos analitikai eszköz, mely tájékoztatást nyújt a sokaság eloszlásáról és gyakorlatilag a változók gyakorisági eloszlásának ábrázolására szolgáló grafikus ábrázolási módszer. A hisztogram osztályközei egy-egy értéktartományt képviselnek, olyan módon hogy az egyes oszlopok magassága jellemzi az előbb említett tényezők (osztályközök) gyakoriságát.

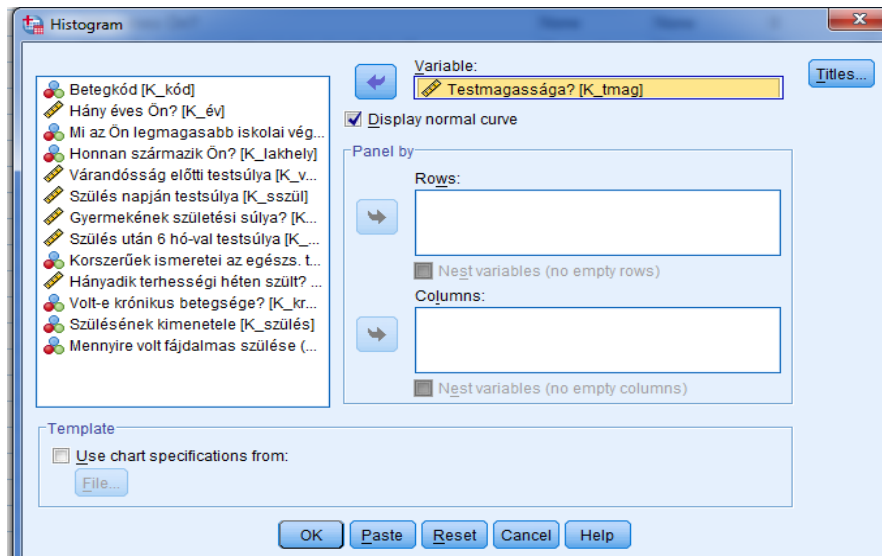
A kész ábráról először szembeütő tulajdonság az eloszlás szimmetriája, ez ránézésre látható és levonható a következtetés, hogy mely nagyságrendek fordulnak elő karakteresen és melyek kevésbé.

A hisztogram koordináta rendszerben történő elhelyezésének tekintetében, az x tengelyen az értékosztályok, míg az y tengelyen az egyes osztályokhoz tartozó gyakoriságok láthatóak. Amennyiben az ábrázolt oszlopok végpontjait egyenes vonallakkal összekötjük, tájékoztatást kaphatunk a teljes sokaság eloszlásáról.

Példánkban ezúttal a várandós kismamák körében vagyunk kíváncsiak a testmagasságok megoszlására.

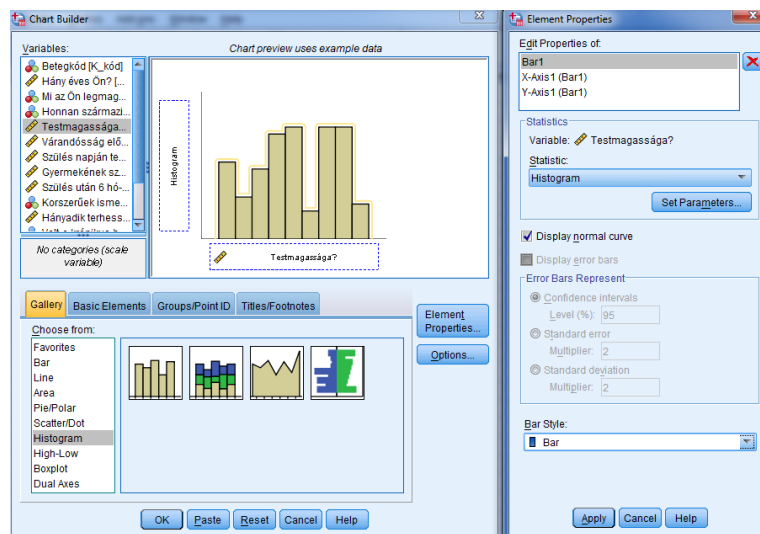
Ezt a diagram formát is kétféle módon érhetjük el, egyrészt a *GRAPH* menüpontban találjuk és a *LEGACY DIALOGS* legördülő listájából a *HISTOGRAM* opciót választjuk. Ekkor legelőször a szerkesztő felületen ki kell választanunk a szemléltetni kívánt – arány skálán mérhető változót, majd ezt át kell helyeznünk a *VARIABLE* mezőbe. Ha a *DISPLAY NORMAL CURVE* rubrikát beikszeljük, abban az esetben megjelenítésre kerül a már előzőekben is említett eloszlási görbe. A *TITLES* parancsgombra kattintva címet, lábjegyzetet

adhatunk meg. Végezetül az *OK* gombra kattintva véglegesíthetjük és kapjuk meg grafikonunkat (Jánosa 2011, Sajtos – Mitev 2007).



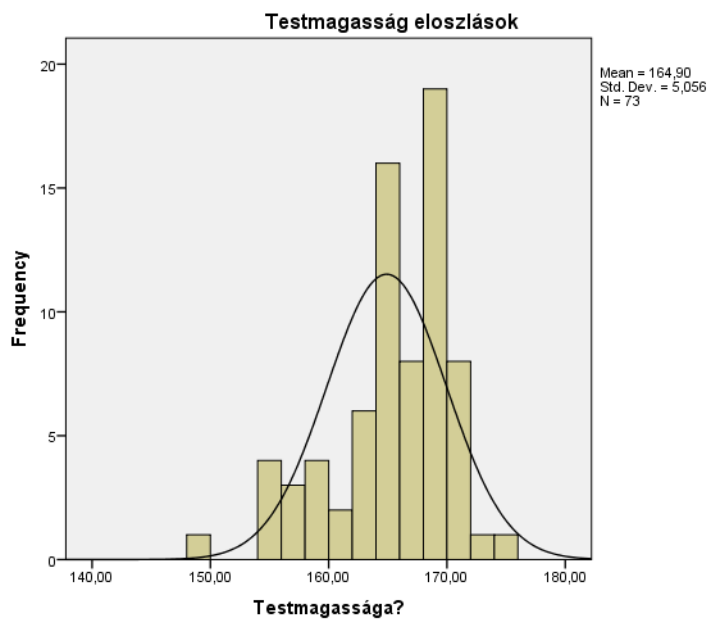
7/40. ábra: Hisztogram szerkesztési panel (*LEGACY DIALOGS* menüopcióval)

Másik lehetőségünk a *CHART BUILDING* verzió. Ebben az esetben belépünk a *CHART BUILDER* menüpontba, itt a *GALLERY*-ben a *HISTOGRAM* opciót választjuk illetve azon belül a számunkra leginkább megfelelő hisztogram típust. Az x tengelyre helyezzük, a vizsgálni kívánt változót (testmagasságok), majd amennyiben a *DISPLAY NORMAL CURVE* rubrikát beikszeljük, abban az esetben megjelenítésre kerül a már előzőekben is említett eloszlási görbe. A szerkesztés befejezéseként az *APPLY* gorsgombbal elfogadjuk a beállított tulajdonságokat majd az *OK* gombra kattintva megjelenik ábránk.



7/41. ábra: Hisztogram szerkesztési panel (*CHART BUILDER*)

A választott módszertől függetlenül az alábbi ábrát kapjuk a várandós édesanyák testmagasságának tekintetében.



7/42.ábra: Várandós anyák testmagasság megoszlása

8. ÖSSZEFÜGGÉS-VIZSGÁLATOK (Ács Pongrác)

8.1. Statisztikai kapcsolatok összefüggéseinek elméleti alapjai

A bennünket környező világ megismerése során a kapcsolatok megismerése igen fontos, és az így nyert információk fontos szerephez jutnak egy-egy döntés előkészítése, illetve a döntések hatásának mérése kapcsán. Nyitott szemmel járva jobban megérthetjük a vizsgált jelenségeket, ha nem kizárólag önmagukban nézzük őket, hanem egymáshoz való viszonyukat, más jelenségekkel való összefüggéseiket is vizsgáljuk.

Összefüggés-vizsgálatok jellemzői:

- Feltáró jellegű kutatásoknál több változó esetén a vizsgált jellemző tulajdonságok között kapcsolatok, összefüggések feltárása a cél. Az összefüggés vizsgálatok legfőbb segítői a jó hipotézisek.
- Az összefüggéseket legalább két azonos (mért, rangsorolt, nominális) adatfajtába tartozó változó között számolunk. Ha az adatfajták különbözőek, akkor különbözőségvizsgálatokat végzünk.

A jelenségeket, folyamatokat leginkább a következőképpen tipizálhatjuk:

- az ismérvek egymástól **függetlenek**, ha az egyik ismérvből semmilyen következtetés nem vonható le a másik ismérvre nézve,
- **sztochasztikus kapcsolatnál** a statisztikai ismérvek között tendenciaszerűen, valószínűségi jelleggel érvényesülő kapcsolatot értünk. Számunkra ez lehet a legérdekesebb eset, hiszen az azt jelenti, hogy az ismérvek körében nagy eséllyel van valamiféle kapcsolat, „együtt alakulás”,
- **függvényszerű, determinisztikus** kapcsolatnál az egyik ismerv meghatározza, hogy a másik ismerv milyen értékeket vehet fel.

A kapcsolatokat megkülönböztetjük a bennük szereplő ismérvfajták alapján, ezek szerint három típust különböztetünk meg:

- Mindegyik (mindkét) ismerv minőségi (kategóriás ismerv, hiszen az ismérvértékek kategóriákat képeznek), azaz nominális típusú, akkor **asszociációs kapcsolatról** beszélünk. A gyakorlati adatelemzés során a módszereket a **keresztábra** elemzésnek is nevezik, illetve a leggyakrabban használt statisztikai mutatószám alapján **khi- négyzet** próbának.
- **Vegyes kapcsolatnál**, az ok szerepét minőségi, de az okozat szerepét mennyiségi ismerv tölti be.

- Ha mindegyik (mindkét) ismérv mennyiségi, akkor **korrelációs kapcsolatról** beszélünk.

Mindhárom típusú kapcsolatot valamilyen mutatószám segítségével célszerű számszerűsíteni. Fontos szerepet töltenek be az ún. kapcsolat intenzitását kifejező mérőszámok, amelynek általános sémáját az alábbiakkal írhatjuk le (az intenzitást általánosságban mérő mutatószám jele legyen T):

$$0 \leq T \leq 1$$

Általánosságban a T mutató abszolút értékére fogalmazzuk meg a fenti intervallumot, de bizonyos esetben - főleg a korrelációs kapcsolatokban - az előjel is fontos információ hordozója, ugyanis a kapcsolat pozitív vagy negatív irányát mutatja. Természetesen ilyen esetben a mutatószám intervalluma: $[-1;1]$.

A mutatószámok értelmezése mindig függ az adott problémától; ismerni kell a vizsgált összefüggés jellegét, természetét. Az általános elemzéshez ad segítséget az alábbi séma:

8/1. táblázat Kapcsolat-szorossági mutatók általános értelmezése

T = 0	nincs kapcsolat
$0 < T < 0,3$	gyenge a kapcsolat
$0,3 \leq T \leq 0,7$	közepes szorosságú a kapcsolat
$0,7 < T < 1$	erős a kapcsolat
T = 1	függvényszerű vagy determinisztikus

Forrás: saját szerkesztés

8.2. Asszociációs kapcsolatok vizsgálata, keresztábra elemzések

A kapcsolat alapfeltétele, hogy az ismérvek mindegyike minőségi legyen. Az asszociációs kapcsolat esetében az adatokat egy kombinációs táblába rendezzük, ha a kombinációs táblában gyakoriságok szerepelnek, akkor ún. kontingencia tábláról beszélünk.

A leggyakrabban használt mutatószámokat az alapján tudjuk csoportosítani, hogy szimmetrikusnak vagy nem szimmetrikusnak tekinthetjük-e őket. A szimmetrikusság jelenti, hogy a változók között nem tudjuk, hogy melyik az ok és melyik az okozat, vagyis nincsen a kapcsolatnak iránya.

8/2. táblázat: A nominális skálán használható mutatószámok összegzése

	Szimmetrikus	Asszimmetrikus
Kizárólag 2*2 tábla esetén	Φ (phi) együttható	
Bármely tábla esetén (2*2-esetén is)	Kontingencia együttható, Cramer együttható	Lambda, Goodman és Kruskal tau (bizonytalansági együttható)

Forrás: saját szerkesztés

Amennyiben a vizsgált változóink dichotóm változók, vagy a kérdőívekben a zárt kérdések magválaszolása során csak kettő (alternatív ismerv), egymást kizáró ismervváltozat van megadva pl. nő-férfi, igen-nem, van-nincs stb., akkor a kétdimenziós kontingencia tábla általános alakja a következő lesz:

8/3. táblázat: A kétdimenziós kontingencia tábla általános alakja

A ismerv Változatai	B ismerv változatai		Összesen:
	B ₁	B ₂	
A ₁	f ₁₁	f ₁₂	S ₁
A ₂	f ₂₁	f ₂₂	S ₂
Összesen:	O ₁	O ₂	n

Forrás: saját szerkesztés

n – az összes elemszám,

f₁₁ – az A ismerv első és a B ismerv első változatához rendelt gyakoriság (hasonlóan értelmezhetők a többi cella gyakoriságai is!),

S₁ – az első sor (az A ismerv első változatához tartozó) gyakoriságok összege,

O₁ – az első oszlop (a B ismerv első változatához tartozó) gyakoriságok összege.

Belátható az alábbi összefüggés:

$$S_1 + S_2 = O_1 + O_2 = n$$

A sorok, illetve az oszlopok összegeit **peremgyakoriságoknak** nevezzük.

Alternatív ismérvek esetén a kapcsolat mérésére alkalmazhatjuk az ún. **Yule-féle mutatót**, ami a táblában szereplő gyakoriságok "kereszt-szorzataiból" állítható elő:

$$Y = \frac{f_{11} \times f_{22} - f_{12} \times f_{21}}{f_{11} \times f_{22} + f_{12} \times f_{21}}$$

A mutatószám – mivel két adat különbségének és ugyanazon adatok összegének hányadosa – minden esetben -1 és +1 közötti értéket vesz fel.

A következő példa a meglévő adatbázisunkból származik.

A példa adatait a következő két kérdés válaszai szolgáltatták:

KORSZERŰNEK TEKINTHETŐK-E AZ ISMERETEI AZ EGÉSZSÉGES TÁPLÁLKOZÁSSAL KAPCSOLATBAN? (tegyen „x”-et a megfelelő kockába!)

Igen

Nem

VOLT-E A KRÓNIKUS BETEGSÉGE? (tegyen „x”-et a megfelelő kockába!)

Igen

Nem

A válaszok eredményeit a kontingencia táblából olvashatjuk le:

8/4. táblázat

		Volt-e krónikus betegség?		Összesen
		Igen	nem	
Korszerűek-e az ismeretei az egészséges táplálkozással kapcsolatban	Igen	11	33	44
	Nem	5	21	26
	Összesen	16	54	70

Forrás: saját szerkesztés

A képletbe behelyettesítve a következő eredményt kapjuk:

$$a = \frac{11 \times 21 - 33 \times 5}{11 \times 21 + 33 \times 5} = 0,16$$

Az eredmény abszolút értéke 0 és 0,3 közé esik, tehát erősségét tekintve gyenge erősségű kapcsolatot találtunk a két ismérv között. A mutatószám alkalmazása során azonban fokozottan figyelni kell arra, hogy valamennyi átlóban lévő elem különbözzön nullától. Ha csak egy esetben nulla a gyakoriság, a mutatószám akkor is determinisztikus kapcsolatot jelez, ha az egyébként nem is áll fenn.

Kettőnél több ismérvváltozat esetén más mérőszámot kell alkalmazni. A **Cramer-együttható** feloldja az alternatív ismérvek dilemmáját és ugyanakkor érzéketlen a kirívó (egyik cellában nulla értékkel bír) esetekkel szemben. Több kutató által a legkedveltebb és legmegbízhatóbb

khi-négyzet alapú mutató, mely szinte minden kereszttábla esetén alkalmazható. Az alap gondolat azt vizsgálja, hogy miként alakulnának az egyes gyakoriságok, ha az ismérvek között nem lenne kapcsolat, tehát függetlenek lennének, vagyis valamely ismérv érték nem vonzaná a másik ismérv valamely adott értékét. A kiinduláshoz itt is a kontingencia táblát használjuk.

A számítás alap gondolata: amennyiben a független viszonyt feltételező gyakoriságok és a tényleges gyakoriságok között eltéréseket találunk, akkor a sztochasztikus kapcsolat meglétére gondolhatunk. A feltételezett gyakoriságok számítása tehát azt jelenti, hogy a sokaságot a peremeloszlások alapján osztjuk szét. Ha a táblát a feltételezett gyakoriságokkal töltjük ki, minden sor megoszlása ugyanolyan lesz, ami megfelel a két ismérv függetlenségének. A 2x2-es kontingencia táblát használva, a peremgyakoriságok segítségével meghatározhatjuk a függetlenség esetén feltételezett gyakoriságokat, amelyeket *-gal különböztetünk meg:

$$\begin{aligned} \frac{S_1 \times O_1}{n} &= f_{11}^* & \frac{S_1 \times O_2}{n} &= f_{12}^* \\ \frac{S_2 \times O_1}{n} &= f_{21}^* & \frac{S_2 \times O_2}{n} &= f_{22}^* \end{aligned}$$

Elsőként minden cellában kiszámítjuk az alábbi relatív különbséget:

$$\frac{(f_{ij} - f_{ij}^*)^2}{f_{ij}^*}$$

ahol: az f_{ij} az i-edik sorának és j-edik oszlopának gyakorisága.

A sztochasztikus kapcsolat létezését jelzi tehát az, ha a ténylegesen megfigyelt és a függetlenség esetére feltételezett gyakoriságok nem egyeznek meg. A kétféle gyakoriság eltérése közötti különbségeket egy mérőszámban kell kifejezni, amit a négyzetes kontingencia mutatójával az ún. χ^2 -értékkel végezhetünk el.

$$\chi^2 = \sum \sum \frac{(f_{ij} - f_{ij}^*)^2}{f_{ij}^*}$$

A χ^2 önmagában még nem felel meg a sztochasztikus kapcsolat mérőszámaival szemben megfogalmazott feltételnek. Alsó határa ugyan 0, de felső határa jelentősen meghaladhatja az 1-et. Ezt a dilemmát oldja fel a Cramer-féle mutatószám (V), amelynek képlete:

$$V = \sqrt{\frac{\chi^2}{n \times (s - 1)}}$$

Ahol a tört nevezőjében az s a változók ismérvváltozatainak minimumát (a kevesebb ismérvváltozat számát) jelöli.

A következő példa során arra voltunk kíváncsiak, hogy van-e kapcsolat a szülés módja (típusa) és a BMI kategóriák között.

8/5. táblázat: A származási hely és a BMI kategóriák kombinációs táblázata

		BMI kategóriák				Összesen
		<i>sovány</i>	<i>normál testalkatú</i>	<i>túlsúlyos</i>	<i>erősen túlsúlyos</i>	
Szülésének kimenetele	hüvelyi	2	13	7	8	30
	császármetszés	3	10	14	13	40
	Összesen	5	23	21	21	70

Forrás: saját szerkesztés

A peremgyakoriságok segítségével a függetlenség esetén feltételezett gyakoriságok:

$$\frac{30 \times 5}{70} = 2,14 \quad \frac{30 \times 23}{70} = 9,86 \text{ stb. stb.}$$

A függetlenség esetén feltételezett gyakoriságokat a következő táblázatban láthatjuk:

8/6. táblázat: A függetlenség esetén feltételezett gyakoriságok

		BMI kategóriák				Össze- sen
		<i>sovány</i>	<i>normál testalkatú</i>	<i>túlsúlyos</i>	<i>erősen túlsúlyos</i>	
Szülésének kimenetele	hüvelyi	2,143	9,857	9,000	9,000	30
	császármetszés	2,857	13,143	12,000	12,000	40
	Összesen	5	23	21	21	70

Forrás: saját szerkesztés

Elsőként minden cellában kiszámítjuk az alábbi relatív különbséget:

$$\frac{(2 - 2,143)^2}{2,143} = 0,010 \text{ stb.}$$

Az így létrejött újabb munkatáblát láthatjuk:

8/7. táblázat: A χ^2 -érték számítása

		BMI kategóriák				Összesen
		<i>sovány</i>	<i>normál testalkatú</i>	<i>túlsúlyos</i>	<i>erősen túlsúlyos</i>	
Szülésének kimenetele	hüvelyi	0,010	1,002	0,444	0,111	30
	császármetszés	0,007	0,752	0,333	0,083	40
	Összesen	5	23	21	21	70

Forrás: saját szerkesztés

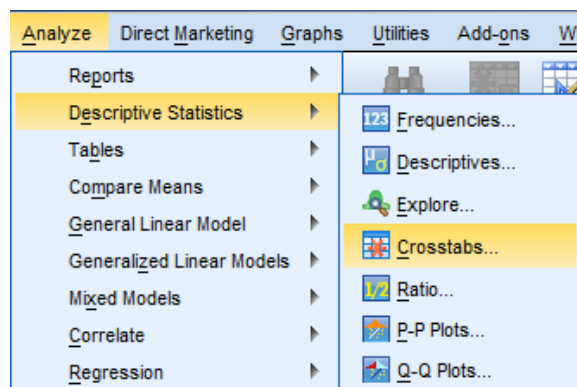
A Cramer-féle (V) mutatószám példánkban:

$$V = \sqrt{\frac{2,743}{70 \times (2 - 1)}} = 0,188$$

a szülés módja (típusa) és a BMI kategóriák között.

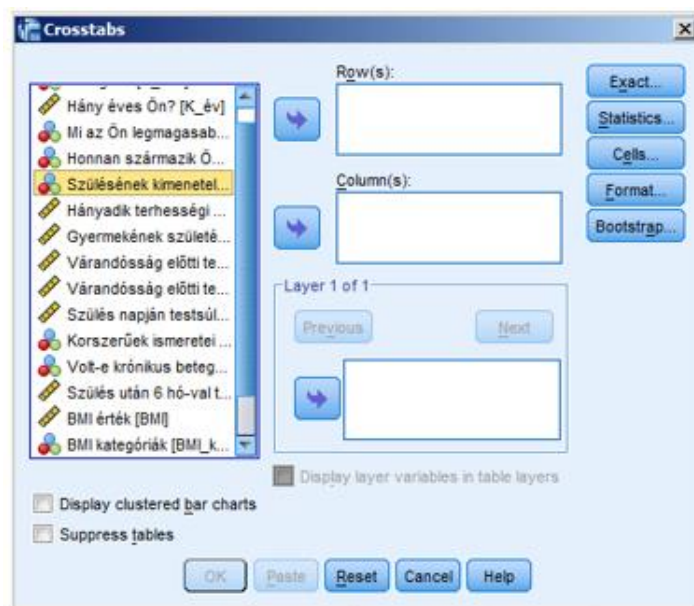
A mérőszám szerint a BMI kategóriák típusa és a szülés módja (típusa) közötti sztochasztikus kapcsolat gyenge erősségűnek mondható. A V^2 mérőszám is értelmezhető, amely azt mutatja meg, hogy – esetünkben – a BMI kategóriák típusa mintegy 3,53 %-ban determinálja a szülés módját, fajtáját.

Az SPSS program segítségével az asszociációs kapcsolat elemzése a keresztábra (CROSSTABS) modul segítségével készíthető, mely az ANALYZE főmenü DESCRIPTIVE STATISTICS almenü részéből érhető el.



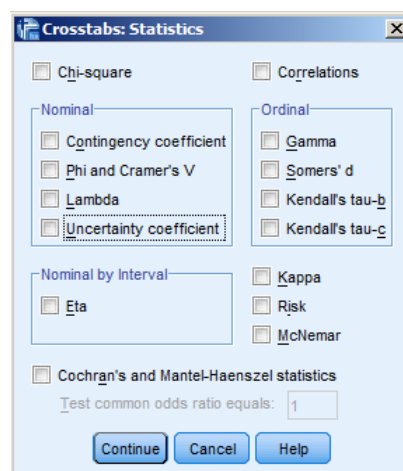
8/1. ábra: A keresztábra készítésének elérési útvonala

Ezt követően a változók beállításainak ablaka következik. Első lépésben a vizsgált változók kijelölése következik, át kell mozgatni a sor (*ROW(S)*) és oszlop (*COLUMN(S)*) ablakba. Arra vonatkozóan kötelező szabályok nem léteznek, hogy melyik változónk legyen a sor és melyik az oszlop, ez a mindenkori kutató szakmai kompetenciájára van bízva. Segítségként és javaslatként megemlíthető, hogy a gyakorlati társadalomtudományokban az oszlopváltozó (*COLUMN(S)*) a függő (amelynek eloszlására kíváncsiak vagyunk), míg sorváltozó (*ROW(S)*) a független (úgy véljük, hogy meghatározó szerepet játszik a függő változóra változó).



8/2. ábra: Az asszociációs kapcsolat (keresztábra) változóinak beállításai

Ezt követően az oldalmenük közül a *STATISTICS* modul mutatjuk be, hiszen az egyik legfontosabb beállításokat tartalmazza, innen érhetőek el a keresztábra elemzések során alkalmazható statisztikai mutatószámok.



8/3. ábra: A kapcsolat-szorossági mutatószámok

A beállítási ablakban jól látszik, hogy a statisztikai mutatószámok a változók skálátípusai alapján külön csoportosításban jelennek meg. Nézzük, most a **nominális adatok** esetén alkalmazott mutatószámokat. A mutatószámok közül a kontingencia, (*CONTINGENCY COEFFICIENT*), a phi (*PHI*) és a Cramer (*CRAMER's V*) mutatószámokat alkalmazzuk szimmetrikus kapcsolatok esetén, míg a lambda (*LAMBDA*), Goodman és Krukal tau, valamint a bizonytalansági együttható (*UNCERTAINTY COEFFICIENT*) aszimmetrikus kapcsolatok esetén választandók. Az asszociációs kapcsolatok során, kiemelten a szimmetrikusok esetében a mutatószámok a khi-négyzet értékhez, illetve annak valamely módosulásából származnak. A baloldalon a legfelső négyzet jelölése (pipa) nyomán a khi-négyzet (*CHI-SQUARE*) érték számítása jelenik meg.

- **Kontingencia-együttható** (*CONTINGENCY COEFFICIENT*): a mutatószám alkalmazható bármekkora változószámú keresztábra esetén, akár a speciális 2X2-es esetében is, azonban értelmezési nehézségek miatt ritkán használjuk, helyette a Cramer együttható (*CRAMER V*) ajánlott. A számításnál a mintanagyságot használja

$$\text{fel. } C = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

- **Phi (PHI) mutatószám (Φ)**: a keresztábrák speciális esetében a 2X2-es táblánál használjuk, mivel itt könnyen értelmezhető, hiszen ilyenkor a felső korlát maximális értéke 1. A számítás során a khi négyzet értékét a mintanagysággal korrigáljuk. Nem javasolt több ismérvváltozat során alkalmazni, mivel ilyenkor nincsen felső korlátja,

így értelmezése nem egyszerű. $\phi = \sqrt{\frac{\chi^2}{N}}$. Amennyiben azt a mutatószámot

használjuk, akkor a folytonossági Yates korrekciós együttható (*CONTINUITY CORRECTION*) nem alkalmazható, ami a khi- négyzet mutatószám 2X2-es táblánál alkalmazott korrekciója.

- **Cramer V együttható** (*CRAMER's V*): a leggyakrabban alkalmazott mérőszám, könnyen értelmezhető. Leggyakrabban a kettőnél több ismérvváltozattal rendelkező

változók esetén alkalmazandó. $V = \sqrt{\frac{\chi^2}{n \times (s-1)}}$. Fontos megemlíteni, hogy vannak

statisztikusok, akik azt tartják, hogy a khi- négyzet alapú mutatószámok - így a Cramer együttható is - nem alkalmazható abban az esetben, ha az ismérvváltozók közül (cellák), több mint 20%-ban az érték 5 alatti.

- **Lambda (LAMBDA):** aszimmetrikus mutatószám. A mutató értelmezése során megkapjuk azt a százalékos értéket, amely megmutatja, hogy a független változó mekkora mértékben képes a függő változót előre jelezni. A számított érték megmutatja, hogy hány százalékkal csökken az előrejelzési hiba, amennyiben az előrejelzés során a feltételezhető okot képező független változót is bevonjuk.

$$\lambda = \frac{SUM(f_i - f_d)}{N - f_d}.$$

- A Lambdához hasonlóan értelmezzük a Goodman és Kruskal tau, valamint a bizonytalansági mutatókat. A maximális értéke 1 azt jelenti, független változó ismérvértékeinek ismeretében 100 %-ban (hibátlanul) becsülhető a függő változó értéke.

A modul jobb oldalában található négyzetben az **ordinális skálán** mért változók esetében használt leggyakoribb mutatószámok kerültek összegyűjtésre.

8/8. táblázat: Ordinalis skálákon alkalmazott kapcsolat-szorossági mutatószámok

Mutatószám (ordinális skálák esetén)	Táblatípusok
<i>Gamma</i>	Bármely táblatípus és méret esetén (könnyen értelmezhető)
<i>Sommers'd</i>	Bármely táblatípus és méret esetén (értelmezése nehézkes)
<i>Kendall tau-b</i>	Szimmetrikus táblák esetén
<i>Kendall tau-c</i>	Nem szimmetrikus táblák esetén

Forrás: saját szerkesztés

A változók sorrendjei között keressük az összefüggéseket, hiszen itt az egyes ismérvkategóriák sorrendje jelentőséggel bír, így itt a kapcsolatok szorosságán kívül az irány is jelentőséggel bír. Pozitív az előjel abban az esetben, ha az egyik változó növekvő értéke növekvő változást idéz elő a másik változóban is. Amennyiben az egyik változó növekedése a másik változóban csökkenést vált ki, negatív kapcsolatról beszélünk. Összességében elmondható, hogy az ordinalis skálán mért változók esetében az asszociációs mérőszámok, esetpárok összehasonlítását célozza. Abban az esetben, amikor egy esetpárban az egyik eset minden változója magasabb, mint a másiké összhangban

lévő, egyező (konkordáns) párról beszélünk. Ha azonos értékek szerepelnek az esetpárokban kötött (tied) esetpárnak nevezzük. Eltérő, vagy széthúzó (diszkordáns) esetpárról abban az esetben beszélünk, ha az egyik változó értéke nagyobb, a másik változó értéke kisebb, mint a pár másik tagjáé. A számítás menete a konkordáns és diszkordáns párok különbségére épül, amelyek a különböző párok között számszerűsítünk. A pozitív asszociációs kapcsolat feltételezi, hogy párok többsége egyező (konkordáns), míg negatív kapcsolat esetében a párok többsége széthúzó (diszkordáns).

- **Gamma (GAMMA) együtthető¹¹:** bármely ordinális adatok és táblák összefüggéseinek vizsgálatánál alkalmazható és viszonylag könnyen értelmezhető. Az értéke -1 és 1 közé esik, ahol a 0 érték a változók függetlenségét mutatja. Ez jelenti, hogy mekkora valószínűséggel találjuk a konkordenciát, vagy a diszkordenciát meghatározónak a kutatott jelenségben. $\gamma = \frac{S - D}{S + D}$. Ha egy gyakorlati példánál a konkordencia az uralkodó jelleggel ($\gamma > 0$), az azt jelenti, hogy az egyik változó magasabb kategóriája a másik változó magasabb kategóriájának együtt járását váltja ki (pl. a szülők iskolai végzettségének vizsgálatakor a pozitív értékét jelentheti, hogy az apa magasabb iskolai végzettsége az anya magasabb iskolai végzettségét eredményezi, tehát magasabb végzettségű párt választ).
- **Somers féle d együtthető:** az ordinális változók közötti kapcsolatot -1 és +1 értéktartományok között méri, bármely tábla esetében, hasonlóan a gamma mutatóhoz. Az abszolút értékben 1-hez közeli érték szoros kapcsolatot jelent, azonban értelmezése a gamma mutatóhoz képest komplikáltabb.
- **Kendall tau-b (KENDALL'S TAU-B):** Szimmetrikus táblák, változók között használandó. A mutató -1 és +1 között értékeket vesz fel, ahol a +1 jelenti, hogy a párok sorrendje hasonló, megegyezik (konkordáns), míg a -1-es értéknél a párok sorrendje ellentétes (diszkordáns).
- **Kendall tau-c (KENDALL'S TAU-C):** Aszimmetrikus tábláknál használandó, értelmezése megegyezik a Kendall tau-b mutatószáméhoz.

A *STATISTICS* ablakban található *CORRELATION* opcióval részletesebben később a mennyiségi ismérvek összefüggéseinek vizsgálatakor foglalkozunk, hasonlóan, mint az ETA mutatóval a vegyes kapcsolatok számszerűsítésekor.

¹¹ Gyakran szerepel a mutatószám Goodman és Kruskal féle Gamma néven is.

- **Kappa (KAPPA)**¹² egy egyetértési mutatószám, mely az értékek (értékelők) egyetértését hivatott mérni. A mérőszám értelmezése Landis-Koch (1977) nyomán a következő:

8/9. táblázat A Kappa mérőszám értelmezése

K	Értelmezés
<0	Szegényes egyetértés
0-0,2	Enyhe egyetértés
0,21-0,4	Ígéretes egyetértés
0,41-0,6	Mérsékelt egyetértés
0,61-0,8	Megalapozott egyetértés
0,81-1	Csaknem tökéletes egyetértés

Forrás: Landis és Koch nyomán saját szerkesztés

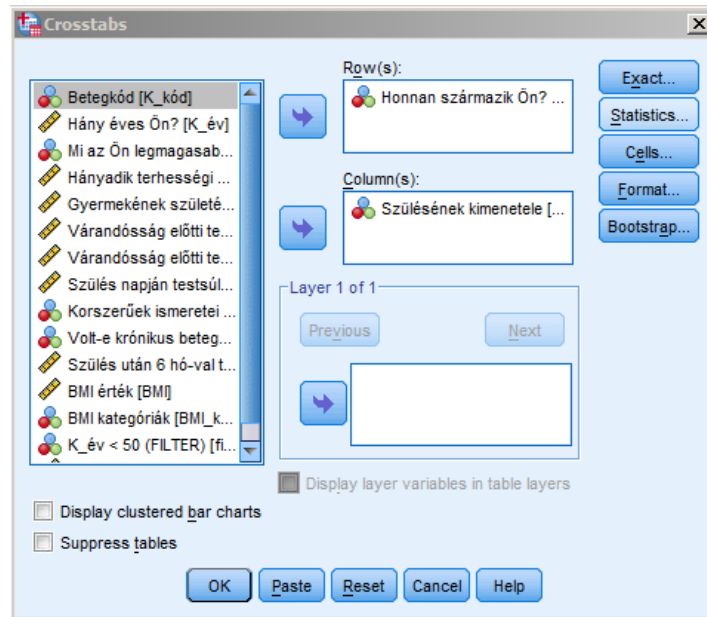
Amennyiben a túlzó kategória számot összevonva kívánjuk értelmezni, akkor azt is mondhatjuk, hogy 0,4 alatt gyenge egyetértésről, 0,4-0,8 között elfogadható egyetértésről, míg 0,8 felett kiváló egyetértésről beszélhetünk. Szimmetrikus tábláknál és az értékelők ugyanazon skálán mért véleményei során használható.

- **Kockázati hányados (RISK)** 2X2-es tábláknál alkalmazott kapcsolat-szorossági mutatószám. A számított eredmény értelmezésekor a mutatószám mellett (0 alsó határral és felső határ nélkül) egy konfidencia intervallumot is ad eredményül. Ha a mutató eredménye 1 és ezt az értéket a konfidencia intervallum is tartalmazza, akkor nem beszélhetünk összefüggésről. Amennyiben 0-nál vagy 1-nél nagyobb, akkor feltételezzük a változók közötti kapcsolatot. Összegezve a vizsgálat a relatív kockázatot és esélyhányadost számít 2X2-es tábláknál dichotóm változók esetében. A vizsgálat során az egyik ismérv okként, a másik egy eseményként értelmezhető.

Gyakorlati feladatként, kíváncsiak vagyunk, hogy van-e oksági kapcsolat a lakhely és a szülési kimenet között. Az eljárás segítségével összehasonlítjuk egy esemény (szülési kimenet) valószínűségét két csoport között.

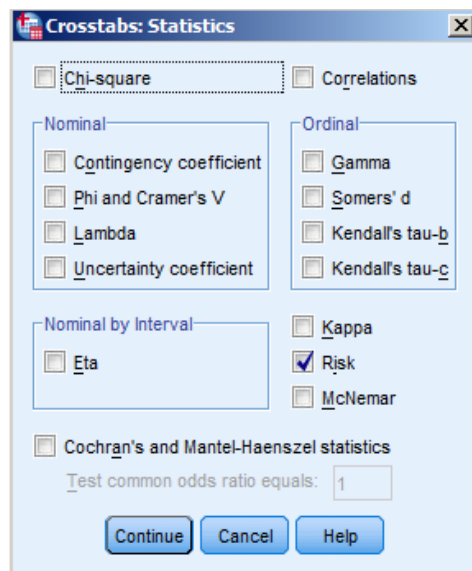
Első lépésként jelöljük ki a vizsgálatba bevonni kívánt változóinkat, ügyelve arra, hogy azok típusa dichotóm legyen, vagyis két kimenetele legyen.

¹² Gyakran szerepel a mutatószám Cohan- féle kappa néven is.



8/4. ábra: A változók kijelölése

A sor változóba a származási hely (város/falu) került, ami a vizsgálatba okként (hüvelyi/császármetszés), az oszlopváltozóba a szülési kimenete, ami eseményként értelmezendő. Ezt követően a *STATISTICS* modulban a mutatószámok kijelölése következik, amelyik esetünkben kizárólag a *RISK* együttható lesz.



8/5. ábra: A kockázati hányados mutatószám beállítása

A kijelölést követően a *CONTINUE* és *OK* gombok lenyomása után eljutunk az eredményekhez. Az eredmények közül jelen esetben kizárólag a harmadik táblázat értelmezésével foglalkozunk.

8/10. táblázat: A relatív kockázati tényezők és az esélyhányados eredményei

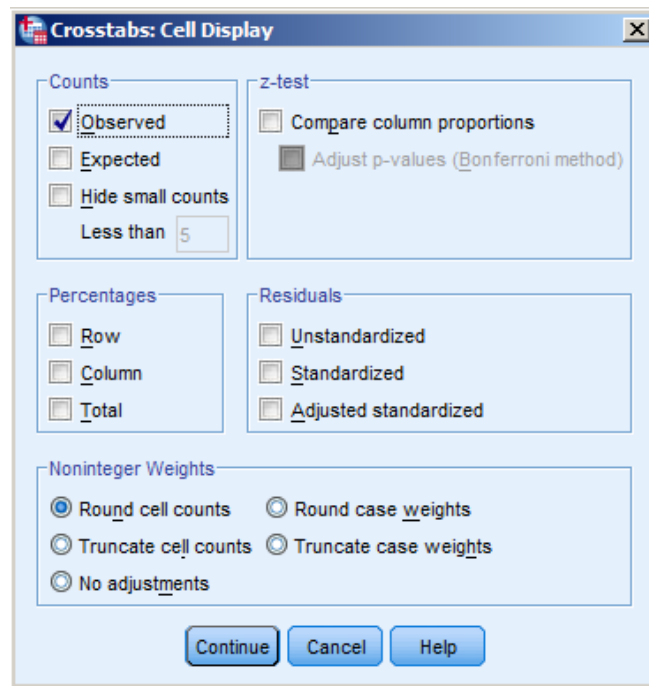
	Risk Estimate		
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for Honnan származik Ön? (város / falu)	2,029	,775	5,314
For cohort Szülésének kimenetele = hüvelyi	1,500	,856	2,627
For cohort Szülésének kimenetele = császármetszés	,739	,487	1,121
N of Valid Cases	70		

Forrás: saját szerkesztés

1. Odds Ratio: $2,029 = \frac{1,500}{0,739}$ azt jelenti, hogy több mint kétszeres (2,029) annak az esélye, hogy a városban lakó kismama hüvelyi úton fog szülni.
2. Relative Risk $_{\text{hüvelyi}} = 1,500$ azt jelenti, hogy a városban a hüvelyi szülés kockázata másfélszeres, mint a faluban.
3. Relative Risk $_{\text{császármetszés}} = 0,739$ azt jelenti, hogy a városban a császármetszéssel való szülés kockázata mekkora a faluhoz képest.
4. Az *OUTPUT* eredményeknél az *Odds Ratio* és *Relative Risk* értékek mellett látható egy számított megbízhatósági tartomány is, amely a kapcsolat létrejötté enged következtetni. Ha a számított intervallumba beleesik az 1, vagy éppen a mutató 1-es értéket vesz fel, akkor nincs kapcsolat a vizsgált változók között. Jelen esetben is ezt tapasztaljuk.

- **McNemar teszt (McNEMAR)** egy dichotóm változók közti kapcsolatot vizsgáló mutatószám, amely ugyanazon vizsgálati csoporton végzett kutatási mérés közötti változást hívatott vizsgálni. Azt méri, hogy a válaszadók hány százaléka választotta ugyanazon opciót a két mérés alkalmával. Gyakorlatban vélemények összehasonlítását kívánja mérni két időszakban (pl.: fogyasztói vélemények, választások stb.).
- **Cochran és Mantel-Haenzsel féle mutatószám (COCHRAN AND MANTEL-HANSSZEL STATISTICS)** két dichotóm változó összefüggését vizsgálja kontrollváltozók együttes hatását feltételezve. A mutatószám előnye, hogy egyszerre veszi figyelembe az összes kontrollváltozó hatását.

A statisztikai mutatószámok tárgyalását követően nézzük meg a *CELL* modulban található fontosabb beállítások értelmezését.



8/6. ábra: A CELL modulban kérhető beállítások

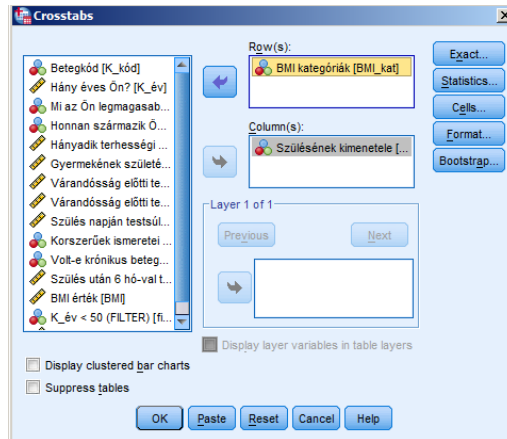
A bal felső nézetben az adatokra vonatkozó beállításokat láthatjuk, ahol az *OBSERVED* az egyedi (valós) megfigyelt értékeket, míg az *EXPECTED* a fiktív, függetlenség esetén feltételezett gyakoriságokat jelenti.

Az alatta lévő dobozban a százalékos arányszámokat kérhetjük (sor százalék=*ROW*; oszlop százalék=*COLUMN*; teljes százalék=*TOTAL*).

A sorszázalék azt jelenti, hogy a cellában tartalmazott gyakoriság a sor gyakoriság hány %-a. Az oszlopszázalék jelenti, hogy a cellában tartalmazott megfigyelt gyakoriság az oszlop gyakoriság hány százaléka. Az összes gyakoriság a cellagyakoriság „sorösszesen”- „oszlopösszesen” és mintanagyság hányadosa.

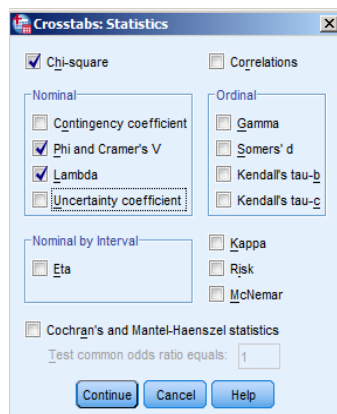
A reziduális értékek (*RESIDUALS*) dobozában található mutatók a tényleges és a függetlenség esetén feltételezett gyakoriságok különbségeiből kerülnek meghatározásra. Amennyiben az érték negatív, akkor a tényleges gyakoriság kisebb, mint a függetlenség esetén indokolt lenne. A három mutató közül az *ADJUSTED RESIDUAL* érték talán a leghasznosabb, hiszen a kapcsolat meglétéén túl, megjeleníti a kategóriák (ismérvváltozók) közül azokat, amelyek okozzák a kapcsolatokat. Amennyiben a számított érték abszolút értéke 2-nél nagyobb, akkor a két ismérvváltozat között szignifikáns kapcsolat van. Ha a számított érték abszolút értéke kettőnél kisebb, akkor a két kategória között nincsen szignifikáns kapcsolat.

Gyakorló feladatként, vizsgáljuk meg a fenti szemléltető példánk alapján, hogy van-e kapcsolat a szülés módja (típusa) és a BMI kategóriák között.



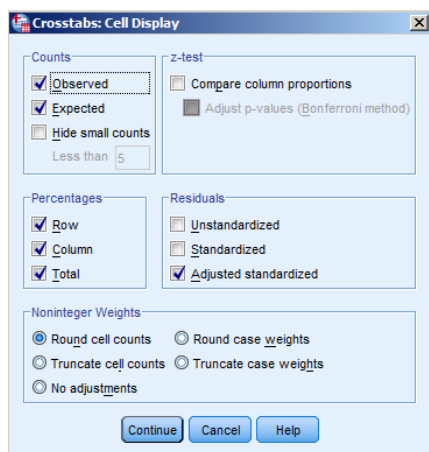
8/7. ábra: A keresztábrában szereplő változók kijelölése

Az oszlopváltozóban a szülés típusa (a függő változó szerepel, amely esemény eloszlására kíváncsiak vagyunk), a sorváltozóban (független, melyről okként vélelmezzük, hogy szerepet játszik a függő változó alakulásában) a BMI kategóriák szerepelnek.



8/8. ábra: A mutatószámok kalibrálása

A változók típusa alapján a bal felső négyzetben található mutatók közül választunk, hiszen mindkét változónk nominális. A khi négyzet értékét, a Cramer és Lambda mutatók kijelölését követően nyomjuk a *CONTINUE* gombot.



8/9. ábra: A CELLS modul beállításai

Amennyiben *COUNTS* dobozban kijelöljük az *OBSERVED* és az *EXPECTED* lehetőséget, úgy a kiszámítja nekünk a tényleges és a fiktív gyakorisági táblát is, amelyben a százalékos arányszámokat is számszerűsíti, amennyiben a *PERCENTAGES* ablakban az kijelöljük. A reziduumok számításánál csak az *ADJUSTED STANDARDIZED* opciót választottuk. A beállítások végeztével a *CONTINUE* és *OK* gombra kattintva a következő eredményeket (output) kapjuk.

Az output nézetben az elemzések elején egy összesítést (*case processing summary*) kapunk arról, hogy mekkora volt a megvizsgált esetek száma (N), hány valós (*valid*) és hány hiányzó értékünk (*missing*) volt és mindezek relatív megoszlása (*percent*) miként alakult.

8/11. táblázat: A relatív kockázati tényezők és az esélyhányados eredményei

	Case Processing Summary					
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
BMI kategóriák * Szülésének kimenetele	70	100,0%	0	0,0%	70	100,0%

Forrás: saját szerkesztés

Látható, hogy az elemszám 70 és minden kismama összes értékét ismerjük, nincsen hiányzó adat, vagyis a 100%-ot tudjuk értékelni.

A következő táblázat, a tényleges kereszttáblát tartalmazza, amelynek segítségével könnyen összehasonlíthatjuk a tényleges és a függetlenség esetén feltételezett (fiktív) gyakoriságokat. Itt is látható, hogy az értékek ugyan nem egyeznek meg, de nagy eltéréseket nem tartalmaz, ami sztochasztikus kapcsolat hiányára utal. A számított táblázatban szereplő értékek megegyeznek a fent számított értékeinkkel.

A táblázatban a következő adatok szerepelnek:

- A megfigyelések valós érték, gyakorisága (*COUNT*)
- A függetlenség esetén feltételezett (fiktív) gyakoriságok (*EXPECTED COUNT*)
- A sormegoszlás (*% WITHIN BMI KATEGÓRIÁK*)
- Az oszlopmegoszlás (*% WITHIN SZÜLÉSÉNEK KIMENETELE*)
- A teljes mintanagyság szerinti megoszlás (*% OF TOTAL*)
- A standardizált, korrigált reziduum (*ADJUSTED RESIDUAL*)

8/12. táblázat: A vizsgált keresztátlában szereplő értékek összefoglaló táblázata

BMI kategóriák * Szülészeti kimenetele Crosstabulation

BMI kategóriák	Szülészeti kimenetele		Total	
	hüvelyi	császármetszés		
Count	2	3	5	
Expected Count	2,1	2,9	5,0	
% within BMI kategóriák	40,0%	60,0%	100,0%	
% within Szülészeti kimenetele	6,7%	7,5%	7,1%	
% of Total	2,9%	4,3%	7,1%	
Adjusted Residual	-,1	,1		
Count	13	10	23	
Expected Count	9,9	13,1	23,0	
% within BMI kategóriák	56,5%	43,5%	100,0%	
% within Szülészeti kimenetele	43,3%	25,0%	32,9%	
% of Total	18,6%	14,3%	32,9%	
Adjusted Residual	1,6	-1,6		
Count	7	14	21	
Expected Count	9,0	12,0	21,0	
% within BMI kategóriák	33,3%	66,7%	100,0%	
% within Szülészeti kimenetele	23,3%	35,0%	30,0%	
% of Total	10,0%	20,0%	30,0%	
Adjusted Residual	-1,1	1,1		
Count	8	13	21	
Expected Count	9,0	12,0	21,0	
% within BMI kategóriák	38,1%	61,9%	100,0%	
% within Szülészeti kimenetele	26,7%	32,5%	30,0%	
% of Total	11,4%	18,6%	30,0%	
Adjusted Residual	-,5	,5		
Total	Count	30	40	70
Expected Count	30,0	40,0	70,0	
% within BMI kategóriák	42,9%	57,1%	100,0%	
% within Szülészeti kimenetele	100,0%	100,0%	100,0%	
% of Total	42,9%	57,1%	100,0%	

56,6% (13/23) azon normál BMI alkatú egyedeknek az aránya, akik hüvelyi úton szültek.

43,3% (13/30) az összes azon egyedből, aki hüvelyi úton szült normál BMI alkatú.

A mintába került összes egyed közül 18,6% (13/70) aki hüvelyi úton szült és normál BMI testalkat kategóriával rendelkezett

Ha mutató +2 feletti értéket vesz fel, akkor ott biztosan van szignifikáns összefüggés, ha -2 alattit akkor biztosan nincsen.

Forrás: saját szerkesztés

A következő táblában számunkra ismert érték tűnik fel a Pearson Chi-Square mutató mellett, hiszen ez megegyezik az általunk számolt négyzetes kontingencia mutató értékével.

8/13. táblázat: A számított khi-négyzet értéket tartalmazó táblázat

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2,743 ^a	3	,433
Likelihood Ratio	2,741	3	,433
Linear-by-Linear Association	,967	1	,325
N of Valid Cases	70		

a. 2 cells (25,0%) have expected count less than 5. The minimum expected count is 2,14.

A df a szabadságfokot jelöli, melynek számítása $df=(sor-1)*(oszlop-1)$. A szabadságfoknak fontos szerepe van az elméleti érték megállapításakor. Hiszen a tapasztalati érték a négyzetes kontingencia mutató (χ^2) és az elméleti érték összevetése után tudjuk megállapítani, hogy a nullhipotézist elfogadjuk vagy elvetjük.¹³ A χ^2 eloszlástáblázat a szabadságfok (3) és egy adott hibavalószínűség (0,05) mellett megadja viszonyítási értéket, mely esetünkben 7,815 (Excelben: =inverz.khi (0,05;3)). Mivel a tapasztalati érték alacsonyabb, mint az elméleti, ezért a nullhipotézist elfogadjuk, vagyis nincsen kapcsolat a két ismerv között. Ez azt jelenti, hogy a szülési kimenetet nem a BMI kategóriák határozzák meg. Ez a szignifikanciára vonatkozó táblázatokból (*DIRECTIONAL MEASURES*, *SYMMETRIC MEASURES*) értékeiből is leolvasható, hiszen nagyobb, mint az általunk választott 5%. A legvégső táblázatban a szimmetrikus mutatókat (*SYMMETRIC MEASURES*) találjuk.

¹³ Részletesen a későbbiekben.

8/14. táblázat: A számított kapcsolat-szorossági mérőszámok

Directional Measures			Value	Asymp. Std. Error ^a	b	Approx. Sig.
Nominal by Nominal	Lambda	Symmetric	,091	,103	,860	,390
		BMI kategóriák Dependent	,085	,100	,820	,412
Goodman and Kruskal tau	Lambda	Szülésének kimenetele Dependent	,100	,152	,627	,530
		BMI kategóriák Dependent	,018	,021		c
	tau	Szülésének kimenetele Dependent	,039	,047		c
		BMI kategóriák Dependent				

Not assuming the null hypothesis.

Using the asymptotic standard error assuming the null hypothesis.

Based on chi-square approximation

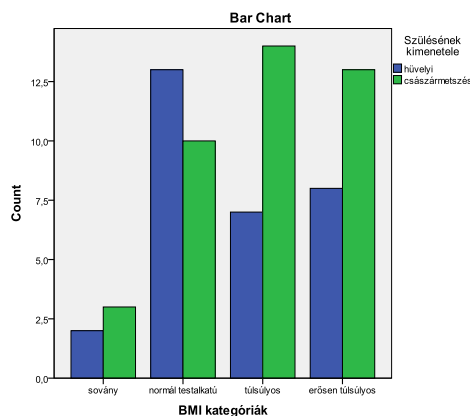
Symmetric Measures			
		Value	Approx. Sig.
Nominal by Nominal	Phi	,198	,433
	Cramer's V	,198	,433
N of Valid Cases		70	

Not assuming the null hypothesis.

Using the asymptotic standard error assuming the null hypothesis.

Forrás: saját szerkesztés

A Cramer-együttható értékéből leolvasható, hogy gyenge kapcsolat van a két ismérv között. A kereszt táblák készítésénél felkínálja a program a grafikus ábrázolást (*DISPLAY CLUSTERED BAR CHARTS*) lehetőségét is. Alaplehetőségként oszlopdiagramon keresztül szemlélteti az asszociációs kapcsolatot.



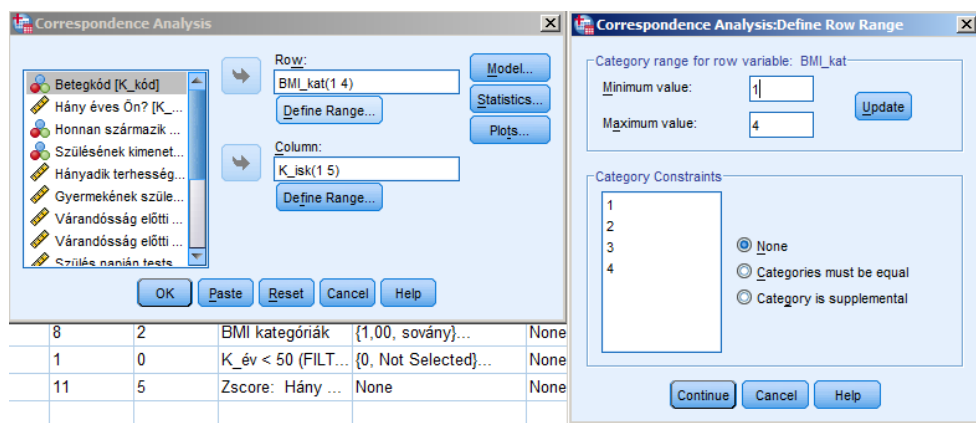
8/10. ábra: A kereszt tábla grafikus megjelenítése

Fontos ismét megemlíteni, a fenti példa a szemléltetés kedvéért került bemutatásra, mivel tudva lévő, hogy nem szerencsés a khi-négyzet alapú mutatók alkalmazása abban az esetben, ha az ismérvváltozók közül (cellák), több mint 20%-ban az érték 5 alatti.

8.3. Korrespondencia- analízis

Az asszociációs kapcsolatot grafikus szemléltetésére alkalmazhatjuk, a manapság egyre népszerűbb *korrespondencia-analízist*. „A korrespondencia-analízis lehetővé teszi, két nominális változó kapcsolatának grafikus megjelenítését egy többdimenziós, de a szemléletesség és a könnyű értelmezhetőség kedvéért kis dimenziószámú térben (általában síkban). Az egymáshoz hasonló kategóriák ezekben az ábrázolásokban is közel kerülnek egymáshoz. Az eredmények értelmezése az alkalmazott normalizáló eljárástól függ. A SPSS-ben az alapértelmezett normalizálás a sor- és az oszlopváltozók kapcsolatát elemzi.” (Ketskeméty – Izsó 2005, 417.o.). A szemléltetés kedvéért vizsgáljuk meg, hogy a BMI kategóriák és az iskolai végzettség között létezik-e összefüggés.

Ez az eljárás a *DATA REDUCTION* főmenü *CORRESPONDENCE ANALYSIS* almenüjében érhető el.



8/11. ábra: A korrespondencia-analízis beállításai az SPSS programban

Először jelöljük ki a sor- (*ROW*) és oszlopváltozókat (*COLUMN*). Ezután minden egyes ismérvet definiálni kell, a benne szereplő ismérvváltozatok számának segítségével.

Jelen esetben a BMI kategóriák típusát definiáltuk 1-től 4-ig, az ismérvváltozatok számának függvényében. Miután mindkét ismérvet meghatároztuk, a többi beállításon ne változtassunk és nyomjuk meg az *OK* gombot. Az eredményeket a következő táblázatokban láthatjuk.

Az első táblázat (*CORRESPONDENCE TABLE*) egy kereszttáblához hasonlóan, a tényleges gyakoriságokat tartalmazza, a másodikban az összesítő eredmények olvashatók. Itt látható, hogy a kapcsolat szignifikáns ($p=0,03$), a khi-négyzet érték magas $\chi^2=22,78$) illetve a létrejövő két dimenzió alkalmas a megjelenítésre, hiszen az értékek szóródásának 94,2%-át magyarázza. A következő két táblázat az egyes ismérvváltozatok koordinátáit tartalmazzák az

alapbeállításként szereplő két dimenzió mentén. Talán a legszemléletesebb lehet számunkra a grafikus megjelenítés (*BIPLOT*), amely segítségével az összetartozó értékek két dimenzió mentén láthatóvá válnak, vagyis egyfajta megjelenítését adja az *ADJUSTED RESIDUAL* értékeknek.

8/15. táblázat: A korrespondencia- analízis eredményei

Correspondence Table

BMI kategóriák	Mi az Ön legmagasabb iskolai végzettsége?					
	egyetem, főiskola	felsőfokú szakképzés	OKJ	érettségi	8 általános	Active Margin
sovány	2	1	0	0	2	5
normál testalkatú	4	11	3	2	3	23
túlsúlyos	4	9	5	0	3	21
erősen túlsúlyos	2	2	5	7	5	21
Active Margin	12	23	13	9	13	70

Summary

Dimension	Singular Value	Inertia	Chi Square	Sig.	Proportion of Inertia		Confidence Singular Value	
					Accounted for	Cumulative	Standard Deviation	Correlation 2
1	,490	,240			,737	,737	,099	-.009
2	,258	,067			,205	,942	,111	
3	,137	,019			,058	1,000		
Total		,325	22,780	,030 ^a	1,000	1,000		

a. 12 degrees of freedom

Overview Row Points^a

BMI kategóriák	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		Total
					1	2	1	2	
sovány	,071	-.375	1,811	,065	,020	,907	,075	,925	1,000
normál testalkatú	,329	-.389	-.206	,038	,102	,054	,648	,096	,743
túlsúlyos	,300	-.549	-.182	,056	,184	,038	,794	,046	,839
erősen túlsúlyos	,300	1,064	-.024	,167	,694	,001	,999	,000	,999
Active Total	1,000			,325	1,000	1,000			

a. Symmetrical normalization

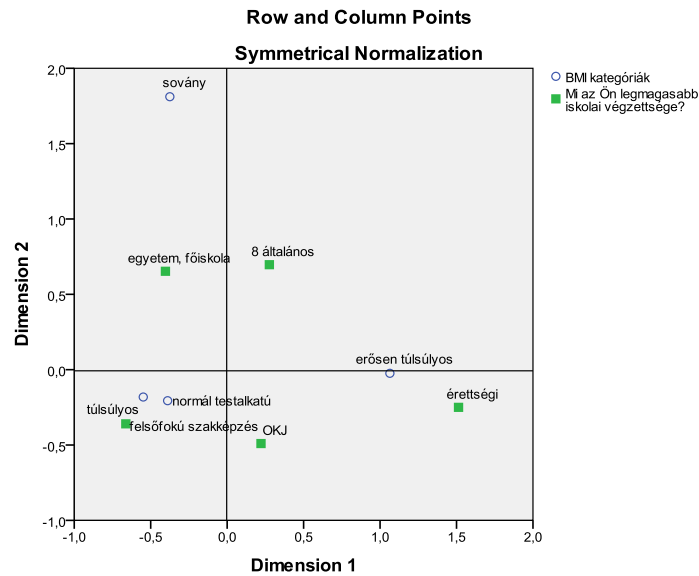
Overview Column Points^a

Mi az Ön legmagasabb iskolai végzettsége?	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		Total
					1	2	1	2	
egyetem, főiskola	,171	-.404	,653	,033	,057	,283	,420	,580	1,000
felsőfokú szakképzés	,329	-.663	-.359	,085	,295	,164	,833	,129	,962
OKJ	,186	,221	-.490	,028	,019	,173	,162	,419	,581
érettségi	,129	1,513	-.250	,150	,601	,031	,961	,014	,975
8 általános	,186	,276	,697	,030	,029	,349	,227	,764	,991
Active Total	1,000			,325	1,000	1,000			

a. Symmetrical normalization

Forrás: saját szerkesztés

A grafikus megjelenítésnél a dimenziók elnevezése kulcsfontosságú, hiszen a megértés segítése a célja, mely a kutató önálló, saját feladata.



8/12. ábra: A korrespondencia- analízis grafikus megjelenítése

A korrespondencia-analízisről az érdeklődő részletesen Jánosa (2011) könyvében olvashat.

8.4. Vegyes kapcsolat vizsgálata

A vegyes kapcsolatokban az ok szerepét mindig a minőségi ismerv tölti be, míg az okozat(ok)ét a mennyiségi ismerv(ek). A vegyes kapcsolat vizsgálata azt célozza, hogy megállapítsuk a kvantitatív változóban rejlő információ mekkora része határozható meg a minőségi ismerv szerinti csoportosítással. A mennyiségi ismerv lehetőséget teremt arra, hogy a számítási eljárásokat kibővítsük. A vegyes kapcsolat szorosságának mérése a szórás felbontásának összefüggésén alapul. A szórásnégyzetek közötti összefüggést a szórásnégyzet összetevőkre bontásának nevezünk. E szerint a teljes szórásnégyzet a belső szórásnégyzet és a külső szórásnégyzet összegeként írható fel.

$$\sigma^2 = \sigma_B^2 + \sigma_K^2$$

A belső- és külső szórásnégyzet kiszámításának módja:

$$\sigma_B^2 = \frac{\sum_{j=1}^m n_j \sigma_j^2}{n}$$

$$\sigma_K^2 = \frac{\sum_{j=1}^m n_j (\bar{x}_j - \bar{x})^2}{n}$$

A szórásnégyzetek összefüggését kifejező képletet átrendezhetjük, ha elosztjuk az egyenlőség mindkét oldalát a teljes szórásnégyzettel:

$$1 = \frac{\sigma_B^2}{\sigma^2} + \frac{\sigma_K^2}{\sigma^2}$$

A csoportosító (minőségi) ismerv – ami egyben a sztochasztikus kapcsolat ok szerepét is betölti – hatását a külső szórás közvetíti. Könnyen belátható, hogy amennyiben a külső szórás nulla, a minőségi ismérvnek semmilyen mérhető hatása nincs; a két ismerv (a minőségi és mennyiségi) független. Az ellenkező szélsőséges esetben – amennyiben a belső szórás nulla – a külső szórás megegyezik a teljes szórással, tehát a kapcsolat determinisztikus. Ezek alapján a külső és a teljes szórás segítségével számszerűsíthető a vegyes kapcsolatot mérő **szóráshányados** mutatója:

$$H = \frac{\sigma_K}{\sigma} = \sqrt{\frac{\sigma_K^2}{\sigma^2}} = \sqrt{1 - \frac{\sigma_B^2}{\sigma^2}}$$

Az eltérés négyzetösszegekre is igaz, hogy a teljes eltérésnégyzet-összeg a külső és a belső eltérésnégyzet-összeg összegeként adódik:

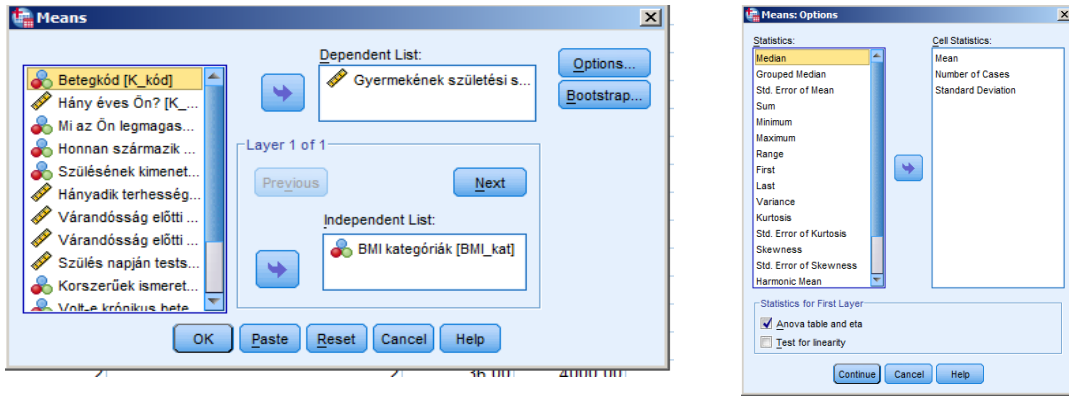
$$SS = SS_B + SS_K = \sum_{j=1}^m \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2 = \sum_{j=1}^m f_j (\bar{x}_j - \bar{x})^2 + \sum_{j=1}^m \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

Így az is igaz:

$$H^2 = \frac{SS_K}{SS}$$

Nézzünk most egy konkrét példát: arra vagyunk kíváncsiak, hogy a kismama szülés előtti BMI kategóriája, mennyire befolyásolja a születési súlyt, vagyis van-e kapcsolat a születési testsúly és a kismama BMI kategóriája között.

Az SPSS program segítségével könnyű dolgunk van, hiszen a változók kalibrálása és az eredmények értelmezése a feladatunk. A beállításokat az *ANALYZE* menü, *COMPARE MEANS* almenüjének *MEANS* moduljában végzünk. A beállítások viszonylag gyorsan végezhetőek, hiszen csak a függő és a független változókat kell először megadni, majd az *OPTIONS* gomb használatával a kért műveleteket tudjuk beállítani.



8/13. ábra: A vegyes kapcsolat beállításai az SPSS programban

A legfontosabb beállítás az *ANOVA TABLE AND ETA* doboz kijelölése. Az eta (η) a tulajdonképpen a már tanult módon írható fel:

$$\eta = \sqrt{\frac{\sigma^2_K}{\sigma^2}}$$

A *CONTINUE* majd az *OK* gombok lenyomását követően kapjuk meg a számított eredményeket, melyek most is összesítő táblákkal kezdődnek. A következő táblázatban (*REPORT*) a részátlagokat, tagszámot és a szórásokat láthatjuk.

8/16. táblázat: A kategóriák alapadatai

Report

Gyermekekének születési súlya?

BMI kategóriák	Mean	N	Std. Deviation
sovány	2840,0000	5	713,61754
normál testalkatú	3276,9565	23	515,78215
túlsúlyos	3221,4286	21	529,69128
erősen túlsúlyos	3612,3810	21	446,66436
Total	3329,7143	70	547,43929

Forrás: saját szerkesztés

Ezt követi az ANOVA tábla, mely a belső és külső eltérésnégyzet-összegeket és a szignifikanciára vonatkozó eredményeket közli.

8/17. táblázat: A vegyes kapcsolat eredményeinek táblázata

		Sum of Squares	df	Mean Square	F	Sig.
Gyermekének születési súlya? * BMI kategóriák	Between Groups (Combined)	3187269,234	3	1062423,078	4,009	,011
	Within Groups	17491325,05	66	265020,077		
	Total	20678594,29	69			

Belső eltérésnégyzet- összeg

Külső eltérésnégyzet- összeg

Forrás: saját szerkesztés

A szignifikancia érték alapján látható ($p < 0,05$), hogy a vizsgált változók szignifikáns összefüggést mutatnak, vagyis az eredmény általánosítható (nem a véletlen műve). Ezt követően megkapjuk a kapcsolat-szorossági mérőszámokat is.

8/18. táblázat: A kapcsolat-szorossági mérőszámok táblázata

	Eta	Eta Squared
Gyermekének születési súlya? * BMI kategóriák	,393	,154

H

H^2

Forrás: saját szerkesztés

A szóráshányados mutató (H) segítségével megállapíthatjuk, hogy közepes szorosságú kapcsolat van a kismama szülés előtti BMI típusa és a kisbaba születési súlya között. A szórásnégyzet- hányados megmutatja, hogy a BMI kategóriák típusa a születési súly 15,4 %-át magyarázza. Vélhetően a fennmaradó 85,6 %-ot más tényezők befolyásolják (pl.: a terhesség alatt elfogyasztott ételek mennyisége és minősége stb.)

8.5. Korrelációs kapcsolat vizsgálata

Amennyiben mind az ok(ok) mind az okozat szerepét is mennyiségi ismérvek közvetítik, **korrelációs kapcsolat**ról beszélünk. A továbbiakban elsősorban egy *tényező*, vagy *magyarázó változó* (X) és egy *eredményváltozó* (Y) közötti kapcsolat mérését mutatjuk be, viszont hangsúlyozni kívánjuk, hogy a valóságban általában nem egy, hanem több tényező együttes, igen összetett hatására alakul ki egy-egy jelenség, folyamat. A korrelációs kapcsolat mérése során azonban több ok együttes hatásának vizsgálatát is viszonylag könnyen meg lehet oldani. A korreláció természete szerint a változók között az alábbi kapcsolatok értelmezhetők: monoton kapcsolat, illetve ezen belül lineáris kapcsolat. Mindkét kapcsolat lehet pozitív, vagy negatív irányú, melynek megítélését segíti a grafikus megjelenítés. Két mennyiségi ismérv

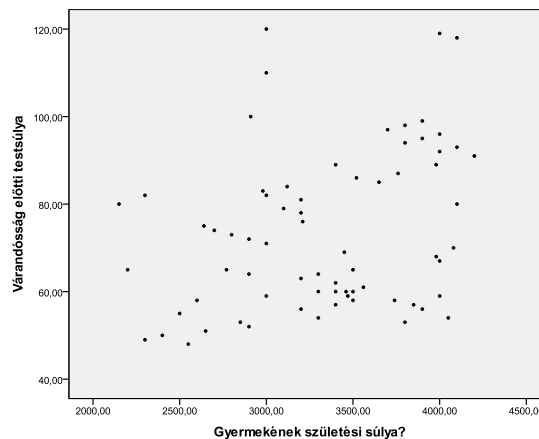
között meglévő kapcsolatot jól ábrázolhatjuk a derékszögű koordináta rendszerben, ún. pontdiagram segítségével. Bővebben az érdeklődő Pintér – Rappai (2007): Statisztika című könyvében olvashat.

A korrelációs kapcsolat mérésének legelterjedtebb mutatószáma a **lineáris korrelációs együttható** (jele: **r**), amelynek alkalmazása során feltételezzük a változók közötti lineáris kapcsolatot, illetve ha a linearitás feltevése nem áll távol a vizsgált problémától. A korrelációs együttható kiszámítása a változók együttmozgását jellemző kovariancia mérőszáma és a változók szórása segítségével történik, az alábbi algoritmus segítségével:

$$r_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y}, \text{ ahol a kovariancia : } C_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{\sum d_x d_y}{n} = \frac{\sum xy}{n} - \bar{x}\bar{y}$$

ahol: σ_x és σ_y a változók szórásai.

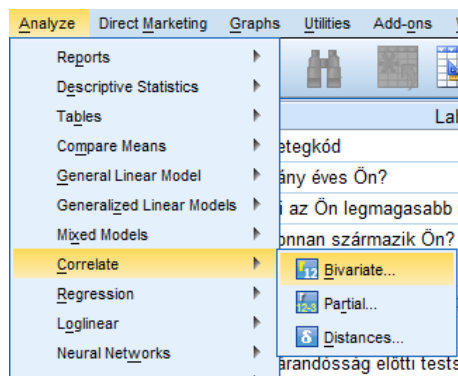
Most a korrelációs kapcsolat szorosságát vizsgáljuk meg egy gyakorlati példa segítségével, vagyis elemezzük a kismama szülés előtti testsúlyának és a kisbaba születési súlyának kapcsolatát. Első lépésben pontdiagram segítségével ábrázoljuk a mért eredményeket.



8/14. ábra: Pontdiagram ábrája

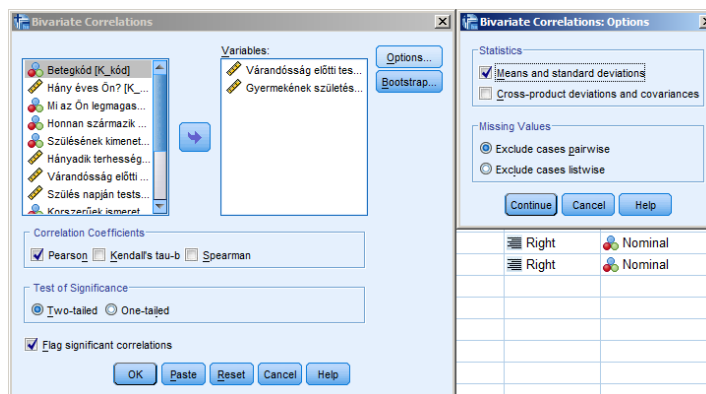
A grafikus ábra alapján feltételezhető a vizsgált változók közötti pozitív irányú lineáris kapcsolat. Növekszik a várandósság előtti testsúly és növekszik a baba születési súlya.

Az ezt követő beállításokat az SPSS programban az **ANALYZE** menüben a **CORRELATE** menüpont **BIVARIATE** moduljával érhetjük el.



8/15. ábra: A korreláció elérési útvonala

Ezt követően a változók vizsgálatba vonása következik. A változókat a középső nyíl segítségével tudjuk a jobb oldali ablakba mozgatni. Jelöljük ki a várandósság előtti testsúly változót, valamint a gyermek születési súlya változókat.



8/16. ábra: A változók bevonása a korrelációs elemzésbe

Látható, hogy a számunkra megfelelő Pearson-féle korrelációs együttható szerepel alapbeállításként. Az *OPTIONS* gombra kattintva kérhetünk további alapstatisztikákat is. Itt jelöljük ki a *MEANS AND STANDARD DEVIATIONS* modult. A *CONTINUE* és *OK* gomb lenyomását követően eljutunk az eredményekhez.

8/19. táblázat: A leíró statisztikát tartalmazó táblázat

Descriptive Statistics

	Mean	Std. Deviation	N
Várandósság előtti testsúlya	73,5286	18,18800	70
Gyermekeinek születési súlya?	3329,7143	547,43929	70

Forrás: saját szerkesztés

Az első táblázatban a számtani átlag, szórás és az elemszám szerepel. Ezt követően a korrelációs együtthatót tartalmazó táblázat következik.

8/20. táblázat: A korrelációs együtthatót tartalmazó táblázat

Correlations

		Várandósság előtti testsúlya	Gyermekének születési súlya?
Várandósság előtti testsúlya	Pearson Correlation	1	,309**
	Sig. (2-tailed)		,009
	N	70	70
Gyermekének születési súlya?	Pearson Correlation	,309**	1
	Sig. (2-tailed)	,009	
	N	70	70

** . Correlation is significant at the 0.01 level (2-tailed).

Forrás: saját szerkesztés

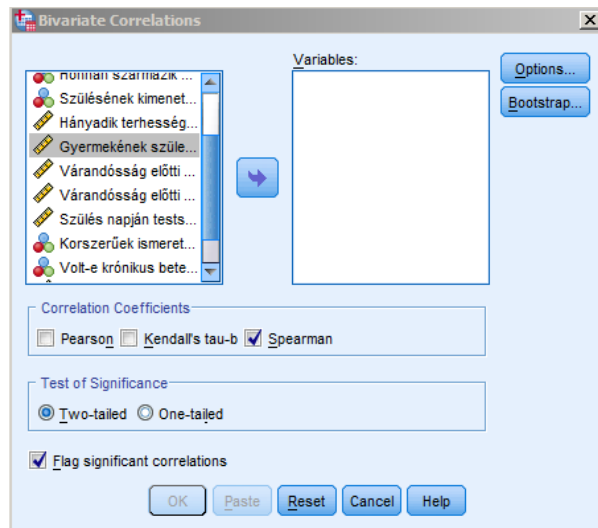
Az eredmények szerint a kapcsolat közepes erősségű ($0,3 < R < 0,7$). A szignifikancia szint alapján elmondható, ($p < 0,05$) hogy az összefüggés nem a véletlen műve, vagyis általánosítható. A korrelációs együttható négyzete a determinációs együtthatót adja eredményül, azaz a determinációs együttható alapján kijelenthető, hogy a várandósság előtti testsúly 9,55 %-ban határozzák meg a kisbaba születési testsúlyát.

Természetesen, ha az adataink sorba lennének rendezve (rangsor), vagyis ordinális skálán szereplnének, akkor is kínál a program számítási lehetőséget. Monoton kapcsolat esetén¹⁴ a kapcsolat szorosságát a Spearman-féle rangkorrelációs együtthatóval mérjük, ami egy robusztusabb mérőszám, azaz érzéketlenebb a kiugró értékekre, hiszen az intervallum- vagy arányskála helyett ordinális skálát használ. Ez azt jelenti, hogy egy magasabb rendű skáláról alacsonyabb rendű skálára is alakíthatjuk át az adatainkat. A rangkorrelációs együttható

$$\text{képlete: } \rho = 1 - \frac{6 \sum_{i=1}^n [R(y_i) - R(x_i)]^2}{n(n^2 - 1)} = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2 - 1)}$$

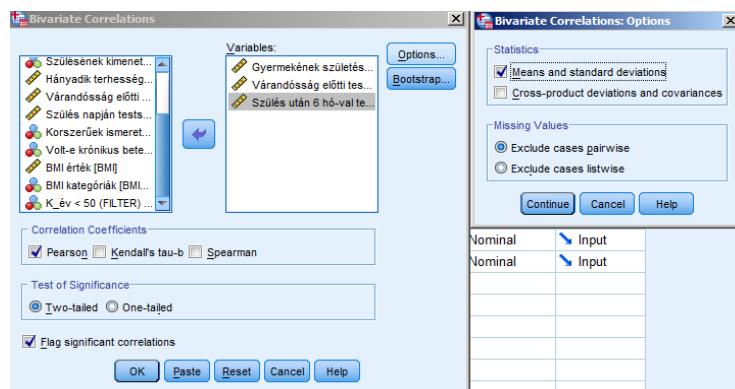
Az SPSS programban az *ANALYZE* menüben a *CORRELATE* menüpont *BIVARIATE* moduljával érhetjük el a rang-korrelációs számítást. A korrelációs együtthatónál a *PEARSON* alapbeállítás helyett a *SPEARMAN* mutatót kell választani. A mutató értelmezése megegyezik a lineáris Pearson együttható során ismertetett értelmezéssel.

¹⁴ Monoton kapcsolatnál Y változásának mértéke X egységnyi elmozdulása esetén nem konstans.



8/17. ábra: A rangkorreláció számítás beállítása

A következő példánál egy korrelációs mátrixot hozunk létre, amikor kettőnél több (mennyiségi, folytonos, skála típusú) változó összefüggés rendszerét elemezzük. Bővítsük vizsgálatunkat, vonjuk be a kismama, szülést követő 6 hónapban mért testsúly változót is.



8/18. ábra: A változók számának bővítése

Ezt követően futassuk le az elemzést, és végeredményként egy korrelációs-mátrixot kapunk.

8/21. táblázat: A korrelációs mátrix

Correlations

		Gyermekének születési súlya?	Várandósság előtti testsúlya	Szülés után 6 hó-val testsúlya
Gyermekének születési súlya?	Pearson Correlation	1	,309**	,237*
	Sig. (2-tailed)		,009	,049
	N	70	70	70
Várandósság előtti testsúlya	Pearson Correlation	,309**	1	,904**
	Sig. (2-tailed)	,009		,000
	N	70	70	70
Szülés után 6 hó-val testsúlya	Pearson Correlation	,237*	,904**	1
	Sig. (2-tailed)	,049	,000	
	N	70	70	70

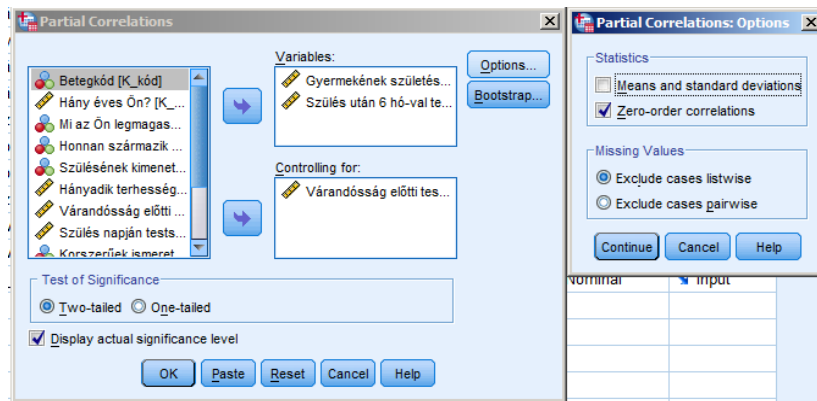
** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Forrás: saját szerkesztés

Az átlóban szereplő értékek értelemszerűen egy, hiszen a változók önmagukkal determinisztikus kapcsolatban vannak. Új eredményként látható, hogy a gyermek születési súlya és a szülést követően a kismama 6 hónapban mért testsúlya között pozitív irányú szignifikáns ($p=0,049$), de gyenge kapcsolatot ($0,24$) számszerűsítettünk, míg a kismamák várandósság előtti és a szülés utáni testsúlya szignifikáns ($p=0,00$) szoros kapcsolatban vannak ($0,9$). Az értékek mögött található csillagok a szignifikáns kapcsolatokat jelentik az értelmezését a későbbiekben még bővebben tárgyaljuk.

Érdekes eredményeket kaphatunk némely esetben, ha a parciális korrelációk számítását is megnézzük. Az SPSS program segítségével ezt az amúgy időigényes számítást viszonylag egyszerűen és gyorsan megtehetjük, az *ANALYZE* menü *CORRELATE* almenüjének, *PARTIAL* modulja segítségével. A *parciális korrelációs együttható* két változó közötti kapcsolatot méri, egy vagy több más változó hatásának kiszűrésével. A módszer segítségével lehetőség van a nem látszólagos összefüggések felfedezésére is. Most vizsgáljuk meg a jelenséget ennek függvényében:



8/19. ábra: A parciális korreláció számítása az SPSS- ben

A beállításoknál a *VARIABLES* dobozba azokat a változókat kell elhelyezni, amelyek közvetlen korrelációs kapcsolatát vizsgáljuk, míg a *CONTROLLING FOR* dobozba az(ok) a változó(k) kerül, amely hatását ki szeretnénk szűrni. Az *OPTIONS* panelen a *ZERO-ORDER CORRELATIONS* jelöljük ki. Ennek tudatában három különböző beállítást végezhetünk, melyek eredményei közül az egyik érdekes eredményt mutat:

8/22. táblázat: A parciális korrelációs együtthatót tartalmazó táblázat

Correlations			Gyermekének születési súlya?	Szülés után 6 hó-val testsúlya	Várandósság előtti testsúlya
Control Variables -none. ^a	Gyermekének születési súlya?	Correlation	1,000	,237	,309
		Significance (2-tailed)	.	,049	,009
		df	0	68	68
	Szülés után 6 hó-val testsúlya	Correlation	,237	1,000	,904
		Significance (2-tailed)	,049	.	,000
		df	68	0	68
Várandósság előtti testsúlya	Correlation	,309	,904	1,000	
	Significance (2-tailed)	,009	,000	.	
	df	68	68	0	
Várandósság előtti testsúlya	Gyermekének születési súlya?	Correlation	1,000	-,104	
		Significance (2-tailed)	.	,394	
		df	0	67	
	Szülés után 6 hó-val testsúlya	Correlation	-,104	1,000	
		Significance (2-tailed)	,394	.	
		df	67	0	

a. Cells contain zero-order (Pearson) correlations.

Forrás: saját szerkesztés

A táblázat alsó része mutat újdonságot számunkra, hiszen itt már a várandósság előtti testsúly kiszűrésével vizsgáljuk a gyermek születési súlyát, illetve a kismama, szülést követő 6. hónapban mért testsúly változók összefüggését. A keletkező eredményt láthatjuk, amely már nem is mutat szignifikáns értéket (Sig.=0,394) és a pozitív irányú gyenge kapcsolat is eltűnt, helyette a kapcsolat iránya negatív lett. Ez az jelenti számunkra, hogy a szülést megelőző testsúly értékre is szükségünk van. A másik két beállításnál ilyen „álösszefüggéseket” nem találtunk.

9. KÖVETKEZTETÉSES VIZSGÁLATOK (Ács Pongrác)

9.1. Bevezetés, elméleti alapok

A hétköznapi életben nagyon sokszor fordul elő, hogy egy eseményről, jelenségről nem rendelkezünk teljes körű információval.

A *következtetési statisztikai módszerek* mindazon eljárások, melyek segítségével a megfigyelhető egyedek egy részének felmérése után, a sokaság valamennyi egyedére vonatkozó megállapításokat teszünk. Elmondható, hogy az így létrejövő adatbázis nem teljes körű, hanem csak egy bizonyos technikával (mintavételi technika) kiválasztott részsokaságra, mintára vonatkozik. Ennek köszönhetően a következtetési statisztikai eljárások során a "bizonytalanság" mindig jelen van, vagyis a hibázási lehetőség rizikója létezik. A jól választott statisztikai módszertannal és a korrekt értelmezéssel, azonban ez a hibázás kontroll alatt tartható.

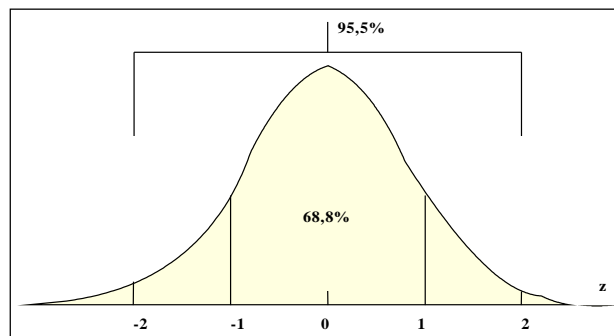
A következtetési statisztikai eljárások két nagy csoportjával a *becslési és a hipotézis ellenőrzési vizsgálatokkal és módszerekkel* fogunk megismerkedni, leginkább a gyakorlati felhasználás aspektusából.

A társadalmi és gazdasági jelenségek, vagy akár a sportteljesítmények és eredmények jelentős köréről feltesszük, hogy folytonos, normális eloszlású valószínűségi változóként viselkednek. A folytonos valószínűségi változók tulajdonsága, hogy egy adott intervallumban végtelen számú értéket vehetnek fel, és annak valószínűsége, hogy egy X változó pontosan x értékét veszi fel, zérus. A valószínűségi eloszlások fontos „azonosítója” a várható érték (μ) és a variancia, szórásnégyzet (σ^2). A normális eloszlás könnyen azonosítható a várható érték és a szórás segítségével, jele: $N(\mu, \sigma)$. A normalitás feltételezésével élünk pl. a súly, a térfogat, magasság, hosszúság, BMI és a teljesítmények esetében.

A várható értékek és a szórások, az elemzés tárgyától függően, igen sokféle értéket vehetnek fel, ami a munkát sokszor megnehezíti, hiszen nagyságuk a változók dimenziójától függ. A *standardizálás* segítségével azonban ez a probléma megoldható, ami azt jelenti, hogy a várható értéket kivonjuk a valószínűségi változó értékéből, és a különbséget elosztjuk a szórással, így egy *standard normális eloszlású valószínűségi változót* (jele: z) kapunk eredményül. Képletben:

$$z = \frac{x - \mu}{\sigma}$$

A standardizálás eredményeként kapott standard normális eloszlású valószínűségi változó várható értéke zérus, szórása egységnyi, azaz $N(0,1)$. Mindkettő - a normális és a standard normális eloszlású - valószínűségi változó sűrűségfüggvénye ún. harang-görbével, **Gauss-görbével** (9/1. ábra) jellemezhető. Standard normális eloszlás esetén mind a valószínűségi változók, mind a hozzájuk rendelhető valószínűségek táblázatba foglalhatók, melyek segítségével a kapott értékek könnyen és viszonylag gyorsan felhasználhatók gyakorlati problémák megoldására, bár a számítógépek korában az ehhez hasonló táblázatok szép lassan kikopnak a gyakorlati alkalmazások repertoárjából, vagyis nagyon ritkán használjuk őket.



9/1. ábra: Fontosabb valószínűségek a z függvényében

A várható értéktől egységnyi szórással eltérő intervallum – és ez nemcsak a standard, hanem az általános normális eloszlás esetére érvényes – és a valószínűségi görbe által bezárt terület 68,8 %-os valószínűséget reprezentál. A kétszeres szórás által meghatározható intervallumhoz tartozó valószínűség 95,5 %; míg a háromszoros szórással lefedhetjük a vízszintes tengely és a görbe által meghatározható teljes területet, szinte a teljes valószínűséget (99,9 %). Mivel a standard normális valószínűségi változó sűrűségfüggvénye szimmetrikus, így elegendő a 0 és a pozitív végtelen közé eső számokhoz tartozó valószínűségi értéket meghatározni.

A statisztikai középértékek – különösen a számtani átlag – kiemelt fontossággal bírnak a következtetési statisztikában is. Közvetlenül adódik annak igénye, hogy a reprezentatív módon kiválasztott minták átlagai és szórásai, valamint az alapsokaság átlaga és szórása között valamilyen összefüggést keressünk. Hangsúlyozni kell, hogy a **centrális határeloszlás** tétele értelmében bármilyen eloszlással rendelkező alapsokaságból egyszerű véletlen mintavétel segítségével nyert minta átlaga valószínűségi változó, mivel értéke mintáról mintára ingadozik, ugyanakkor az **átlagok normális eloszlású valószínűségi változók**. Mindez természetesen fokozottan aláhúzza a normális eloszlás gyakorlati hasznosíthatóságát, elterjedtségét.

A következtetési statisztika igényli különböző összefüggések felismerését a mintaátlagok, azok szórása és az alapsokasági átlag és szórás között. Könnyen belátható, hogy amennyiben ismerjük valamennyi minta átlagát, a minták átlagából képzett átlag megegyezik az alapsokasági átlaggal. A mintaátlagok szórása azonban eltér az alapsokaság szórásától.

Létezik azonban – bizonyítás nélkül közöljük – egy olyan összefüggés, amelynek segítségével közvetlen kapcsolat írható fel az alapsokasági szórás (szórásnégyzet) és a mintaátlagok szórása (szórásnégyzete) között:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \left[\frac{N-n}{N-1} \right]$$

ahol: n a mintaelemek száma és N az alapsokaság elemeinek száma.

Itt jegyezzük meg, hogy a kifejezés második tagját, az $\left[\frac{N-n}{N-1} \right]$ tényezőt korrekciós tényezőnek vagy véges szorzónak hívja az irodalom. A visszatevés nélküli kiválasztás¹⁵ esetén játszik fontos szerepet, visszatevéses mintavétel alkalmazása során nem szerepel a képletben. Itt kell szólni arról, hogy a korrekciós tényezőt elhagyhatjuk visszatevés nélküli kiválasztás, azaz egyszerű véletlen mintavétel esetén is, amennyiben az alapsokaság (N) nagysága jelentősen eltér a minta (n) nagyságától, mivel ilyen esetekben a tényező 1-hez közeli értékkel bír.

A mintaátlag szórásnégyzete $\sigma_{\bar{x}}^2$, egy olyan átlagos négyzetes hiba, amelyet akkor követünk el, amikor következtetéseink során a sokasági várható értéket mindig a mintaátlaggal helyettesítjük. A statisztikai módszerek között kiemelkedő fontossággal bír a mintaátlag szórása ($\sigma_{\bar{x}}$), amit a mintaátlag *standard hibájának* neveznek. Az alapsokaság szórásának ismeretében tehát könnyen kiszámítható a mintaátlagok szórása.

A véletlen minta elemei véletlen változók, ezért bármely transzformációjuk, így a belőlük számított számtani átlag is, véletlen változó lesz.

¹⁵ A visszatevés nélküli mintavétel (pl. egyszerű véletlen mintavétel) a gyakorlatban igen népszerű, mivel alkalmazása nem jár információvesztéssel.

9.2. A statisztikai becslések

A statisztikai becslés az ismeretlen alapsokaság valamely konstans paraméterének közelítő jellegű meghatározása. Ilyen paraméterek: várható érték (véges alapsokaságnál, átlag), szórás és az arány.

Láttuk azonban, hogy az alapsokaság átlaga, valamint a mintaátlagok között közvetlen, a szórás és a mintaátlagok szórása között is jól kifejezhető összefüggés írható fel. Különösen fontos szerepet tölt be a standard hiba, a mintaátlagok szórása. Ez a szóródási mérőszám lehetőséget ad arra, hogy a becslésünket egy olyan intervallummal adjuk meg, aminek a bekövetkezése, adott valószínűségi szinten, garantálható.

A $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \left[\frac{N-n}{N-1} \right]$ képlet alapján szükségünk van az alapsokasági szórás ismeretére, ha mintánk van, akkor a korrigált mintabeli szórást használjuk, melynek képlete:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

A korrigált mintabeli szórás segítségével felírható a gyakorlatban jól használható standard hiba képlete is, melynél a véges szorzót

$$\left(1 - \frac{n}{N} \right)$$

akkor használjuk, ha a mintánk nagysága meghaladja az alapsokaság nagyságának 5%-át.:

$$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Hangsúlyoznunk kell, hogy a fenti standard hiba képlete csupán az átlagok szóródását jellemzi. Más paraméterekre pl. értékösszeg, arány is felírhatók a megfelelő szórások, más néven standard hibák.

Azokat a mintából származó statisztikákat, melyeket az alapsokasági paraméterek közelítő meghatározására használnak, becselőfüggvénynek nevezik. A becselőfüggvény egy adott mintára vonatkozó konkrét értékét, **pontbecslésnek** hívják. A becslés során elkövethető véletlen hiba átlagos nagyságát a standard hiba (becselőfüggvény szórása) szolgáltatja. A következő táblázat a leggyakrabban használt alapsokasági paraméterbecslések fő jellemzőit tartalmazza.

9/1. táblázat: Legfontosabb sokasági paraméterek becslőfüggvényi és azok jellemzői

Alapsokasági paraméter	Torzítatlan becslőfüggvény	Standard hiba	Becslőfüggvény eloszlása
várható érték	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	$S_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}}$	kis minta (n<50) t- eloszlás nagy minta (n≥50) normális
arány	$p = \frac{k}{n}$	$S_p = \sqrt{\frac{p(1-p)}{n}}$	kis minta (n<50) binomiális nagy minta (n≥50) normális

Forrás: Pintér – Rappai (2001)

A gyakorlatban jól használható információt nyerünk azonban akkor, ha intervallumbecslést végzünk. Az **intervallumbecslés** során felhasználjuk azt, hogy a minta-paraméterek valamilyen ismert eloszlású valószínűségi változók, és így az adott eloszlás értékének felhasználásával egy **adott megbízhatósági szinten** állapíthatunk meg egy intervallumot. Ezt az intervallumot **konfidencia intervallumnak** hívjuk. Az intervallumok meghatározásához szükséges kritikus érték - a normális eloszlás szimmetrikus voltából adódóan - a 0-ra szimmetrikusan helyezkedik el. A pontbecslés, a standardhiba és az eloszlás típusának ismeretében a konfidencia intervallumot (ez egy pontbecslés, amely köré mindkét irányba felvesszük a hibahatárt) már felírhatjuk. A hibahatár tartalmazza az általunk pozitív és negatív irányba tolerált maximális „pontatlanságot”. Az átlagbecslés esetén a konfidencia intervallum:

$$\bar{x} \pm z \times \sigma_{\bar{x}}$$

ahol: z a standard normális eloszlás adott értéke, melyek közül a fontosabbakat az alábbiak:

9/2. táblázat: Gyakran használt kritikus értékek

α	$1-\alpha$	$Z_{(\alpha/2)}$	$Z_{(1-\alpha/2)}$
0,01	0,99	-2,576	2,576
0,05	0,95	-1,96	1,96
0,1	0,9	-1,645	1,645

Forrás: saját szerkesztés

Ennek tükrében a becslések gyakorlata 6 lépésből áll. Az első három elméleti feladat (mintavétel, becslőfüggvény konstruálása, becslőfüggvény megítélése), míg a második három a gyakorlati munka (pontbecslés, standard hiba meghatározás, intervallumbecslés)¹⁶.

Ennek segítségével nézzünk egy konkrét példát:

Gyakorlati feladatként ismerjük az adatbázisunkból 70 kisbaba születéskor mért testsúlyát (gramm). Becsüljük meg 95 %-os megbízhatóság mellett a kisbabák átlagos testsúlyának értékét!

¹⁶ Pintér – Rappai (2001)

Az átlag (mintaátlag) számítása:

$$\bar{\mu} = \frac{\sum x}{n} = 3329,71$$

és a szórás (korrigált szórás, hiszen mintáról van szó):

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} = 547,44$$

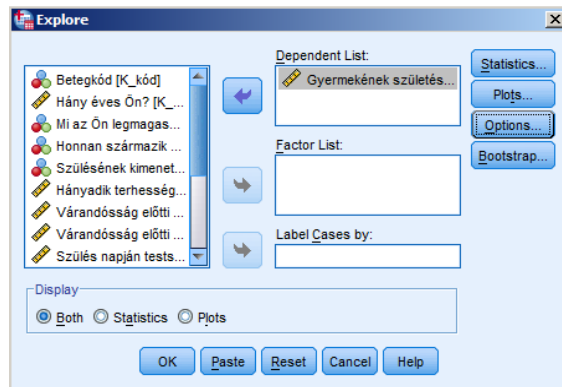
eredményeit felhasználva, kiszámíthatjuk a standard hibát is (Véges szorzót nem használunk, mivel a minta nagysága az alapsokaság nagyságának 5%-át nem haladja meg):

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = 65,43$$

Mivel nagy mintánk van ($n > 50$) ezért a normális eloszlást táblaértékeiből keressük ki a kritikus értéket, melyet a fenti táblázatban láthatunk : $z = 1,96$. ($Z(1-\alpha/2)$). A hibahatár értéke ($z \times \sigma_{\bar{x}}$): 128,25, mely után a végeredmény a következő: $3329,71 \pm 128,25$

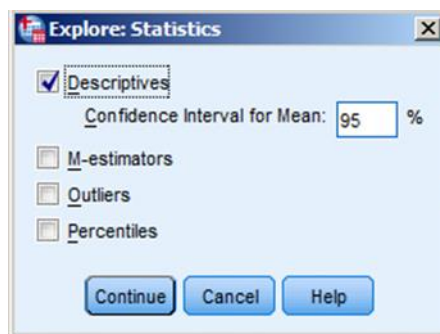
Itt fontos megjegyezni, hogy az SPSS program nem fog különbséget tenni az elemszám alapján a becslőfüggvény típusa között, és itt is a t-eloszlásból fog számolni. Ez a végeredményben is kisebb (nem számottevő) változást fog mutatni: $z = 1,9949$ (tszf, megbízhatóság), vagyis a hibahatár értéke 130,53 és a végeredmény $3329,71 \pm 130,53$. Ez azt jelenti, hogy a 95%-os megbízhatósági szinten kijelenthető, hogy a kisbabák születési súlya 3199,18 és 3460,25 gramm között lesz (intervallumbecslés). Ehhez hasonló számítási módszerrel számítható, becsülhető akár a többi folytonos változó is (pl.: a szülés várható hete, a kismama 6 hónap utáni testsúlya stb.). Természetesen ezt a számítást és a hozzá tartozó grafikus megjelenítést a SPSS programon kívül az Excel programban is ábrázolhatjuk, melyet az érdeklődő részletesen megtalál Oláh tankönyvében (Oláh 2008).

Az SPSS program is viszonylag gyorsan elvégzi az egyszerű becsléseket. Az *ANALYZE* menü, *DESCRIPTIVE STATISTICS, EXPLORE MODULBAN*.



9/2. ábra. A változók kijelölésének modulja

A fent számított gyakorlati feladatot számszerűsítsük az SPSS program segítségével. Az első lépésként a vizsgálatba célváltozóját (függő változóját) mozgassuk a középső nyíl segítségével a *DEPENDENT LIST* ablakba. Ezt követően a *STATISTICS* modult választjuk, ahol az általunk választott vizsgálatot fogjuk kijelölni.



9/3. ábra: A statisztika becslések beállításai

A *DESCRIPTIVES* modult választottuk, ahol a megbízhatósági tartomány értékét tudjuk beállítani. A feladat alapján a 95%-os megbízhatóságot választottuk. Ezt követően nyomjuk meg a *CONTINUE* és *OK* gombokat. A létrejövő eredmények számunkra már ismerősen fognak visszaköszönni.

9/3. táblázat: A becslés végeredményét tartalmazó táblázat

Descriptives			Statistic	Std. Error
Gyermekének születési súlya?	Mean		3329,7143	65,43151
	95% Confidence Interval for Mean	Lower Bound	3199,1820	
		Upper Bound	3460,2466	
	5% Trimmed Mean		3346,5079	
	Median		3400,0000	
	Variance		299689,772	
	Std. Deviation		547,43929	
	Minimum		2150,00	
	Maximum		4200,00	
	Range		2050,00	
	Interquartile Range		905,00	
	Skewness		-,294	,287
	Kurtosis		-,856	,566

Forrás: saját szerkesztés

A végeredmény egyező lesz az általunk korábban számszerűsített eredménnyel, természetesen azzal a megjegyzéssel, hogy a program t-eloszlással számolt.

Az átlagbecsléséhez hasonlóan becsülhető az alapsokaság valamilyen ismerv szerinti aránya (megoszlási viszonzyszáma) is. Valamely tulajdonsággal bíró egyed arányát jelöljük az alapsokaságban P -vel. A **P arány pontbecslése**:

$$p = \frac{k}{n}$$

ahol: k a mintában az adott tulajdonsággal bíró egyedek száma, n a minta elemszáma.

A mintabeli aránynak a mintából számítható standard hibája:

$$\sigma_p = \sqrt{\frac{p(1-p)}{n}}$$

$$\sigma_p = \sqrt{\frac{p(1-p)}{n} \left(1 - \frac{n}{N}\right)}$$

Nagy minta esetén joggal feltételezzük, hogy p eloszlása közelíthető a normális eloszlással, ezért a konfidencia intervallum szerkesztéséhez felhasználhatjuk a standard normális eloszlás értékeit.

A konfidencia intervallum:

$$p \pm z \times \sigma_p$$

Gyakorlati feladatként kutatási adatbázisunk segítségével 95,5 %-os megbízhatóság mellett határozzuk meg a császármetszéssel születendő kisgyermek arányát!

$$p = \frac{k}{n} = \frac{40}{70} = 0,5714 = 57,14\%$$

$$0,5714 \pm 2 \times \sqrt{\frac{0,5714 \times 0,4286}{70}}$$

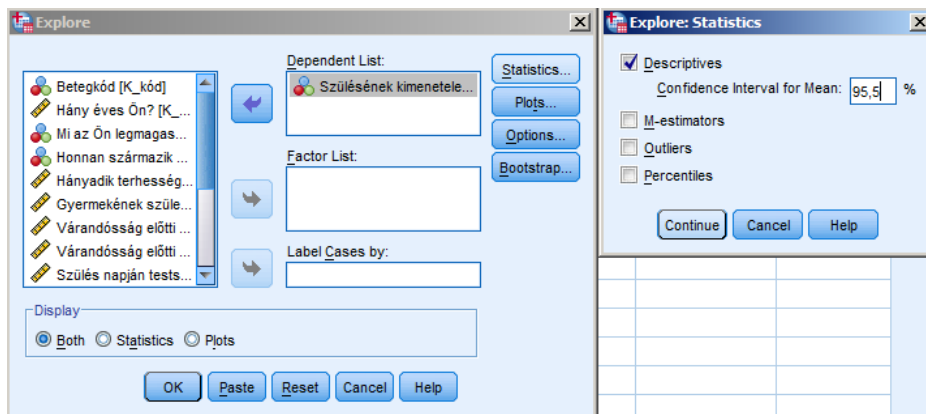
$$0,5714 \pm 0,1182$$

45,31%

68,97%

Tehát 95,5%-os megbízhatósági szint mellett megállapíthatjuk, hogy a császármetszéssel születendő gyermekek aránya 45,31%-a és legfeljebb 68,97 %.

Az aránybecslésnél sincsen bonyolult dolgunk, hiszen ebben az esetben a függő változónk a szülési kimenet módja (szülésének kimenete) lesz, azonban a *STATISTICS* modulban a konfidencia intervallumban a 95,5%-ot jelöljük.



9/4. ábra: A statisztika aránybecslések beállításai

A *CONTINUE* és *OK* gombok lenyomását követően eljutunk a végeredményhez, mely szintén egyezik az általunk korábbiakban számolt eredménnyel.

9/4. táblázat: Az aránybecslés eredménye

Descriptives			Statistic	Std. Error
Szülésének kimenetele	Mean		1,57	,060
	95.5% Confidence Interval for Mean	Lower Bound	1,45	
		Upper Bound	1,69	
	5% Trimmed Mean		1,58	
	Median		2,00	
	Variance		,248	
	Std. Deviation		,498	
	Minimum		1	
	Maximum		2	
	Range		1	
	Interquartile Range		1	
	Skewness		-,295	,287
	Kurtosis		-1,970	,566

Forrás: saját szerkesztés

Tehát 95,5%-os megbízhatósági szint mellett megállapíthatjuk, hogy a császármetszéssel születendő gyermekek aránya 45%-a és legfeljebb 69%.

9.3. Különbözőség-vizsgálatok

A különbözőségvizsgálatok olyan matematikai-statisztikai elemző eljárások, melyek segítségével (szubjektív hatásokat kiszűrve) meg tudjuk állapítani, egy adott változót tekintve kettő vagy több rész minta között van-e jelentős különbség. A különbözőség-vizsgálatok során különböző típusú változók elemzése történik. Fontos megemlíteni, hogy a különbözőség-vizsgálatok is a hipotézisvizsgálatok és elemzések csoportjába tartozik, azonban az egyik leggyakrabban előforduló módszertanok között tartjuk számon, ezért egy rövid elméleti áttekintést kívántunk itt is adni. Természetesen a részletes gyakorlati alkalmazások bemutatása a hipotézisvizsgálatok során fog megtörténni.

Amennyiben a különbözőség- vizsgálatok során szignifikancia vizsgálatot is végzünk, akkor nemcsak a mintákra, hanem az általuk reprezentált populációk közötti különbözőségekkel kapcsolatosan is pontos eredményhez jutunk.

Az eljárás segítségével megállapíthatjuk, hogy a vizsgálat során megfigyelt különbség az általunk meghatározott valószínűségi szinten más esetben is kimutatható-e. Ha a különbség igazolhatóan nem a véletlen műve, akkor szignifikáns különbségnek nevezzük.

Általában két mintát 95%-os valószínűségi szint felett tekintünk szignifikánsan különbözőnek, vagyis a tévedés lehetősége nem nagyobb, mint 5%. Az ilyen szignifikancia szint jelölése $p < 0,05$. (100 esetből kevesebb, mint 5 esetben hibázunk).

A különbözővizsgálatok leggyakrabban alkalmazott módszerei:

1. **Önkontrollós vizsgálat:** egy minta vagy részminta legalább egy változóját két különböző időpontban rögzítjük. A statisztikai módszertanban gyakran említik, párosított mintás próbának is.
2. **Kontrollcsoportos vizsgálat:** két egymástól független minta vagy részminta összehasonlítása ugyanazzal a feltáró eszközzel történő vizsgálat alapján. (pl. a fiúk és lányok dohányzási szokásai különböznek-e).
3. **Összetett kontrollcsoportos vizsgálat:** két vagy több részmintánál ugyanaz a változó alapján létezik-e különbség.

9/5. táblázat: A vizsgálatok lehetséges csoportosítása

		Változók típusa		
		Mért, intervallumskálán értelmezett adatok (paraméteres vizsgálatok)	Rangsorolt, ordinális adatok (nem paraméteres vizsgálatok)	Megállapítható, nominális adatok
Minták száma és viszonya	egy minta, (két időpontban)	egy mintás t-próba, párosított mintás t-próba (önkontrollós vizsgálat)	Wilcoxon-próba (önkontrollós vizsgálat)	Khi- négyzet próba
	kettő minta	kétmintás t-próba, F- próbával (Kontrollcsoportos vizsgálat)	Mann-Whitney próba (Kontrollcsoportos vizsgálat)	Khi- négyzet próba
	több, mint kettő	Varianciaanalízis (ANOVA, Összetett kontrollcsoportos vizsgálat)	Kruskal-Wallis próba (Összetett kontrollcsoportos vizsgálat)	Khi- négyzet próba

Forrás: saját szerkesztés

9.4. Hipotézis-vizsgálatok (paraméteres és nem paraméteres próbák a gyakorlatban)

A következtetési statisztika egyik leggyakrabban alkalmazott módszereinek összefoglaló neve. A hipotézisellenőrzés (feltevés-vizsgálat) olyan statisztikai módszer, mely alkalmas egy választott statisztikai próba (teszt) segítségével egy-egy feltevés elfogadásáról vagy elvetéséről való döntés meghozatalában. Tehát a feltevések (hipotézisek), egy-egy sokaság jellemzőjét (átlagát, arányát stb.), eloszlási paraméterét (pl. várható érték), az alapsokaság eloszlását (pl. normális eloszlás) tartalmazzák többnyire egzakt matematikai-statisztikai formában. Így lehetővé válik az, hogy a hipotéziseket a matematikai-statisztika eszközeivel, meghatározott valószínűség figyelembevételével ellenőrizzük, és végezetül a feltevést elfogadjuk, vagy elvessük.

A hipotézisellenőrzés mindig egy adott hipotézisrendszerre vonatkozik, amely mindig egy nullhipotézisből (H_0) - kiinduló feltevés - és egy vele szemben álló alternatív hipotézisből (H_1) áll. A hipotézisellenőrzés végeredménye mindig egy igen-nem (elfogadom-elvetem) típusú döntés, mely csak egy adott hibaváltoszínűség mellett érvényes. Egy dolgot feltétlenül

tartsunk szem előtt, hogy a döntés kizárólag a nullhipotézisre vonatkozik. A későbbiekben jelöljük Θ -val (theta) az ismeretlen alapsokasági értéket és Θ_0 -val a feltételezett értéket!

A legtöbb tudomány területén gyakran a kísérleti módszer alkalmazásánál valósul meg, ami a gyakorlatban sokszor azt jelenti, hogy valamely új, vagy más módszerről kívánjuk eldönteni, hogy van-e valamilyen hatása a vizsgálat tárgyát képező egyedekre. Két vagy több csoportot képzünk, és azt vizsgáljuk, hogy a csoportoknál például az eltérő foglalkoztatás (edzés módszer) eredményeként milyen változás tapasztalható. A betegeknél két csoportot képzünk és két különböző fizioterápiás módszerrel foglalkoztatjuk őket. A nullhipotézis azt mondja, hogy a két csoport eredményei között nincs lényeges eltérés, vagyis a két módszer fejlesztő hatásában nincsen különbség. Kiinduló *nullhipotézisünket* az alábbi módon írhatjuk fel:

$$H_0: \Theta = \Theta_0$$

Természetesen ez a kifejezés önmagában még nem értelmezhető, meg kell fogalmazni ellentét-párját, azaz az *alternatív hipotézis*, így lesz hipotézisrendszer, amely az egész „eseményteret” lefedi. Az előző példánál maradva: a két csoport eredményei között van különbség (nem a véletlen hatása), viszont azt nem tudjuk melyik a jobb, ilyenkor az alternatív hipotézis *kétoldali*:

$$H_1: \Theta \neq \Theta_0$$

illetve *egyoldali*, ha állást foglalunk, hogy az eredmények alapján melyik csoport eredményei jobbak, melyik módszer a jobb :

$$H_1: \Theta < \Theta_0$$

vagy

$$H_1: \Theta > \Theta_0$$

A fent megfogalmazott hipotézisek ellenőrzését matematikai függvények ún. *próbafüggvények* segítségével végezhetjük el. A függvény lehetővé teszi az ismert statisztikai eloszlástípusoknak megfelelő elméleti értékkel való összevetést. Egy adott valószínűségi szint ún. *szignifikancia szint* mellett a számított értéket az elméleti értékkel összehasonlítva, a hipotézist vagy elvetjük, vagy elfogadjuk; ezáltal teszteljük az adott alapsokaságra megfogalmazott állításunkat.

A vizsgálat menete így négy lépésben folyik:

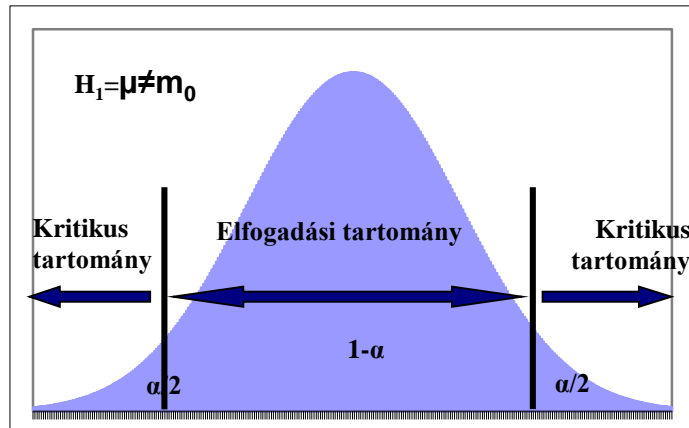
1. Az első lépésben fel kell állítani a hipotézisrendszert (H_0 és a H_1 meghatározása).
2. A megfelelő próbafüggvény kiválasztása.
3. A mintaelemek alapján számított (empirikus) próbafüggvény-érték meghatározása.
4. Döntés.

A próbafüggvény kiválasztása szempontjából fontos az alapsokaság eloszlása, a mintavétel módja és a minta nagysága. Leggyakrabban független azonos eloszlású (FAE) mintát feltételezünk, viszont az alapsokaság eloszlásáról gyakran csak elképzelésünk van (ennek vizsgálata: illeszkedésvizsgálat).

A hipotézisellenőrzés során a próbafüggvény értéktartományát logikailag két egymást kizáró tartományra osztjuk: az elfogadási tartományra és a kritikus (elutasítási) tartományra. Meg tudjuk mondani, hogy a próbafüggvényünk értéke milyen valószínűséggel esik az elfogadási tartomány határai közé. Döntést úgy hozunk, ha a próbafüggvény a kritikus tartományba esik, akkor a nullhipotézist elvetjük, különben elfogadjuk. A próbafüggvény kritikus (elutasítási) tartományba esésének valószínűségét nevezzük szignifikancia szintnek.

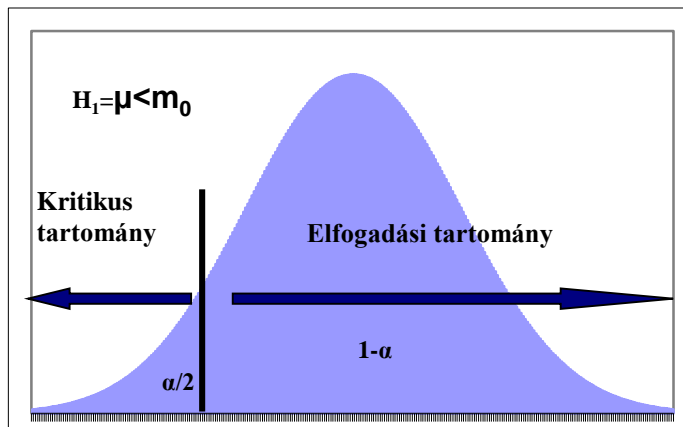
A döntéshozatal másik módszere a *szignifikancia érték (p érték)* alapján történik, ami azt mutatja meg, hogy az nullhipotézis elvetése milyen valószínűséggel okoz hibát. A döntés, ha a p érték alacsony szám, akkor kicsi az elsőfajú hiba elkövetésének valószínűsége, ezért célszerű elutasítani a nullhipotézist. Ez szemben, ha a p érték nagy elfogadjuk a nullhipotézist. A tartományok elhelyezkedése többféle lehet, mely az alternatív hipotézistől függ. Ha a kritikus tartományba esés valószínűsége α , akkor az elfogadási tartományba esése: $1-\alpha$.

Tételezzük fel, hogy a hipotézisünk a várható érték (μ) és egy feltételezett érték (m_0) egyenlőségére vonatkozik. Az egyik leggyakrabban alkalmazott hipotézisvizsgálati probléma annak vizsgálata, hogy a sokasági várható érték egy előre adott konstanssal egyezik-e, az ilyen próbát *egymintás várható érték* próbának nevezzük. Ilyenkor egy sokaság várható értékének egy konkrét számmal történő egyezőségét teszteljük, különböző alternatív hipotézisekkel szemben.

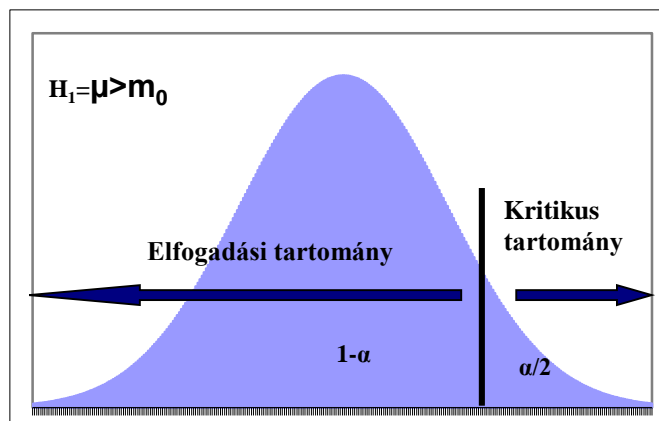


9/5. ábra: Elfogadási és kritikus tartomány kétoldali (two tailed) alternatív hipotézis esetén

A kritikus tartományba esés valószínűsége α , mivel két egyenlő nagyságú részből áll a kritikus tartomány ezért, egyes részekbe $\alpha/2$ valószínűséggel esik a függvény. Ha a nullhipotézissel szemben azt állítjuk, hogy a várható érték nemcsak, hogy nem egyenlő, hanem nagyobb vagy kisebb, akkor egyoldalas jobb széli (right tailed), vagy bal széli (left tailed) kritikus tartományt kapunk.



9/6. ábra: Elfogadási és kritikus tartomány bal oldali alternatív hipotézis esetén



9/7. ábra: Elfogadási és kritikus tartomány jobb oldali alternatív hipotézis esetén

A próbákat leggyakrabban egy- vagy kétmintásnak nevezzük és vonatkozhatnak a sokasági várható értékekre, szórásra, illetve arányra is, ennek megfelelően a leggyakoribb egymintás tesztek próbafüggvényei:

Nullhipotézis	Nagyminta ($100 \leq n$)	Kisminta ($n < 100$)
$H_0 : \mu = \mu_0$	$\bar{x} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} H_0 \sim N(0;1)$	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} H_0 \sim_{n-1} t$
$H_0 : P = P_0$	$\bar{x} = \frac{P - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} H_0 \sim N(0;1)$	
$H_0 : \sigma^2 = \sigma_0^2$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} H_0 \sim_{n-1} \chi^2$	

Nézzünk most a próbát a gyakorlati alkalmazás során!

A szakdolgozat írása során egy hallgató egy külföldi tudományos cikkben azt olvasta, hogy a svéd kismamáknak a szülés előtti BMI indexének átlaga 25,6. Az adatbázisunk segítségével ismerjük meg, a mintába került kismamák ($n=70$) átlagos BMI értékét, 26,99 és a korrigált szórását: 6,08. Kíváncsiak vagyunk, hogy van-e a svéd és magyar kismamák BMI értékében különbség! Elfogadhatjuk-e, hogy a magyar kismamák BMI értéke a svéd kismamák mért testtömeg-index eredményét nem haladhatja meg?

$$H_0: \mu=25,6$$

$$H_1: \mu>25,6$$

A nullhipotézisben tehát azt feltételezzük, hogy a magyar kismamák BMI értékének várható átlagos értéke megegyezik a svéd kismamák várható testtömeg index értékével (a svéd átlagtól való eltérés csak véletlen tényezőknek tekinthető). Az alternatív hipotézisben pedig azt fogalmazzuk meg, hogy ez az érték nagyobb lehet 25,6-nál, amiből arra

következtethetünk, hogy az eltérés valamilyen szisztematikus dologgal magyarázható, ami a testsúlyra fejt ki kedvezőtlen hatást¹⁷ (pl.: egészségtelen táplálkozás túl nagy volumene, vagy a mozgásszegény életmód).

A gyakorlati esetek során legtöbbször nem áll módunkban nagy elemű minta segítségével a hipotéziseinket ellenőrizni, hanem kis mintával kell dolgoznunk. Kis minta esetén a standard normális eloszlás nem alkalmazható, ilyenkor a *Student-féle t-eloszlást* és ennek az eloszlásnak a táblázatát kell alkalmaznunk. A t-eloszlás alkalmazása során figyelembe kell venni az ún. *szabadságfokot*, amely a minta elemszámának 1-gyel csökkentett értéke. Egy adott rendszer szabadságfokán azt a számot értjük, ahány érték ebben a rendszerben szabadon megválasztható (t- és χ^2 – eloszlás esetén egy, F-eloszlásnál két szabadságfokot határozunk meg).

$$t = \frac{26,99 - 25,6}{\frac{6,08}{\sqrt{70}}} = 1,91$$

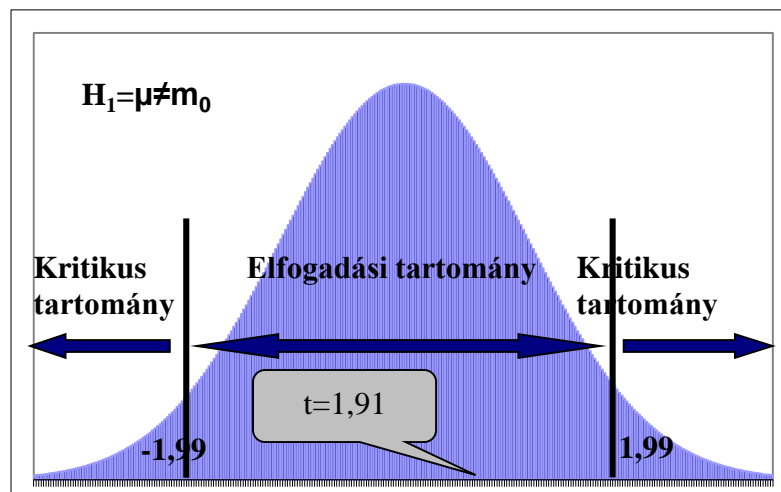
A t-eloszlás kritikus értéke 5 %-os szignifikancia szinten 69 szabadságfok (n-1) mellett a t-eloszlást tartalmazó táblázatból olvasható:1,99.

Mivel a számított érték kisebb, mint a táblabeli, ezért elfogadjuk a nullhipotézist (nincs okunk rá, hogy elutasítsuk), tehát elutasítjuk az alternatív hipotézist (a grafikus ábrán az elfogadási tartományban jelenik meg). Ez azt jelenti, hogy a mintabeli és az elvárt érték közötti eltérést – 5 %-os szignifikancia szinten – a véletlen okozhatta.

Ha csak arra lennénk kíváncsiak, hogy az általunk mért érték megegyezik-e vagy sem a svéd kismamák értékével, akkor kétoldalú hipotézis-ellenőrzést hajtanánk végre. Ilyenkor az alternatív hipotézis a következő lesz:

$$H_1: \mu \neq 25,6$$

¹⁷ Ezt a feltételezést szakértői véleményünkre alapozzuk, melynek irodalmi kutatásától jelen esetben eltekintünk.



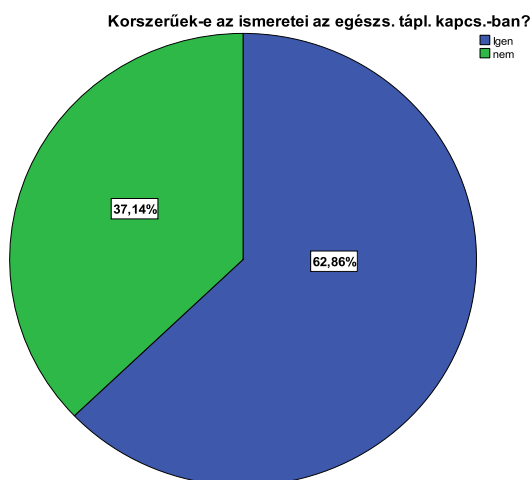
9/8. ábra: A döntést segítő grafikus ábra kétoldali alternatív hipotézis esetén

A fenti alternatív hipotézis esetén a sűrűségfüggvény mindkét oldalát figyelembe kell venni, így a kritikus érték (5%-os szignifikancia szinten): $\pm 1,99$. Ehhez viszonyítva is el kell fogadni a 25,6-os BMI-re vonatkozó hipotézist.

Hasonlóan kell eljárunk, ha nem az alapsokasági átlagra, hanem az alapsokasági arányra vonatkozóan fogalmazunk meg feltevést. Itt is meg kell jegyeznünk, hogy csak nagy minta esetén használható a tesztelésre a standard normális eloszlás.

$$z = \frac{p - P_0}{\sqrt{\frac{P_0(1 - P_0)}{n}}}$$

Gyakorlati példaként vizsgáljuk meg, hogy elvárható-e a kismamák körében, hogy a 60%-uk korszerű táplálkozási ismerettel rendelkezzen. Először a feltett kérdésre (Korszerűek-e az ismeretei az egészséges táplálkozással kapcsolatban?) történő válaszok megoszlását kell vizsgálnunk. A válaszok megoszlását a következő ábra mutatja:



9/9. ábra: A válaszok megoszlásának grafikus ábrája

A kismamák közül 44 igennel, 26 nemmel válaszolt a kérdésre. Vizsgáljuk meg elvárható-e a kismamák között, hogy az egészséges táplálkozásról megfelelő ismerettel rendelkezők aránya 60%-os legyen.

$$H_0: P=0,6$$

$$H_1: P<0,6$$

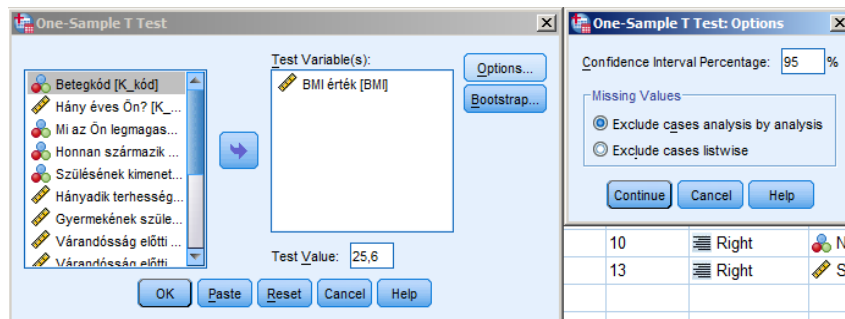
Az alternatív hipotézisben azt a feltevést fogalmaztuk meg, amely szerint 60%-nál kisebb a megfelelő ismeretekkel rendelkezők aránya.

$$p = \frac{44}{70} = 0,628$$

$$u = \frac{0,628 - 0,6}{\sqrt{\frac{0,6 \times (1 - 0,6)}{70}}} = 0,93$$

A táblabeli érték (a negatív oldalt figyelembe véve) -1,645. Az empirikus érték az elfogadási tartományba esik, tehát elfogadjuk a nullhipotézist, vagyis a megfelelő, korszerű táplálkozási ismeretekkel rendelkező kismamák aránya 60%-nál magasabb lesz.

Az SPSS programmal viszonylag egyszerűbben jutunk eredményhez. Az *ANALYZE* menü, *COMPARE MEANS* almenü, *ONE- SAMPLE T TEST* modulja segítségével.



9/10. ábra: Az egymintás t- próba az SPSS programmal

Először válasszuk ki a BMI érték változónkat, majd a *TEST VALUE* dobozba írjuk a nullhipotézisben megadott értéket (25,6), az összehasonlítás alapjául szolgáló konstans értéket. Ezután az *OPTIONS* dobozba megadhatjuk a szignifikancia szintet (95%). Ezt követően az OK gomb megnyomásával jutunk el a végeredményhez az *OUTPUT VIEW*-ban:

A végeredmény két táblából áll. Az elsőben a leíró statisztikai adatokat láthatóak.

9/6. táblázat: A leíró statisztika adatai

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
BMI érték	70	26,9891	6,08032	,72674

Forrás: saját szerkesztés

A második táblázat a hipotézis eldöntését segíti.

9/7. táblázat: A t-érték és szignifikanciát tartalmazó táblázat

One-Sample Test

	Test Value = 25.6					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
BMI érték	1,911	69	,060	1,38907	-,0607	2,8389

Forrás: saját szerkesztés

Látható, a számított t-érték, a szabadságfok, viszont nem látjuk a kritikus értéket. A döntést azonban a következő adat szolgáltatja, amely azt mutatja, hogy a t számított értéke milyen szinten szignifikáns. Általában a Sig=5% (0,05) alatt a nullhipotézist elutasítjuk. A nullhipotézisünk igaz, hiszen a valószínűség magasabb, minden 100 esetből, több mint 6 eset, vagyis a nullhipotézist elfogadjuk. A konfidencia-intervallum azokat a határértékeket mutatja, melyek közé a követelménytől vett eltérések értékei 95%-os valószínűséggel esnek. Itt szeretnénk megjegyezni, hogy a szignifikancia érték a határérték közeli, ilyenkor az általánosságok megfogalmazásával óvatosan kell bánni.

Gyakorlatban nagyon sokszor fordul elő, hogy két különböző sokaságból veszünk véletlen és független mintát. Ilyen esetekben a két sokaság ugyanazon paramétereit hasonlítjuk össze, teszteljük különbségeiket, azonosságukat. A gyakorlati alkalmazások során számtalanszor találkozunk a két alapsokasági várható érték egyezőségének, minta alapján történő tesztelésével. Ahogyan már eddig is megszokhattuk, az állítást általánosságban nullhipotézisben, konkrét formában az alternatív hipotézisben található.

$$H_0: \mu_1 - \mu_2 = \delta$$

Az alternatív hipotézis különböző módon történő felírása lehetővé teszi a várható értékek nagyságrendi relációiról történő döntést is, vagyis:

$$H_1: \mu_1 - \mu_2 < \delta \quad ; \quad H_1: \mu_1 < \mu_2 \text{ (bal oldali)}$$

$$H_1: \mu_1 - \mu_2 > \delta \quad ; \quad H_1: \mu_1 > \mu_2 \text{ (jobb oldali)}$$

$$H_1: \mu_1 - \mu_2 \neq \delta \quad ; \quad H_1: \mu_1 \neq \mu_2 \text{ (két oldali)}$$

Most nézzük a legelterjedtebb **kétmintás t-próbát**, melynek két előfeltétele van: mindkét sokaság eloszlása legyen normális (külső, egyéb információ szükséges), illetve az alapsokasági szórásnégyzetek legyenek egyenlők.

Ha normális eloszlású alapsokaságból vett kis mintánk van, és a sokasági szórás ismeretlen, de feltételezzük az azonosságot, akkor egyszerűen t-próbát használunk (Nem követünk el nagy hibát azonban, ha nagyobb minta esetén is ezt a próbát¹⁸ használjuk):

Az alkalmazható próbafüggvény:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

A szabadságfok: $n_1 + n_2 - 2$

Ebben az ún. közös szórás (s_p) négyzetének képlete:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 1}$$

Amennyiben az alapsokasági szórásnégyzeteket nem ismerjük, akkor teszteljük. Ismeretes, hogy a mintabeli korrigált varianciák hányadosa - valószínűségi változó -, ha az alapsokasági varianciák egyenlő F-eloszlást követ, $n_1-1; n_2-1$ szabadságfok- párral.

$$F = \frac{s_1^2}{s_2^2} \Big|_{H_0} \approx F_{n_1-1; n_2-1}$$

A két várható érték különbözőségét a következő példán keresztül értelmezzük. Az adatbázisunkból ismerjük a testtömeg index lakhely szerinti leíró statisztikai értékeit. A városban élő kismamák (35 fő) BMI értékének átlaga 24,14, szórása 4,38, míg a falun élők (35 fő) indexének átlaga 29,84 kg, szórása 6,25.

Gyakorló feladat: 5%-os szignifikancia szinten van-e meghatározó különbség a két mintát alkotó kismamák átlagos testtömeg-indexét tekintve?

Mivel kétmintás próbáról van szó, először megvizsgáljuk, hogy a szórások egyenlőnek tekinthetők-e, illetve teszteljük a normalitást (lásd később).

Ezután következik a kétmintás t-próba:

$$H_0: \mu^1 = \mu^2$$

$$H_1: \mu^1 \neq \mu^2$$

¹⁸Amennyiben a t-eloszlás nagyobb szabadságfokú értékeit a standard normális eloszlás hasonló adataival összevetjük, szembevető a hasonlóság.

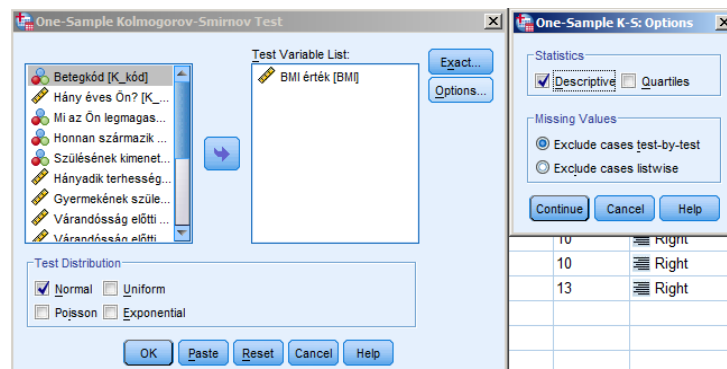
$$s_p^2 = \frac{(35 - 1) \times 4,38^2 + (35 - 1) \times 6,25^2}{35 + 35 - 2} = 29,12$$

$$s_p = \sqrt{29,12} = 1,29$$

$$t = \frac{24,14 - 29,84}{5,396 \times \sqrt{\frac{1}{35} + \frac{1}{35}}} = -4,41$$

A t-eloszlás táblabeli értéke 68-mas szabadságfok esetén, mivel kétoldali a hipotézisünk, 1,99. A számított érték az elutasítási tartományba esik, tehát 5-os szignifikancia szint mellett a két lakhelyen élő kismamák BMI indexe szignifikáns különbséget mutat ($p < 0,05$).

Az SPSS program segítségével a kutatónak kényelmes dolga van, hiszen könnyen számítható az eljárás. Az előfeltételek közül első lépésben vizsgáljuk meg, hogy a BMI index normális eloszlást követ-e. Ennek eldöntésére a Kolmogorov-Szmirnov próba áll rendelkezésünkre, melyet a korábbiakban tárgyaltunk, itt most részletes ismertetésétől eltekintünk. Az eljárást az *ANALYZE* menü *LEGACY DIALOGS* almenüjének *1-SAMPLE K-S* opciójából indíthatjuk. A *TEST VARIABLES* dobozba a BMI érték változót mozgassuk, majd az *OPTIONS* gombra kattintva kérjük a leíró statisztikát (*DESCRIPTIVE*). Ezt követően a *CONTINUE* és *OK* gombok lenyomása következik.



9/11. ábra: A BMI változó normalitás-vizsgálatának beállításai

A program két táblázatot közöl velünk, melyből a lényegi információt számunkra a második tartalmazza.

9/8. táblázat: A BMI normalitás vizsgálata

One-Sample Kolmogorov-Smirnov Test

			BMI érték
N			70
	a,b	Mean	26,9891
		Std. Deviation	6,08032
Most Extreme Differences		Absolute	,084
		Positive	,084
		Negative	-,062
Kolmogorov-Smirnov Z			,707
Asymp. Sig. (2-tailed)			,700

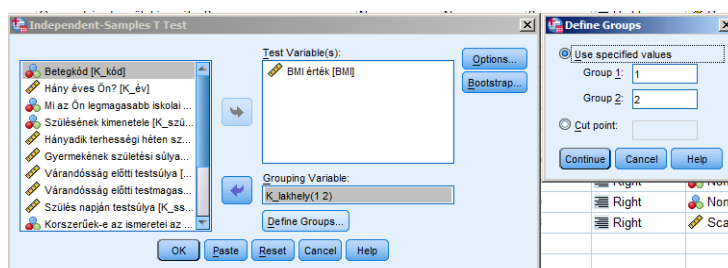
Test distribution is Normal.

Calculated from data.

Forrás: saját szerkesztés

A táblázat alján található szignifikancia érték ($p=0,7$) nyomán megállapítható, hogy a BMI index a normális eloszlást követi ($p<0,05$), vagyis a paraméteres eljárásokat lehet alkalmazni.

Az *ANALYZE* főmenü, *COMPARE MEANS* almenü, *INDEPENDENT-SAMPLES T TEST* modulban hajtható végre a művelet. Az első lépésben a tesztváltozó dobozba a BMI értékek kerüljenek, majd a csoportosító változónak megadjuk lakhely változónkat. Itt jelölni kell az ismérvváltozatokat, vagyis hogy milyen kódokkal jelöltük a két csoportot.



9/12 .ábra: A kétmintás- t próba beállításai

A beállításokat követően a következő végeredményhez jutunk, melynek első táblázata ismét a leíró statisztikai adatokat tartalmazza:

9/9. táblázat: A leíró statisztikai adatok

Group Statistics					
		N	Mean	Std. Deviation	Std. Error Mean
BMI érték	város	35	24,1393	4,38158	,74062
	falú	35	29,8389	6,25372	1,05707

Forrás: saját szerkesztés

A kétmintás t-próba előfeltétele a szórások azonossága, melyet az SPSS a Levene (Levin) teszttel ellenőriz. A sajátos F-próbaként felfogható teszt értékelése megegyezik az eddig tárgyaltakkal, hiszen itt is a nullhipotézis arra vonatkozik, hogy a szórások azonosak. Mivel a megfigyelt szignifikanciaszint értéke nagyobb, mint 0,05, ezért a szórások azonosságára vonatkozó nullhipotézist elfogadjuk, ilyenkor a táblázat felső sorát kell nézni. Ha elvetnénk a nullhipotézist (a szórások nem egyenlők), akkor a második sort lenne szükséges elemezni. A program kétszélű próbát végez, alternatív hipotézist ellenőriz.

9/10. táblázat

		Independent Samples Test								
		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
BM érték	Equal variances assumed	1,804	,184	-4,416	68	,000	-5,69963	1,29071	-8,27519	-3,12407
	Equal variances not assumed			-4,416	60,899	,000	-5,69963	1,29071	-8,28064	-3,11862

Forrás: saját szerkesztés

A további eredmények megegyeznek a már tárgyaltakkal, tehát a nullhipotézist elvethetjük. A t értéknél szereplő negatív előjel arra enged következtetni, hogy a vizsgálatba szereplő első ismérvváltozat átlaga az alacsonyabb. Ez azt jelenti, hogy szignifikáns eltérés van a lakhelyek vonatkozásában a vizsgált BMI változóban.

A normalitás sérülése esetén nem tudjuk a kétmintás t-próbát alkalmazni, ilyenkor nem-paraméteres vizsgálatok közül a Mann-Whitney próbát szükséges alkalmazni.

Gyakorlati feladatként, kíváncsiak vagyunk van-e különbség lakhely szerint a születések időpontjában, vagyis meghatározza-e a szülés idejét (hányadik héten szül) a lakóhely.

A normalitásvizsgálat arról tanúskodik, hogy a szülési hét változó nem követi a normális eloszlást ($p < 0,05$).

9/11. táblázat: A születési időpont (szülés hete) változó normalitás vizsgálata

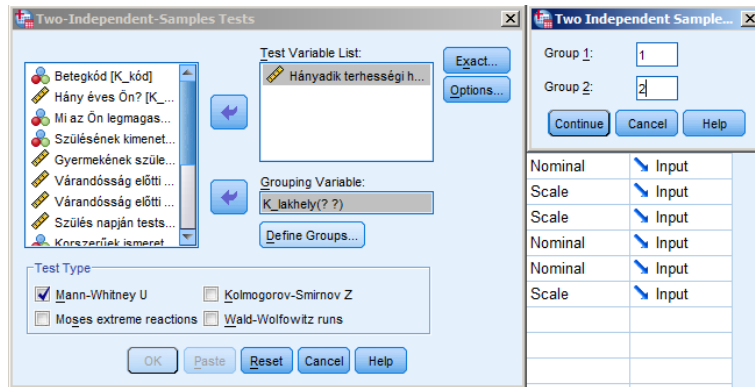
One-Sample Kolmogorov-Smirnov Test			Hányadik terheségi héten szült?
N			70
	a,b	Mean	38,2714
		Std. Deviation	1,49319
Most Extreme Differences		Absolute	,187
		Positive	,113
		Negative	-,187
Kolmogorov-Smirnov Z			1,566
Asymp. Sig. (2-tailed)			,015

Test distribution is Normal.

Calculated from data.

Forrás: saját szerkesztés

Az előfeltétel nem teljesülése esetén a nem-paraméteres vizsgálatot végzünk. A kétmintás t-próba nem paraméteres párja a Mann-Whitney teszt, melyet az *ANALYZE* menü *LEGACY DIALOGS* almenüjének *2 INDEPENDENT SAMPLES* opciójából indíthatjuk.



9/13. ábra: a Mann-Whitney próba beállításai

A *TEST VARIABLES* dobozba a szülési időpontra vonatkozó változót mozgassuk (Hányadik terhességi héten szült), majd *GROUPING VARIABLE* dobozba a lakhely változó kerüljön. A csoportosító változónál fontos kalibrálni az ismérvváltozatok kódjait. Ezt követően *OPTIONS* gombra kattintva kérjük a leíró statisztikát (*DESCRIPTIVE*), majd a *CONTINUE* és *OK* gombok lenyomása következik.

9/11. táblázat: A Mann-Whitney próba eredményei

	Hányadik terhességi héten szült?
Mann-Whitney U	529,500
Wilcoxon W	1159,500
Z	-1,000
Asymp. Sig. (2-tailed)	,318

Grouping Variable: Honnan származik Ön?

Forrás: saját szerkesztés

A táblázatban látható szignifikancia alapján ($p=0,318$) elmondható, a szülési hét változó vonatkozásában a városban és falun élő kismamák között nem volt szignifikáns különbség, ez azt jelenti, hogy a különbség a véletlen műve.

A következőkben a kettőnél több sokaság esetén alkalmazható módszerek közül először a varianciaanalízis módszerét mutatjuk be. A módszer segítségével megkíséreljük egy vagy több minőségi ismerv alapján képzett részmintákban a kiválasztott mennyiségi ismerv szerinti különbözőségét számszerűsíteni. Az eljárás nemzetközileg használt rövidítése: ANOVA.

Szokás még „egyutas”, „One-way ANOVA”-nak is nevezni. Az eljárás hazai pontos elnevezése: **egyszempontos varianciaanalízis**. A varianciaanalízis (Analysis Of Variance=ANOVA) célja az átlagok összehasonlítása, viszont eszköze a varianciák vizsgálata. A varianciaanalízis feltételezi, az alapsokaságon és valamennyi csoporton (részsokaságon) belül a mennyiségi ismérv normális eloszlást követ. A másik tesztelendő előfeltétel: a varianciahomogenitás, vagyis a csoportok szórásai egyenlők (homoszkedasztikus) legyenek.

A statisztikai összehasonlításoknál is megkülönböztethetjük a függő és független változókat. A független változó mindig a csoportosítás szempontja – akár szerepel ez külön csoportosítási változóként/faktorként az adatbázisunkban, akár nem. A független változó értékeit itt tehát maguk a minták jelentik. A mintáktól függő változó pedig a vizsgált folytonos változó, amelynek az átlagait hasonlítjuk össze. Mindezeket azért lehetnek fontosak, mert a statisztikai programok használatakor a felnyíló ablakok kérhetnek ilyen változó kijelölést. Abban az esetekben, ha a kutatási adatbázisunk még nem tartalmazza „csoportosítási” változót, akkor létre kell hoznunk azt, generálnunk kell.

A varianciaanalízisnek többszempontos változatai is léteznek. Ezek tárgyalása meghaladja a jelen kiadvány kereteit, bővebben Pintér – Rappai 2007-es könyvében olvashat az érdeklődő.

Az egyszempontos varianciaanalízis alkalmazásának három legtipikusabb esete:

- kettőnél több (rész) sokaság várható értékének egyezőségére vonatkozó hipotézis ellenőrzése;
- homogenitás-vizsgálat;
- vegyes kapcsolat (kvalitatív és kvantitatív változó közötti kapcsolat) szignifikáns voltának tesztelése.

A varianciaanalízis modellje:

$$x_{ji} = \mu + \tau_j + \varepsilon_{ji}$$

ahol a j-edik csoport i-edik eleme (x_{ji}) , a teljes sokaságra vonatkozó várható érték (μ) , a j-edik osztály csoporthatása (τ_j) és az ε_{ji} véletlen hatás összegeként adódik. A vizsgálat során a következő hipotézisrendszert teszteljük:

$$H_0 : \mu_1 = \mu_2 = \dots \mu_m = \mu$$

$$H_1 : \mu_j \neq \mu$$

A nullhipotézis elfogadása a várható értékek egyezőségének, a részekre bontott sokaság homogenitásának, valamint a vegyes kapcsolat hiányának (függetlenség) kimondását jelenti. A csoportosított sokaságra vonatkoztatva, egy adott mintáról elmondható, hogy háromféle átlagtól vett eltérés számítható, mely az alábbi összefüggésből keletkezik:

$$\sum \sum (x_{ij} - \mu)^2 = \sum n_j (\mu_j - \mu)^2 + \sum \sum (x_{ij} - \mu_j)^2$$

ahol a képlet a teljes eltérés- négyzetösszeget felbontja külső (csoportok közötti), illetve belső (csoportokon belüli) eltérés- négyzetösszegekre.

9/12. táblázat

Eltérés- négyzetösszeg	Eltérés- négyzetösszeg típusa
$SS = \sum_{j=1}^m \sum_{i=1}^{n_j} (x_{ji} - \bar{x})^2$	Teljes eltérés- négyzetösszeg
$SS_K = \sum_{j=1}^m (\bar{x}_j - \bar{x})^2$	Külső eltérés- négyzetösszeg
$SS_B = \sum_{j=1}^m \sum_{i=1}^{n_j} (x_{ji} - \bar{x}_j)^2$	Belső eltérés- négyzetösszeg

Forrás: saját szerkesztés

Az eltérésnégyzet összegekből képezhető próbafüggvény F eloszlást követ, ahol a számláló szabadságfoka m-1 (m a csoportok száma), a nevező szabadságfoka n-m (n a sokaság tagszáma). A próbafüggvény, egyoldalú nagyobb alternatív hipotézist feltételezve alkalmas a varianciaanalízis végrehajtására, vagyis ha F számított értéke nagyobb, mint a kritikus érték, akkor a nullhipotézist elvetjük.

$$F = \frac{\frac{SS_K}{m-1}}{\frac{SS_B}{n-m}}$$

A döntést segítő összefüggéseket a következő táblában olvashatjuk:

9/13. táblázat

Tényezők	Eltérésnégyzet- összeg (SS)	Szabad- sádfok (df)	Átlagos eltérés- négyzet- összeg(MS)	F
Csoportok között	$SS_K = \sum_{j=1}^m n_j (\bar{x}_j - \bar{x})^2$	m-1	$\frac{SS_K}{m-1}$	$\frac{MS_K}{MS_B}$
Csoporton belül	$SS_B = \sum_{j=1}^m \sum_{i=1}^{n_j} (x_{ji} - \bar{x}_j)^2$	n-m	$\frac{SS_B}{n-m}$	
Összesen	$SS = \sum_{j=1}^m \sum_{i=1}^{n_j} (x_{ji} - \bar{x})^2$	n-1	$\frac{SS}{n-1}$	

Forrás: saját szerkesztés

Legyen a példafeladatunk a következő:

Gyakorlati feladatként, határozzuk meg, vajon eltérőek-e a terhesség előtti különböző BMI kategóriába (sovány, normál testalkatú, túlsúlyos, erősen túlsúlyos) tartozó kismamák megszületett gyermekeinek az átlagos születési súlya, vagyis azonosnak (homogénnek) tekinthetők-e a különböző BMI kategóriákba tartozó kismamák gyermekeinek az átlagos születési súlya (gramm)?

Első lépésben a folytonos változó eloszlását kell ismét tesztelnünk, hiszen a normalitásvizsgálat a varianciaanalízis egyik előfeltétele. A normalitásvizsgálat alkalmas arra, hogy két valószínűségi változó eloszlását összehasonlítsuk, vagy ellenőrizzük, hogy egy valószínűségi változónak csakugyan az az eloszlása, amit feltételeztünk. Jelen esetben arra vagyunk kíváncsiak, hogy a születési testsúly változó normális eloszlásból származik-e. A változó kiugró/*outlier* értékeinek ellenőrzését követően vizsgáljuk a normalitást. Amennyiben a számított szignifikancia érték magasabb, mint 5%, akkor elmondható, hogy a változó normál eloszlást követ. A felhasznált adatbázis elemszáma 70 fő, így a nem paraméteres eljárások közül ismét Kolmogorov- Szmirnov próbát alkalmazzuk a normalitás tesztelésére. A megfelelő beállítások elvégzése után a következő eredményekhez jutunk:

9/14. táblázat

	N	Mean	Std. Deviation	Minimum	Maximum
Gyermekének születési súlya?	70	3329,7143	547,43929	2150,00	4200,00

		Gyermekének születési súlya?
N		70
Normal Parameters ^{a,b}	Mean	3329,7143
	Std. Deviation	547,43929
Most Extreme Differences	Absolute	,091
	Positive	,069
	Negative	-,091
Kolmogorov-Smirnov Z		,758
Asymp. Sig. (2-tailed)		,614

a. Test distribution is Normal.

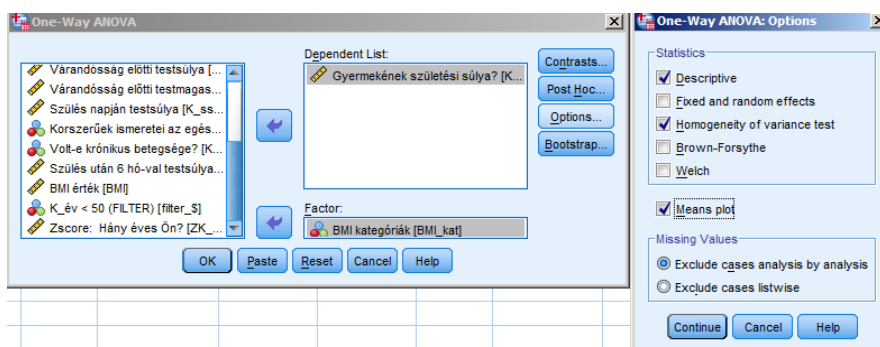
b. Calculated from data.

Forrás: saját szerkesztés

A Kolmogorov-Szmirnov Z értéke, valamint a hozzá tartozó szignifikancia szint alapján ($p=0,614$) megállapítható, hogy a változó normál eloszlást követ, így a varianciaanalízis, mint paraméteres próba elvégezhető. Abban az esetben, ha a Z értékhez tartozó szignifikancia kisebb, mint 5%, akkor nem paraméteres Kruskal-Wallis próbát kell végrehajtani (lásd később). Amennyiben problémát jelent a normalitásvizsgálat, akkor tekintjük úgy, mintha nem normál eloszlású lenne a mintánk, mert ha normál eloszlású adatokon nem parametrikus tesztet végzünk, gyakorlatilag a parametrikus teszttel azonos eredményt kapunk, míg fordított esetben ez nem áll fent!

Az elemzést az *ANALYZE* menü, *COMPARE MEANS* almenüjének, *ONE-WAY ANOVA* moduljával indíthatjuk.

Először az elemzésben megjelenő változókat jelöljük. A függő változó (*DEPENDENT LIST*) a gyermek születési súlya változó lesz, míg a független változó (*FACTOR*), vagyis csoportosító változó a BMI kategória lesz.



9/14. képernyőnét: A varianciaanalízis beállításai az SPSS programban

Ezt követően a *OPTIONS* panelen jelöljük be a *DESCRIPTIVE*, a *HOMOGENEITY OF VARIANCE*, illetve a grafikus megjelenítést biztosító *MEANS PLOT* lehetőségeket. A *HOMOGENEITY OF VARIANCE* (varianciahomogenitás) opciót mindig használjuk, hiszen így tudjuk tesztelni a második előfeltételt, a varianciák egyezőségét. A kellő beállításokat követve (*CONTINUE* majd *OK*) eljutunk a következő eredményekhez.

A beállítások nyomán először a leíró statisztikát tartalmazó (elemszám, átlag, szórás, korrigált szórás, konfidencia intervallum alsó és felső értékei, minimum és maximum érték) táblázatot kapjuk meg:

9/15. táblázat

Descriptives

Gyermekeinek születési súlya?								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
sovány	5	2840,0000	713,61754	319,13947	1953,9268	3726,0732	2300,00	4050,00
normál testalkatú	23	3276,9565	515,78215	107,54801	3053,9156	3499,9974	2200,00	4080,00
túlsúlyos	21	3221,4286	529,69128	115,58811	2980,3160	3462,5412	2150,00	4100,00
erősen túlsúlyos	21	3612,3810	446,66436	97,47015	3409,0618	3815,7001	2910,00	4200,00
Total	70	3329,7143	547,43929	65,43151	3199,1820	3460,2466	2150,00	4200,00

Forrás: saját szerkesztés

A táblázatból látható, hogy a mintában 5 fő volt, aki a sovány kategóriába tartozik. A terhesség előtt erősen túlsúlyos BMI kategóriában lévő kismamák gyermekeinek átlagos születési súlya 3612,38 gramm, melyhez 547,44 szórás érték tartozott. Az is látható, hogy a születési súly 95%-os konfidencia intervalluma 3409,06 és 3815,70 gramm közé esik.

A következő táblázat a szórás-homogenitást vizsgálja az úgynevezett Levene-teszt segítségével.

9/16. táblázat

Test of Homogeneity of Variances

Gyermekeinek születési súlya?

Levene Statistic	df1	df2	Sig.
,210	3	66	,889

Forrás: saját szerkesztés

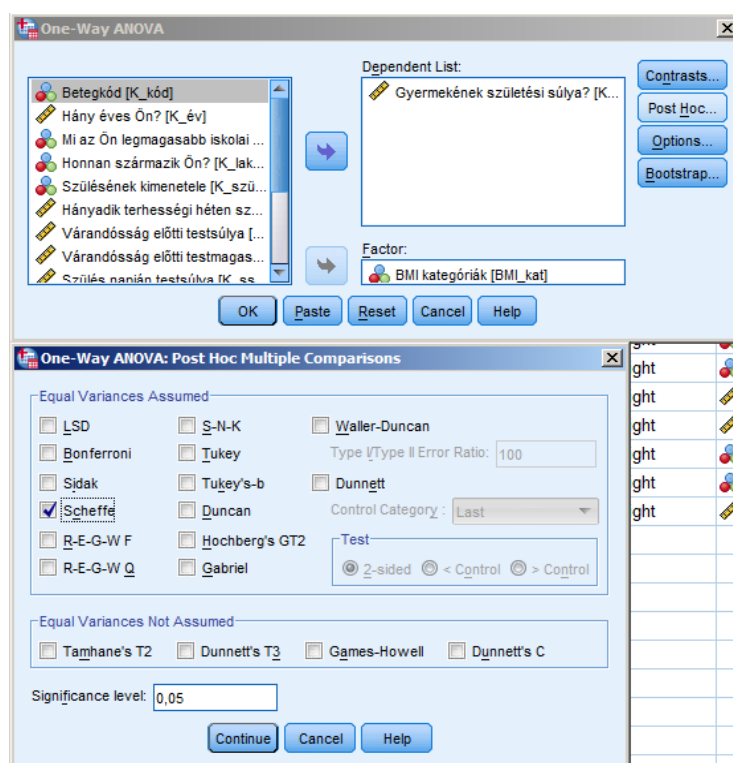
Amennyiben a teszt szignifikancia szintje alacsony, mint 0,05, akkor el kell utasítani a szórások azonosságára vonatkozó hipotézist, ha azonban magasabb, mint 0,05, akkor fennáll a szórások azonossága. Amennyiben a szórások egyezősége nem állna fenn, úgy az Options panelből a Brown-Forsythe, valamint a Welch próbákat kellene alkalmazni, mivel az F-próba nem szolgáltatna releváns eredményt. Jelen esetben ($p=0,889$) a szórásokat azonosnak tekinthetjük. Ezt követően a varianciaanalízis táblázata következik.

9/17. táblázat ANOVA

Gyermekének születési súlya?					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3187269,234	3	1062423,078	4,009	,011
Within Groups	17491325,05	66	265020,077		
Total	20678594,29	69			

Forrás: saját szerkesztés

A táblázat első oszlopában láthatóak a csoportok közötti, csoportokon belüli és a teljes eltérés négyzetösszegek. A második oszlop a szabadságfokokat tartalmazza, amelyekkel az eltérésnégyzet-összegeket elosztva a csoportok közötti és csoportokon belüli eltérésnégyzet-összegeket kapjuk. A F értékét megkapjuk ($F=4,01$), ha a csoportok közötti és csoportokon belüli eltérésnégyzeteket egymáshoz hasonlítjuk ($1062423,08/265020,08$). Itt is látható, hogy a szignifikancia szint kisebb, mint 0,05, ezért a nullhipotézist elutasítjuk, vagyis a terhesség előtti BMI értékhez tartozó kismamák babáinak születési súlya különbözik egymástól. A különbözőség tényét konstatálva, lehetőségünk van az egyes kategóriák átlagai közötti különbségek vizsgálatára, vagyis arra, hogy mely kategóriák különböznek egymástól. Ennek vizsgálatát segíti a Post Hoc (utólagos) modul, melynek feltétele, hogy legalább három kategóriánk legyen.



9/15. ábra: A Post Hoc tesztek beállítási modulja

A modult a *ONE-WAY ANOVA* panelből érhetjük el, ha ott a *POST HOC* lehetőségre klikkelünk. A csoportok közötti különbségek tesztelésére több utóteszt is létezik. Nincs egyetlen általánosan elfogadott eljárás, amit mindenki használ. A post-hoc tesztek elsősorban aszerint vannak csoportosítva, hogy a szóráségyezés feltétele teljesül-e. A próba kiválasztásánál két fontos szempontot kell figyelembe vennünk: mennyire könnyen lehet vele különbséget kimutatni (mennyire engedékeny), illetve mennyire megbízható. A Post Hoc tesztek első csoportja az egyenlő szórásnégyzeteknél alkalmazható tesztek tartalmazza.

A *Post Hoc tesztek* közül bemutatunk pár gyakran alkalmazott analízist. A szórások egyezőségekor gyakran használják a Bonferroni és Scheffe tesztek. A Bonferroni teszt páronkénti átlagok különbségének vizsgálatára használható, amikor a két csoport elemszáma különböző is lehet. Lényege, hogy az α -hibához tartozó t-értéket korrigálja a független összehasonlítások számának megfelelően. A Bonferroni teszt statisztikája:

$$L = t(\text{táblázatbeli}) \sqrt{S_p^2 \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

A *Scheffe-teszt* a hagyományos tesztek közé tartozik. Ez már valóban a H_0 hipotéziseket vizsgálja. Az egyszerű F-próba akkor utasítja el a H_0 -hipotézist, ha létezik egy $a > 0$ vektor, amelynél a konfidencia-intervallum nem tartalmazza a 0-t. Ha k darab összehasonlítandó csoport van, akkor $k(k-1)/2$ összehasonlítást kell végezni. A statisztikája:

$$L = \sqrt{s_p^2 (k-1) F_{(táblázatbeli)} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

A *Dunnnett-teszt* egy kijelölt csoportot (kontroll) hasonlít össze a többivel. Eredetileg egyenlő elemszámokra volt érvényes, de később elkészült az általánosítása egyenlőtlen elemszámokra is. Lényegét tekintve páronkénti összehasonlítást végez szimultán, de meg kell adni egy kezdő, kontroll csoportot, és ehhez hasonlítja a többi csoport átlagát. Dunnnett-tesztet az *ANALYZE, COMPARE MEANS, ONE-WAY ANOVA, POST HOC* parancsok után érhetjük el. A teszt alkalmazása előtt ki kell választani a kontroll csoportot (*CONTROL CATEGORY*). A párbeszéd ablakból csak az első vagy utolsó csoportot tudjuk kiválasztani a legördülő listából. Továbbá meg kell adni, hogy az összehasonlítás egyoldalú vagy kétoldalú legyen. Alapbeállításként kétoldalú szimmetrikus összehasonlítás történik. Ebben az esetben nincs semmiféle előzetes információnk az összehasonlítandó párokról, bármelyik csoport lehet nagyobb, vagy kisebb, mint a kontroll. Egyoldalú próba esetében előzetesen már van információnk arról, hogy az összehasonlítandó csoport vagy csak nagyobb, vagy csak kisebb

lehet, mint a kontroll csoport. Amennyiben nincs információnk a csoportok közötti relációról, mindig a kétoldalú próbát használjuk.

Statisztikája:

$$\bar{x}_i - \bar{x}_o \pm |d|s_p \sqrt{\frac{2}{n}}$$

$$\bar{x}_o = \textit{kontroll csoport}$$

Amennyiben a szórásnégyzetek különböznek Tamhane-tesztet, illetve Dunnett's T3 próbákat alkalmazhatunk.

Az adatbázisunk segítségével a Scheffe post-hoc próba eredményeinek értelmezése következik. A post-hoc beállítások közül jelen esetben a Scheffe-próbát választottuk, melynek legfontosabb adatait a következő output tábla mutatja.

9/18. táblázat: A varianciaanalízis post hoc tesztjeinek táblázata
Post Hoc Tests

Multiple Comparisons

Dependent Variable: Gyermekeinek születési súlya?
Scheffe

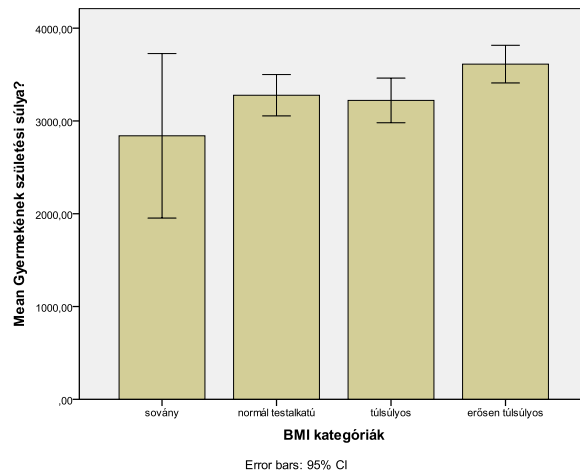
(I) BMI kategóriák	(J) BMI kategóriák	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
sovány	normál testalkatú	-436,95652	254,02092	,405	-1165,7413	291,8282
	túlsúlyos	-381,42857	256,17185	,533	-1116,3843	353,5272
	erősen túlsúlyos	-772,38095*	256,17185	,035	-1507,3367	-37,4252
normál testalkatú	sovány	436,95652	254,02092	,405	-291,8282	1165,7413
	túlsúlyos	55,52795	155,37894	,988	-390,2535	501,3094
	erősen túlsúlyos	-335,42443	155,37894	,209	-781,2058	110,3570
túlsúlyos	sovány	381,42857	256,17185	,533	-353,5272	1116,3843
	normál testalkatú	-55,52795	155,37894	,988	-501,3094	390,2535
	erősen túlsúlyos	-390,95238	158,87104	,120	-846,7526	64,8478
erősen túlsúlyos	sovány	772,38095*	256,17185	,035	37,4252	1507,3367
	normál testalkatú	335,42443	155,37894	,209	-110,3570	781,2058
	túlsúlyos	390,95238	158,87104	,120	-64,8478	846,7526

*. The mean difference is significant at the 0.05 level.

Forrás: saját szerkesztés

Az első oszlopban láthatók a viszonyítás alapját (BMI kategóriák I), a második oszlopban a viszonyítás tárgyát képező ismérvváltozók (BMI kategóriák J). A Scheffe post hoc analízis különbséget jelez a sovány és az erősen túlsúlyos kategória között ($p < 0,05$). A felhasználó segítségére van és a szignifikáns eltéréseket jelzi a harmadik oszlopban található (Mean Difference I-J) számított értékek mögött található apró csillag.

Az eredmények prezentációjánál jelen esetben az oszlopdiaagram a javasolt megjelenítési forma, melyre érdemes egy konfidencia intervallumot illeszteni. A beállításokat a korábbiakban tárgyaltuk, így most a végeredményként megjelenő grafikus ábrát mutatjuk be.



9/15. ábra: A gyermekek születési súlya az édesanyák terhesség előtti BMI kategóriáinak vonatkozásában

Abban az esetben, ha az előfeltételek során a normalitás vizsgálat nem teljesül, vagyis a folytonos változó nem normál eloszlásból származik, a fent bemutatott paraméteres próbát (varianciaanalízist) nem alkalmazhatjuk. Amennyiben a mintánk nem az ismert normális eloszlásból származik és háromnál több csoportot tartalmaz akkor a *Kruskal-Wallis* nem paraméteres próbát alkalmazzuk. A nem paraméteres eljárásoknak a közös tulajdonsága, hogy nem tételezik fel, hogy az adatok egy adott populáció egy specifikus eloszlásához illeszkednek, szemben a paraméteres módszerekkel, melyek esetében fontos előfeltétel, hogy eloszlásuk a módszerben feltételezett tulajdonságokkal rendelkezzen. Ezért szokták ezeket a módszereket összevontan eloszlásmentes módszereknek is nevezni.

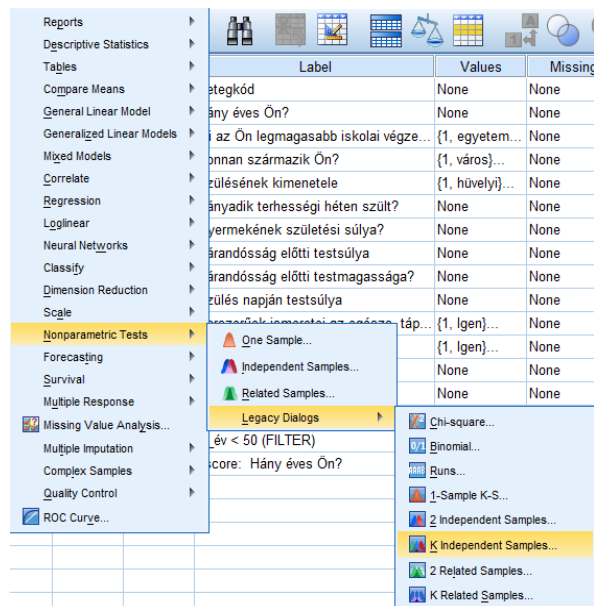
Az egyszempontos varianciaanalízis nem paraméteres megfelelő módszereként általában a *Kruskal-Wallis* próbát szokták emlegetni. A *Kruskal-Wallis* próba végrehajtása hasonló a *Mann-Whitney U* próbához, sőt, két független csoport esetén mindkét módszer azonos eredményt adhat. Gyakorlatban mondható, hogy a *Kruskal-Wallis* próba, a *Mann-Whitney* teszt általánosítása három vagy több független minta esetén. A *Kruskal-Wallis* próba a mintákat egyesíti, kiszámítja a rangokat, majd csoportonként átlagolja. Ha a mediánok egyenlők, akkor a rangok átlagai nem térnek el lényegesen egymástól.

Legyen a példafeladatunk a következő:

Gyakorlati feladatként, határozzuk meg, vajon eltérőek-e a terhesség előtti különböző BMI kategóriába (sovány, normál testalkatú, túlsúlyos, erősen túlsúlyos) tartozó kismamák szülési időpontjai (születési hét), vagyis azonosnak (homogénnek) tekinthetők-e a különböző BMI kategóriákba tartozó kismamák szülési időpontjai?

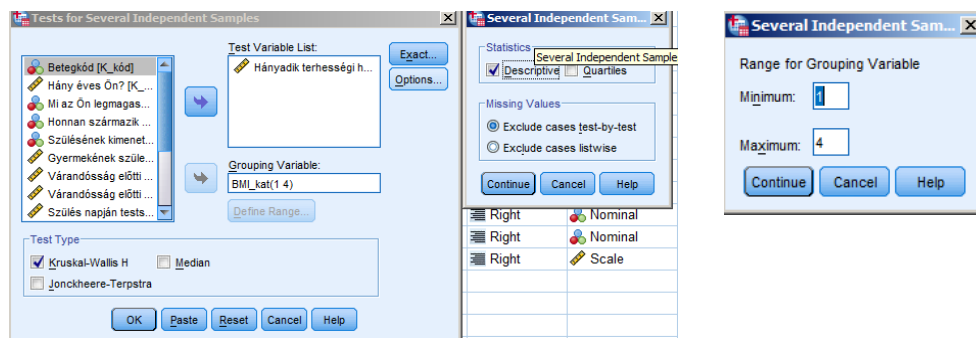
Első lépésben a szülési változó normalitás vizsgálatát végezzük el, melyet már a korábbiakban megtettünk. A Kolmogorov-Szmirnov próba segítségével számított szignifikancia ($p=0,015$) alapján látszik, hogy a változó nem követi a normális eloszlást.

A Kolmogorov-Szmirnov próba beállításai, így a nem paraméteres próbák közül a Kruskal-Wallis tesztet kell alkalmazni. A statisztikai próba melyet az *ANALYZE* menü *LEGACY DIALOGS* almenüjének *K INDEPENDENT SAMPLES* opciójából indíthatunk.



9/16. ábra: A Kruskal-Wallis próba beállításai

Ezt követően a célváltozóként a szülés hetének változóját (hányadik terhességi héten szült) az válasszuk, míg a csoportosító változó jelen a terhesség előtt BMI kategória legyen. Általános beállításaként kérjük az *OPTION* modulból a leíró statisztika (*DESCRIPTIVE*) opciót. Ezt követően a csoportosító változónál meg kell jelölni az ismérvváltozatok számát, ami a felhasznált adatbázisunkban négy, azaz négy kategóriába soroltuk a BMI kategóriákat.



9/17. ábra: A Kruskal- Wallis próba változóinak beállításai

A lefutott próba eredményei alapján látható, hogy szignifikáns eltérés nem tapasztalható a szülési hét tekintetében az egyes BMI kategóriák között.

Az első táblázat a leíró statisztikai eredményeket mutatja. Itt leolvasható az egyes változók elemszámai, átlagai, szórásai, valamint a legkisebb és legnagyobb értékei.

9/19. táblázat: a leíró statisztikai eredményei

Descriptive Statistics					
	N	Mean	Std. Deviation	Minimum	Maximum
Hányadik terhességi héten szült?	70	38,2714	1,49319	34,00	41,00
BMI kategóriák	70	2,8286	,94748	1,00	4,00

Forrás: saját szerkesztés

A második táblázat az egyes ismérvváltozók elemszámait, valamint a rendezett rangszám értékeit tartalmazza.

9/20. táblázat: az ismérvváltozók elemszáma, illetve rangszám értékeinek táblázata

Ranks			
	BMI kategóriák	N	Mean Rank
Hányadik terhességi héten szült?	sovány	5	35,80
	normál testalkatú	23	26,59
	túlsúlyos	21	39,86
	erősen túlsúlyos	21	40,83
	Total	70	

Forrás: saját szerkesztés

A harmadik táblázat hordozza számunkra a legfontosabb információkat, hiszen itt látható a szignifikancia érték, mely alapján eldönthető, hogy az egyes kategóriák (ismérvváltozatok) közötti különbség létező-e.

9/21. táblázat

a,b

	Hányadik terhességi héten szült?
Chi-Square	7,167
df	3
Asymp. Sig.	,067

Kruskal Wallis Test

Grouping Variable:
BMI kategóriák

Forrás: saját szerkesztés

10. BEVEZETÉS A REGRESSZIÓ ANALÍZISBE (Ács Pongrác)

10.1. Kétfváltozós lineáris regresszió

A mennyiségi ismérvek közötti összefüggések vizsgálatára a korrelációs számítás mellett a leggyakrabban alkalmazott statisztikai módszer a regresszió számítás. A regresszió számítás a jelenségek tendenciáit vizsgálja, megpróbálja a kapcsolat természetét valamilyen jól megfogható és értelmezhető függvény formájában megragadni. Ezeket a függvényeket regresszió-függvényeknek nevezzük. A regresszió számítás bemutatását először az alapesetnél, a kétfváltozós lineáris regresszióval tesszük meg. Jelen esetben a vizsgálni kívánt változó, az eredményváltozó alakulása (növekedése vagy csökkenése) sztochasztikusan függ az egyetlen magyarázó, függő változótól. Fontos megjegyezni, hogy a változók közötti kapcsolat a gyakorlatban sokszor nem lineáris. Ilyen esetben mind a szorosság mérésének, mind a kapcsolat törvényszerűségét felíró matematikai modellnek a felépítése viszonylag bonyolult matematikai-statisztikai eljárásokat igényel, melyekkel most a jelen elektronikus tankönyvben nem foglalkozunk. Amennyiben a változók között lineáris sztochasztikus kapcsolatot tételezünk fel, egy viszonylag egyszerű matematikai modellel, egy lineáris függvény (egyenes) paramétereinek meghatározásával jól felhasználható regresszió-függvényt számszerűsíthetünk: $\hat{Y} = b_0 + b_1 X$

Az egyenes konstans paramétereinek becslését az ún. legkisebb négyzetek módszerével (LNM)¹⁹ végezhetjük el.

A két paraméter az alábbi képletekkel határozható meg:

$$b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$
$$b_0 = \bar{Y} - b_1 \bar{X} = \frac{\sum Y}{n} - \frac{b_1 \sum X}{n}$$

A gyakorlatban a tényezőváltozó paraméterének (b_1) különösen kiemelt a szerepe, regressziós együtthatónak nevezi a statisztika, míg a b_0 paramétert tengelymetszetnek vagy konstansnak hívják. A regressziós együttható a tényezőváltozó egységnyi növekedését kísérő várható eredményváltozó változást számszerűsíti az eredetileg megadott mértékegységben, vagyis a magyarázó változó értékének egy egységnyi változása az eredményváltozót b_1

¹⁹A becslési módszer elvi leírásától tananyagunkban eltekintünk, csupán a módszer alkalmazásával nyert megoldó-képleteket használjuk fel.

egységnyivel változtatja. A regressziós együttható ismerete lehetővé teszi, hogy lineáris összefüggés esetén is kvantifikáljuk az elaszticitást (a rugalmasságot), amely a változás relatív (százalékban kifejezett) mértékét fejezi ki. A mutató választ ad, arra, hogy az X magyarázó változó 1%-os változásához az Y eredményváltozó hány %-os változása járul várhatóan. Az átlagos elaszticitás a változók átlagai segítségével az alábbi módon határozható meg:

$$El = b_1 \frac{\bar{x}}{\bar{y}}$$

A regresszió-analízis módszertanával kapcsolatosan számtalan irodalom áll rendelkezésre, melyek közül további ismeretekhez juthat az olvasó Pintér – Rappai (2007), Mundruczó (1981), Ramanathan (2003) könyveiben.

A következőkben gyakorlati példán szeretnénk bemutatni először a kétváltozós lineáris regressziót. A következőkben egy új adatbázis felhasználásával²⁰ kívánjuk bemutatni a módszert, mely adatbázis természetesen más feladatokhoz gyakorlasként is jól felhasználható. Az adatbázist Kehl Dániel a Pécsi Tudományegyetem Közgazdaságtudományi Karának adjunktusa állította össze és a Sporttudományi kutatások módszertana című könyvben is már szemléltető adatbázisként szerepelt. Természetesen a következő illusztratív és gyakorlatorientált példafeladatok teljesen újak, de az adatbázis felhasználásával számtalan lehetőség áll a gyakorolni vágyó hallgató rendelkezésére.

Az adatbázist úgy jött létre, hogy abban a világ legjelentősebb motor márkái, illetve a márkák legnépszerűbb típusai szerepelnek. A végleges adatbázis ötvenhárom motort tartalmaz. Az adatok szekunder módon a „Motor katalógus” 2003-as kiadványából lettek kigyűjtve.

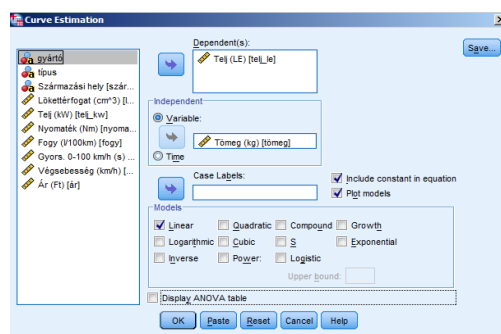
²⁰ Az adatbázis (motor.sav) elérhető a Pécsi Tudományegyetem Egészségtudományi Kar honlapján, illetve a Sporttudományi kutatások módszertana című tankönyv DVD mellékletén. Az adatbázis felhasználásával sokváltozós módszerek illusztrálása is megtörtént Ozsváth- Ács elektronikus tankönyvében.

Az egyes motorokról a következő adatok állnak rendelkezésre:

1. gyártó: általában egy gyártónak több terméke is bekerült az adatbázisba,
2. típus: az adott gyártók által használt típusjelzés,
3. származás: a gyártó cég nemzetisége,
4. lökettérfogat (cm^3)
5. teljesítmény (KW)
6. teljesítmény (lőerő, LE)
7. tömeg (kg)
8. fogyasztás (l/100km)
9. gyorsulás (0-100 km/h (s))
10. végsebesség (km/h)
11. ár (Ft).

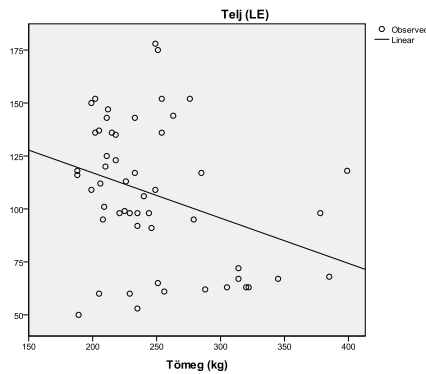
Gyakorlati feladat segítségével kívánjuk a módszert szemléltetni! Vizsgáljuk meg a teljesítmény (LE) és a tömeg (kg) kapcsolatrendszerét. A példánkban két folytonos, metrikus változónk van, melyek közül a függő (*DEPENDENT*) változó a motor teljesítménye (LE), míg a független változó a motor tömege lesz. Ez azt jelenti, hogy a tömeg függvényének hatására fogjuk vizsgálni a teljesítményt (LE).

A regresszió-analízis vizsgálatának első lépései között a kutatók gyakran az alapadatokat pontdiagram segítségével jelenítik meg, melyből láthatóvá válhat a kapcsolat típusa és iránya. Az SPSS program *ANALYZE* menüpontjának *REGRESSION* almenüjének *CURVE ESTIMATION* opciójával tudjuk megtenni, ahol a pontdiagramon kívül grafikus illeszkedésvizsgálatot is megjeleníthetünk.



10/1. ábra: A regresszió grafikus illeszkedésének vizsgálatának beállításai

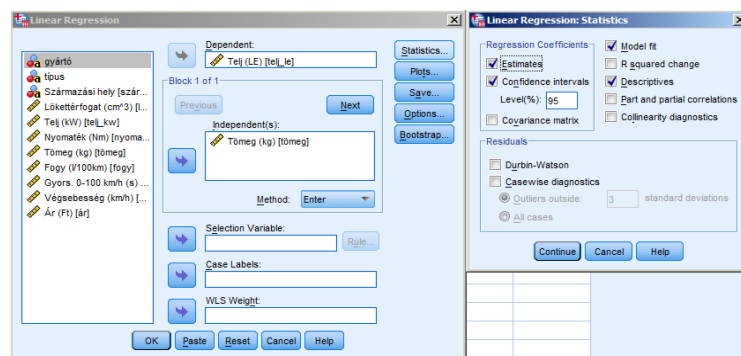
A függő változó (*DEPENDENT(s)*) dobozban a teljesítmény változót mozgattuk, míg a független változóként (*INDEPENDENT*) a tömeget választottuk. Ezt követően az *OK* gomb megnyomásával jutunk el a grafikus ábrához.



10/2. ábra: A regresszió pontdiagram ábrája

A két változó kapcsolatáról a grafikus ábra alapján azt lehet megállapítani, hogy negatív irányultságú, de ezen felül kvantitatív eredményeket nem láthatunk.

Ezt követően számszerűsítsük és értelmezzük a regressziós egyenes értékeit, melyet az *ANALYZE* menüpontjának *REGRESSION* almenüjének *LINEAR* opciójával tehetünk meg. A beállításokat a következő képernyőnézet szemlélteti.



10/3. ábra: a lineáris regresszió beállításai

A függő változónak a teljesítményt (LE), míg független változónak a tömeget (kg) jelöljük, majd a *STATISTICS* gomb lenyomását követően az alapbeállításokhoz válasszuk még a *CONFIDENCE INTERVALS*, *MODEL FIT* és a *DESCRIPTIVES* lehetőségeket. A *CONTINUE* és *OK* gombok választását követően az számítási eredményeket az *OUTPUT VIEW* nézetben olvashatjuk.

10/1. táblázat: A regressziós modell összefoglaló adatai

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,327 ^a	,107	,089	31,932

a. Predictors: (Constant), Tömeg (kg)

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6207,500	1	6207,500	6,088	,017 ^b
	Residual	52003,254	51	1019,672		
	Total	58210,755	52			

a. Dependent Variable: Telj (LE)

b. Predictors: (Constant), Tömeg (kg)

Forrás: saját szerkesztés

Az első táblázatban a PEARSON féle korrelációs együtthatót láthatjuk, mely alapján közepes erősségű kapcsolatról beszélhetünk ($r=0,327$). Ezt követően látható a kapcsolat erejét számszerűsítő determinációs együttható ($r^2=0,107$), mely azt jelenti teljes szórás közel 11%-át tudja a független változó magyarázni, vagyis a motor teljesítményének változásában a tömeg csupán 11%-ban játszik szerepet. Fontos megemlíteni, hogy minél magasabb az r^2 , annál jobban illeszkedik az egyenesünk a ponthalmazra. A grafikus ábrán szerkeszthető és kérhető a determinációs együttható megjelenítése. Ezt követi a becslés standard hibája, mely az előrejelzés pontosságának proxy mutatója (minél magasabb, annál kevésbé képes a modell jól becsülni). Ezt követően a varianciaanalízisnél megismert ANOVA táblázat következik, mely az F-próba értékét és a kapcsolat meglétét igazoló ($p<0,05$) szignifikancia értéket tartalmazza. A táblázat alapján megállapítható, hogy a két változó közötti kapcsolat létezik és nem a véletlen műve. Ezt követően a regressziós egyenes paramétereit is tartalmazó koefficiens táblázat következik.

10/2. táblázat: A regressziós egyenes paramétereit

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	159,824	21,967		7,276	,000	115,725	203,924
	Tömeg (kg)	-,214	,087	-,327	-2,467	,017	-,388	-,040

a. Dependent Variable: Telj (LE)

Forrás: saját szerkesztés

A modell értelmezése előtt látható, hogy mindkét változóhoz tartozó t-érték, illetve az ezekhez tartozó szignifikancia értékek ($p<0,05$) is a modell meglétét igazolja. A paraméterbecslés rész végén lehetőség nyílik az általunk beállított megbízhatósági szinten a paraméterekre való intervallumbecslés értelmezésére is.

A nem standardizált koefficiensek segítségével leolvasható a regressziós egyenes egyenlete:

$$b_1 = -0,21$$

$$b_0 = 159,82$$

A regresszió-függvény:

$$\hat{Y} = 159,82 - 0,21 X$$

A regressziós együttható ismeretében elmondható, hogy a teljesítmény egységnyi növekedése átlagosan várhatóan 0,21 kg-mal csökkentheti a tömeget.

A modell segítségével megbecsülhetjük egy 150 kg tömegű motor teljesítményét (LE):

$$\hat{Y} = 159,82 - 0,21 \times 150 = 128,32$$

10.2. Többváltozós lineáris regresszió

Szintén gyakran előforduló módszer több független változós lineáris regresszió-analízis. Az előzőekben tárgyalásra került a kétváltozós lineáris regresszió, azonban meg kell jegyeznünk, hogy az életben viszonylag kevés alkalommal tudunk egy jelenséget két változóval leírni és magyarázni, legtöbbször a változóink száma meghaladja a kettőt.

Ismételjük át, hogy mire is használható a regresszió analízis:

- Annak meghatározására, hogy a független változók hatással vannak-e a függő változóra: Létezik-e az összefüggés?
- Annak meghatározására, hogy a független változók milyen mértékben magyarázzák a függő változó ingadozását: A kapcsolat szorossága.
- A kapcsolat formájának és matematikai struktúrájának meghatározása.
- Előrejelzésre: függő változó értékeinek az előrejelzése.

A lineáris regressziós analízis módszerét tehát kiterjeszthetjük két vagy több független változóra. Ezt nevezzük többszörös lineáris regressziónak. A **többszörös lineáris regresszió** egy függő változó (Y) és kettő vagy több független változó (magyarázó változó) (X_1, X_2, \dots, X_i) közötti kapcsolat leírására szolgáló statisztikai eljárás. Hasonlóan a lineáris regresszióhoz az eljárás választ próbál adni arra, hogy a független változók egységnyi változása, a függő változó milyen mérvű megváltozását vonhatja maga után.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_i X_i + \varepsilon$$

- Y : függő változó
- $X_1, X_2, X_3, \dots, X_i$: független változók avagy magyarázó változók
- i : a magyarázó változók száma
- β_0 : (más jelöléssel α) konstans, állandó érték
 - Megadja a regressziós egyenes és a koordináta-rendszer függőleges (y) tengelyének metszéspontját.
- $\beta_1, \beta_2, \beta_3, \dots, \beta_i$: konstans regressziós együtthatók
 - Megadja a regressziós egyenes meredekségét.
 - Grafikusan jelzi, hogy a független változók egységnyi változása várhatóan milyen mértékben változtatja meg a függő változót.
- ε : hibatenyező, hibatag

Feltételei:

- a sokaság elemei függetlenek legyenek egymástól
- csak normális eloszlású sokaságok hasonlíthatók össze
- a sokaságok szórásai a mintán belül egyformák
- A magyarázó változók egymástól nem függenek (nincs köztük multikollinearitás).

Multikollinearitásnak nevezzük a tényezőváltozók közötti lineáris kapcsolatot. Ellenőrzése többféleképpen történhet:

- Többszörös determinációs együtthatóval
- F-próbával ($F > F_{krit}$)
- VIF-mutató

Jelen esetben a VIF mutató számszerűsítését mutatjuk be.

VIF mutató:

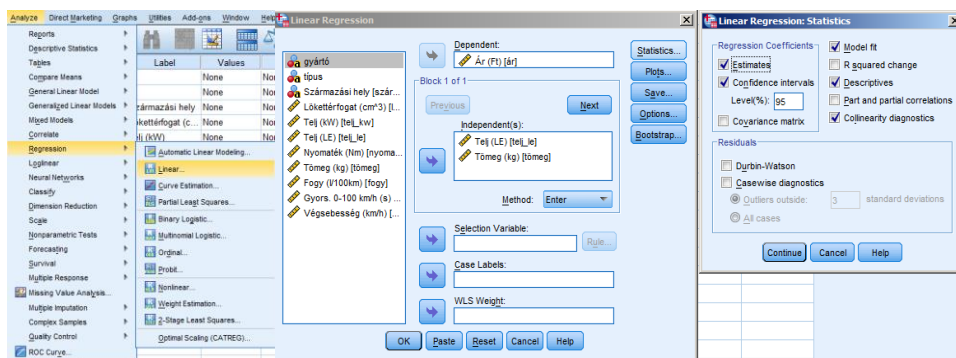
- Variancianövelő tényező $VIF_j = \frac{1}{1 - R_j^2}$
- $1 < VIF \leq \infty$
- $VIF=1$ ha $R_{j2}=0$ (amikor a j. magyarázó változó nem korrelál a többi magyarázó változóval)
- $VIF \rightarrow \infty \Rightarrow R_{j2}=1$ (a j. magyarázó változó pontosan kifejezhető a többi lineáris kombinációjaként)
 - $1 < VIF \leq 2$ - gyenge multikollinearitás
 - $2 < VIF \leq 5$ - erős zavaró multikollinearitás
 - $5 < VIF$ - nagyon erős, káros multikollinearitás

Multikollinearitás kizárása: két független változó közti korrelációs együttható nem haladhatja meg a 0,7-es, míg a determinációs együttható: 0,5-ös értéket.

Gyakorlati feladat segítségével vizsgáljuk meg, hogy motorok teljesítménye és tömegének nagysága kapcsolatban van-e a motorok árával. Az motorok ára forintban értendő. A vizsgálatba bevont független változóink a teljesítmény (LE) és a tömeg (kg) voltak.

Végezzük el a beállításokat:

A statisztikai próbát ismét az *ANALYZE* menüpont *REGRESSION* moduljának *LINEAR* eljárásával érjük el. (10/4. ábra). A függő változónk legyen az ár, míg a független változók a teljesítmény és a tömeg. Ezt követően az *STATISTICS* menüpontban kérjük a kijelölt beállításokat (*ESTIMATES*, *CONFIDENCE INTERVALS*, *MODEL FIT*, *DESCRIPTIVES*, *COLLINEARITY DIAGNOSTICS*).



10/4. ábra: Többváltozós lineáris regresszió beállításai

A *CONTINUE* és *OK* panelek lenyomását követően az alábbi eredményekhez jutunk:

Az első táblázat a leíró statisztikát tartalmazza a három vizsgált változó tekintetében. Megállapítható, hogy a motorok átlagos ára: 3 440 618 forint, az átlagos teljesítmény 106 (LE), és a motorok átlag tömege 250,5 kg. A táblázat a számtani átlag mellett a szórás értékeket, valamint az elemszámot tartalmazza.

10/3. táblázat

Descriptive Statistics

	Mean	Std. Deviation	N
Ár (Ft)	3440618,00	1343598,867	50
Telj (LE)	106,16	34,164	50
Tömeg (kg)	250,54	51,765	50

Forrás: saját szerkesztés

A következő táblázat egy korrelációs mátrix, mely a kapcsolat-szorossági mérőszámokat, valamint a szignifikancia értékeket tartalmazza. Fontos kitétel, hogy a változók közötti szignifikáns kapcsolatban a korrelációs együtthatók értéke sehol nem haladja meg a 0,7-es értéket.

A következő táblázatban a modell magyarázó erejére vonatkozó adatokat (többszörös korrelációs együttható, determinációs együttható, korrigált R^2 érték, valamint a regressziós modell standard hibája) láthatjuk. Látható, hogy a független változók a függő változó, közel 50%-át (48) magyarázza, vagyis a motorok árát 48%-ban magyarázzák.

10/5. táblázat

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,693 ^a	,480	,458	989521,817

a. Predictors: (Constant), Tömeg (kg), Telj (LE)

Forrás: saját szerkesztés

A következő táblázat a modell tesztelésére szolgáló ANOVA értékeket tartalmazza. A számított értékekből látszik, hogy a regressziós modell elégedettség szóródását ($p < 0,05$), vagyis alkalmas a függő változó becslésére. A táblázat utolsó oszlopa szolgáltatja a lényegi információt, mely szerint elvetjük a nullhipotézist, vagyis a modell alkalmas a függő változó magyarázatára.

10/6. táblázat

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4,244E+13	2	2,122E+13	21,670	,000 ^b
	Residual	4,602E+13	47	9,792E+11		
	Total	8,846E+13	49			

a. Dependent Variable: Ár (Ft)

b. Predictors: (Constant), Tömeg (kg), Telj (LE)

Forrás: saját szerkesztés

Az utolsó táblázatban szereplő t-próbához tartozó értékek alapján látható, hogy a modellben szerepeltetett változók megfelelőek ($p < 0,05$), vagyis a lineáris kapcsolat van. Az oszlop utolsó oszlopában szereplő VIF mutató a multikollinearitás vizsgálatára szolgál, és azt mutatja, hogy a modellben a multikollinearitás nem zavaró.

10/7. táblázat

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	-2267145,374	984129,116		-2,304	,026	-4246957,788	-287332,960		
	Telj (LE)	9004,631	4365,278	,229	2,063	,045	222,824	17786,438	,898	1,113
	Tömeg (kg)	18966,360	2880,989	,731	6,583	,000	13170,557	24762,162	,898	1,113

a. Dependent Variable: Ár (Ft)

Forrás: saját szerkesztés

Az utolsó táblázatban szereplő B értékek alapján felírható az egyenlet:

$$\text{ár} = -2267145,37 + 9004,63 \cdot \text{teljesítmény (LE)} + 18966,36 \cdot \text{tömeg (kg)}$$

Az együtthatók alapján a változók közötti kapcsolatokról azt tudjuk elmondani, hogy a teljesítmény és tömeg is hatással van az elégedettségre, hatásuk iránya egyező. A teljesítmény és a tömeg növekedésével nő a motorok ára. Ezt követően becsülhetjük a 136 lóerős, 215 kg tömegű motor árát. $(3035252 = -2267145,37 + 9004,63 \cdot 136(\text{LE}) + 18966,36 \cdot 215(\text{kg}))$

12. IRODALOMJEGYZÉK

1. Ács P. (2009) Sporttudományi kutatások módszertana. Pécsi Tudományegyetem Természettudományi Kar Testnevelés- és Sporttudományi Intézet, Pécs
2. Alexin Z. – Lelovics Zs. (2009): Etikai kérdések a beavatkozással nem járó humán orvosi kutatásokban. Orvosi Hetilap 150. évfolyam 37. szám 1749-1752.o.
3. Babbie E. (2004): A társadalomtudományi kutatás gyakorlata. Balassi Kiadó, Budapest
4. Bajsz V. (2013) Egy multinacionális cég egészségfelmérése. Szakdolgozat. PTE ETK, Pécs
5. Bakonyi G. – Kokas K. (2006): Bevezetés a könyvtári informatika alapjaiba. Jatepress, Szeged 63-76. o.
6. Barton D. – Pál V. (2013): Open Access és osztályozás. Könyvtár figyelő. 59. évfolyam 4. szám 700–709. o.
7. Bauer András – Berács József – Kenesei Zsófia (2007): Marketing alapismeretek, Aula Kiadó, Budapest, o. 124
8. Berhidi A. (2008): Az orvosi szakirodalom adattárai 2. A tudomány hálójában: Web of Science. Nőgyógyászati Onkológia 13. évfolyam 3. szám 154-158.o.
9. Berhidi A. (2008): Az orvosi szakirodalom adattárai 3. Scopus: Elsevier „madárka” a tudomány „fészkében”. Nőgyógyászati Onkológia 13 évfolyam 4. szám 186–190.o.
10. Bertáné Németh Á. – Farkas Fné. (2013): A PhD. disszertációval szemben támasztott formai követelmények. Online elérhető: <http://ktk.pte.hu/sites/default/files/mellekletek/2014/04/disszertacio.pdf> (2014-08-10.)
11. Betlehem J. – Boncz I. – Oláh A. (2010): Tudományos közlések az egészségtudományban. Nővér 23. évfolyam 6. szám 4-11.o.
12. Bíbor M. – Gulyás B. – Földes Zs. – Hegyi Á. – Kiss Farkas G. – Lacházi Gy. – Orosz A. – Parádi A. (2005): A magyar irodalom filológiája. Online elérhető: <http://www.tankonyvtar.hu/hu/tartalom/tkt/magyar-irodalom/ch07s11.html> (2014-08-01)
13. Bíróiné Nagy E. – Bognár J. – Farkas J. – Gombocz J. – Hamar P. – Kovács A.T. – Mészáros J. – Ozsváth K. – Rétsági E. – Rigler E. – Salvára I. M. – Szabó B. – Tihanyiné Hős Á. – Vináné Kokovay Á. (2011): Sportpedagógia – Kézikönyv a testnevelés és sport pedagógiai kérdéseinek tanulmányozásához. Online elérhető:

http://www.tankonyvtar.hu/en/tartalom/tamop425/0025_Birone_Nagy_Edit-Sportpedagogia/ch01s07.html (2014-08-24)

14. Boncz I. – Döbrössy L.-Péntek Z. – Kovács A. – Budai A. – Imre L. – Vajda R. – Sebestyén A. (2013): A szervezett országos emlőszűrési program negyedik (2008-2009) szűrési körének részvételi arányai. Orvosi Hetilap 154. évfolyam 50. szám 1975-1983.o
15. Csajbók E. (2009): Az orvosi szakirodalom adattárai 4. Az orvostudomány határán. Nőgyógyászati Onkológia 14. évfolyam 3. szám 133-137. o.
16. Cserné Adermann G. (1999): A tanulás- és kutatómódszertan alapjai. JPTE-FEEFI, Pécs.
17. Demsey P.A. – Dempsey A.D. (1999): Kutatómunka az ápolásban. Medicina Könyvkiadó Rt, Budapest, 37-45.o.
18. Drótos L. – Martin Rebecca A. (2012): Értéknövelt szolgáltatás az olvasóknak: ingyenes és nyílt hozzáférésű források megtalálása. Tudományos és műszaki tájékoztatás 59. évfolyam 10. szám 437-439.o.
19. Martin Rebecca A. (2012): Értéknövelt szolgáltatás az olvasóknak: ingyenes és nyílt hozzáférésű források megtalálása. Tudományos és műszaki tájékoztatás 59. évfolyam 10. szám 437-439.o.
20. Falus I. (szerk) (2004): Bevezetés a pedagógiai kutatás módszereibe. Műszaki Könyvkiadó, Budapest
21. Falus I. (szerk.) (2000): Bevezetés a Pedagógiai kutatás módszereibe. Pedagógus Könyvek. Budapest. Műszaki Könyvkiadó. 540. o.
22. Farkas Livia (2012): A Google továbbra is a barátod: kérdőívek készítése
23. Fésüs L. (2014): Tudományetikai kihívások és válaszok hazánkban és Európában. Magyar Tudomány 175. évfolyam 6. szám 645-650. o.
24. Fidy J. – Makara G. (2005): Biostatisztika. Online elérhető: <http://www.tankonyvtar.hu/hu/tartalom/tkt/biostatisztika-1/ch11.html> (2014-06-30)
25. Gőcze I. (2011): A tudományos kutatás módszerei. Hadtudományi Szemle 4. évfolyam 3. szám 157-166. o.
26. Hajdu O. (1987): Sokváltozós statisztikai módszerek gyakorlati alkalmazása. Prodinform Műszaki Tanácsadó Vállalat. Budapest
27. Héra G. – Ligeti Gy. (2006): Módszertan. Bevezetés a társadalmi jelenségek kutatásába. Osiris Kiadó, Budapest.

28. Hornyacsek J. (2013): A tudományos kutatás elméleti és gyakorlati kérdései 2. (A tudományos kutatás folyamata). Online elérhető: http://hhk.uni-nke.hu/downloads/kiadvanyok/mkk.uni-nke.hu/kulonszam2013julius/eloadasokpdf/02hornyacsek_tudkut_final%20x.pdf (2014-08-24)
29. Hunyadi L. (2001): Statisztikai következtetésemélet közgazdászoknak. KSH, Budapest
30. Hunyadi L. (2002): Grafikus ábrázolás a statisztikában, Statisztikai Szemle 2002/1. 22-53. o.
31. Istvánfi Cs. (2000): Gondolatok a sporttudományokról. Kalokagathia 1-2 szám 7-18. o.
32. Jackson P.A. – Reay J.L. – Scholey A.B. – Kennedy D.O. (2012): DHA-rich oil modulates the cerebral haemodynamic response to cognitive tasks in healthy young adults: a near IR spectroscopy pilot study. British Journal of Nutrition 107. évfolyam 8. szám
33. Jánosa A. (2005): Adatelemzés számítógéppel, Perfekt Kiadó. Budapest, 271. o.
34. Jobbágy Á. – Csordás P. – Mersich A. – Lupkovics G. – Sztaniszláv Á. (2008): Vérnyomás otthoni monitorozása. IME 7. évfolyam 10. szám 36-40.o.
35. Karamánné Pakai A. – Peterka G. – Dér A. – Bujtor A. – Dancsné Balogh K. – Czömpöl O. (2006): Egy kifejlesztés alatt álló otthon monitorozó készülék iránti igények felmérése Zala Megyében. Nővér, 19. évfolyam 5. szám 30-37. o.
36. Kecskeméty L. – Izsó L. (2005): Bevezetés az SPSS programrendszerbe, ELTE- Eötvös Kiadó, Budapest, 460. o.
37. Kehl D. – Rappai G. (2006): Mintaelem-szám tervezése Likert-skálát alkalmazó lekérdezésekben. Statisztikai Szemle 84. évfolyam 9. szám 848- 876. o.
38. Kerlinger F. (1980): Analysis Of Covariance Structure Tests Of A Criterial Referents Theory Of Attitudes. Multivariate Behavioral Research 15 évfolyam 4. szám 403- 422. o.
39. Kerpel-Fronius S. (2008): A nürnbergi orvosper. A kényszereutanázia-program örök érvényű társadalmi tanulságai. LAM 18. évfolyam 2. szám 94-96. o.
40. Kopp M. – Kovács M.E. (2006): A magyar népesség életminősége az ezredfordulón. Semmelweis Kiadó, Budapest
41. Kovács J. (2007): Bioetikai kérdések a pszichiátriában és a pszichoterápiában. Medicina, Budapest
42. Lám J. – Balázsné Szelei E. – Rózsa E. – Bódi M. (2011): Gyógyszerosztással összefüggő gyógyszerelési hibák direkt megfigyeléses vizsgálatának szervezése, és a

- vizsgálat legfontosabb tanulságai. Semmelweis Egyetem, Egészségügyi Menedzserképző Központ. Online elérhető: http://semmelweis.hu/emk/files/2012/02/7et_gyogyszereles.pdf (2014-08-25.)
43. Lampek, K. – Kívés, Zs. (2012): Kutatásmódszertani és biostatistikai ismeretek (pp.177-208). In: Oláh A. (szerk.) Az ápolástudomány tankönyve. Medicina Könyvkiadó, Budapest, 177-210. o.
 44. Lógó Emma (2007): Kérdőíves technikák, módszerek (előadás)
 45. Majoros P. (2004): A kutatómódszertan alapjai: Tanácsok, tippek, trükkök - nem csak szakdolgozat-íróknak. Perfekt, Budapest
 46. Móczár Cs. – Borda F. – Faragó K. – Borgulya G. – Brauniczer F. – Vörös V. (2007): Egészséges életmód hatása túlsúlyos és elhízott betegeken. Orvosi Hetilap 148. évfolyam 2. szám 65–69. o.
 47. Mozaffarian D. – Hao T. – Rimm E.B. – Willett W.C. – Hu F.B. (2011): Changes in Diet and Lifestyle and Long-Term Weight Gain in Women and Men. The New England Journal of Medicine 364. évfolyam 25. szám 2392-2404. o.
 48. Mundruczkó Gy. (1981): Alkalmazott regressziószámítás. Akadémiai Kiadó, Budapest
 49. Oláh A. (szerk.) (2008): Az ápolástudomány tankönyve. Medicina Könyvkiadó, Budapest
 50. Online elérhető: <http://erg.bme.hu/oktatas/teir/GT52AT02/kerdoiv.pdf> (2014-07-15)
 51. On-line elérhető:
<http://www.informaworld.com/smpp/title~content=t775653673~db=all~tab=issueslist~branches=15 - v15> (2014-07-15)
 52. Online elérhető: <http://www.urban-eve.hu/2012/01/28/a-google-tovabbra-is-a-baratod-kerdoivek-keszítése/> (2014-07-12)
 53. Ozsváth Károly, Ács Pongrác (2011): Bevezetés a sporttudományos kutatásba http://sek.nyime.hu/layouts/1038/Sport/ozsvath_acs_sportudkut121010.pdf (Letöltés dátuma: 2014.09.01.)
 54. Pakai A. – Kívés Zs. (2013): Kutatásról ápolóknak 2. rész: Mintavétel és adatgyűjtési módszerek az egészségtudományi kutatásokban. Nővér 26. évfolyam 3. szám 20-43. o.
 55. Pakai A. – Tóth M. – Várady Horváth Á. – Oláh A. – Horváth Ö. P. (2013): Lehetséges protektív tényezők a sebgyógyulásban egy felmérés tükrében. Nővér 26. évfolyam 2. szám 8-15. o.
 56. Pálvölgyi M. (2011): Információkereső nyelvek III. Eger: Eszterházy Károly Főiskola.
Online elérhető:

- http://www.tankonyvtar.hu/hu/tartalom/tamop425/0005_13_infkereso_nyelvek_iii_scor_m_05/531_online_katalogusok.html (2014-08-12.)
57. Papp L. (2013): Kutatásról ápolóknak 1. rész: A kutatás tervezése. Nővér 26. évfolyam 2. szám 3-7.o.
 58. Pintér J. – Rappai G. (2001): A mintavételi tervek készítésének néhány gyakorlati megfontolása. Marketing & Menedzsment 35. évfolyam 4. szám 4-11. o.
 59. Ramanathan R. (2003): Bevezetés az Ökonometriába alkalmazásokkal. Panem Kft. Budapest
 60. Rappai G. (2001): Üzleti statisztika Excellel. Központi Statisztikai Hivatal, Budapest
 61. Reiczigel, J. (2005): Válogatott fejezetek a biostatistikából. SZIE ÁOTK, Online elérhető: <http://www.univet.hu/users/jreiczig/valfej/val-fej-jegyzet-2005-02-05.pdf> (2014-06-30)
 62. Sajó A. (1999): Elektronikus folyóiratok az Interneten. Tudományos és Műszaki tájékoztatás 46. évfolyam 7. szám 275-280. o.
 63. Sajtos L. – Mitev A. (2007): SPSS kutatási és adatelemzési kézikönyv. Alinea Kiadó, Budapest, 402. o.
 64. Salavecz Gy. – Neculai K. – Rózsa S. – Kopp M. (2006): Az Erőfeszítés-Jutalom Egyensúlytalanság Kérdőív magyar változatának megbízhatósága és érvényessége. Mentálhigiéné és pszichoszomatika 7. évfolyam 3. szám. 231-246. o.
 65. Siegrist J – Klein D – Voigt KH (1997): Linking sociological with physiological data: the model of effort-reward imbalance at work. Acta Physiol Scand 161. 112–116. o.
 66. Szabó J. – Gerevich J. (2009): Kapcsolatok a felépülésben, felépülés a kapcsolatokban. A társas támogatottság mérése alkoholbetegek önéletrajzaiban. Lege Artis Medicinae 19. évfolyam 1. szám 67-72. o.
 67. Szabó K. (2002): Kommunikáció felsőfokon. Kossuth Kiadó. Budapest. 2.Kiadás. 404 o.
 68. Székelyi M. – Barna I. (2005): Túlélőkészlet az SPSS-hez. Typotex Kiadó, Budapest, 455. o.
 69. Szepesi J.: HunKat, a HunTékát használó könyvtárak közös katalógusa. Online elérhető: <http://www.bdf.hu/konyvtar/tempus/HunKat.pdf> (2014-08-15)
 70. Tomcsányi P. (2000): Általános kutatómódszertan. Szent István Egyetem, Gödöllő
 71. Vajda R. – Karamánné Pakai A. – Éliás Zs. – Sélleyné Gyuró M. – Tamás P. – Várnagy Á. – Kívés Zs. (2014): A méhnyakrákkal kapcsolatos ismeretek és a

szűrővizsgálaton való részvételi mutatók vizsgálata. LAM 24. évfolyam 3. szám 118-125. o.

72. Vasas L. (2008): Az orvosi szakirodalom adattárai 1. MEDLINE: mindig és mindenhol
73. Vízvári D. (2010): Könyvek, folyóiratok, adatbázisok. Egészségügyi Stratégiai Kutatóintézet, Budapest

13. MELLÉKLETEK

1. Melléklet

Forrás: Egészségügyi Tudományos Tanács honlapján a www.ett.hu

Hazai vonatkozásban a kutatás megkezdése előtt az alábbi törvényi szabályozás áttekintése ajánlott:

- Egészségügyi törvény: **1997. évi CLIV törvény** VIII-IX. fejezet
- Genetikai törvény: a humán genetikai adatok védelméről, a humán genetikai vizsgálatok és kutatások, valamint a biobankok működésének szabályairól szóló **2008. évi XXI. törvény**
- Gyógyszertörvény: az emberi alkalmazásra kerülő gyógyszerekről és egyéb, a gyógyszerpiacot szabályozó törvények módosításáról szóló **2005. évi XCV. törvény**
- A humán reprodukciós eljárásokkal kapcsolatos, kötelezően nyilvánosságra hozandó érvényességi adatok, statisztikák köréről, a nyilvánosságra hozatal módjáról és helyéről, továbbá az ellenőrzés módjáról szóló **339/2008 (XII.30.) Korm. Rendelet**.
- Az emberen végzett orvostudományi kutatások, az emberi felhasználásra kerülő vizsgálati készítmények klinikai vizsgálata, valamint az emberen történő alkalmazásra szolgáló klinikai vizsgálatra szánt orvostechonikai eszközök klinikai vizsgálata engedélyezési eljárásának szabályairól szóló **235/2009. (X.20.) Korm. Rendelet**.
- Az emberi reprodukcióra irányuló különleges eljárások végzésére vonatkozó, valamint az ivarsejtekkel és embriókkal való rendelkezésre és azok fagyasztva tárolására vonatkozó részletes szabályokról szóló **30/1998. (VI.24.) NM rendelet**.
- **23/2002 (V.9.) EüM rendelet** az emberen végzett orvostudományi kutatásokról.
/módosította: 31/2009. (X.20.) EüM rendelet/
- **28/2014. (IV. 10.) EMMI rendelet** Az Egészségügyi Tudományos Tanácsról.
- **35/2005. (VIII.26.) EüM rendelet** Az emberi felhasználásra kerülő vizsgálati készítmények klinikai vizsgálatáról és a helyes klinikai gyakorlat alkalmazásáról.
/módosította: 32/2009 (X.20.) EüM rendelet/
- **33/2009. (X.20.) EüM rendelet** Az orvostechonikai eszközök klinikai vizsgálatáról.
- **34/2009. (X.20.) EüM rendelet** Az Állami Népegészségügyi és Tisztiorvosi Szolgálat egyes közigazgatási eljárásaiért és az igazgatási jellegű szolgáltatásaiért fizetendő díjakról szóló 1/2009 (I.30.) EüM rendelet a népjóléti ágazatba tartozó egyes

államigazgatási eljárásokért és igazgatási jellegű szolgáltatási díjakról szóló 50/1996 (XII.27.) NM rendelet, valamint az emberen végzett orvostudományi kutatásokról szóló 23/2002 (V.9.) EüM rendelet módosításáról.

- **ADATVÉDELEM 2011. évi CXII. tv.** Az információs önrendelkezési jogról és az információszabadságról. (Ezzel az 1992. évi LXIII. tv. "A személyes adatok védelméről és a közérdekű adatok nyilvánosságáról" hatályát veszítette.) **1997. évi XLVII. tv.** Az egészségügyi és a hozzájuk kapcsolódó személyes adatok kezeléséről és védelméről. /II. fejezet 21. §. tudományos kutatás céljából történő adatkezelés/
- **BÜNTETŐ TÖRVÉNYKÖNYV 2012. évi C. törvény XVI. fejezet** "az egészségügyi beavatkozás és kutatás rendje elleni bűncselekmények, valamint az egészségügyi önrendelkezési elleni bűncselekmények".
- **SZAKÉRTŐI TESTÜLET 33/2007. (VI.22.) IRM rendelet** Az Igazságügyi Szakértői Testületek szervezetéről és működéséről

14. FÜGGELÉK (TÁBLÁZATOK)

1. Standard normális eloszlás

Sűrűségfüggvény értékei

z	0	1	2	3	4	5	6	7	8	9
0,0	0,500	0,496	0,492	0,488	0,484	0,480	0,476	0,472	0,468	0,464
0,1	0,460	0,456	0,452	0,448	0,444	0,440	0,436	0,433	0,429	0,425
0,2	0,421	0,417	0,413	0,409	0,405	0,401	0,397	0,394	0,390	0,386
0,3	0,382	0,378	0,374	0,371	0,367	0,363	0,359	0,356	0,352	0,348
0,4	0,345	0,341	0,337	0,334	0,330	0,326	0,323	0,319	0,316	0,312
0,5	0,309	0,305	0,302	0,298	0,295	0,291	0,288	0,284	0,281	0,278
0,6	0,274	0,271	0,268	0,264	0,261	0,258	0,255	0,251	0,248	0,245
0,7	0,242	0,239	0,236	0,233	0,230	0,227	0,224	0,221	0,218	0,215
0,8	0,212	0,209	0,206	0,203	0,200	0,198	0,195	0,192	0,189	0,187
0,9	0,184	0,181	0,179	0,176	0,174	0,171	0,169	0,166	0,164	0,161
1,0	0,159	0,156	0,154	0,152	0,149	0,147	0,145	0,142	0,140	0,138
1,1	0,136	0,133	0,131	0,129	0,127	0,125	0,123	0,121	0,119	0,117
1,2	0,115	0,113	0,111	0,109	0,107	0,106	0,104	0,102	0,100	0,099
1,3	0,097	0,095	0,093	0,092	0,090	0,089	0,087	0,085	0,084	0,082
1,4	0,081	0,079	0,078	0,076	0,075	0,074	0,072	0,071	0,069	0,068
1,5	0,067	0,066	0,064	0,063	0,062	0,061	0,059	0,058	0,057	0,056
1,6	0,055	0,054	0,053	0,052	0,051	0,049	0,048	0,047	0,046	0,046
1,7	0,045	0,044	0,043	0,042	0,041	0,040	0,039	0,038	0,038	0,037
1,8	0,036	0,035	0,034	0,034	0,033	0,032	0,031	0,031	0,030	0,029
1,9	0,029	0,028	0,027	0,027	0,026	0,026	0,025	0,024	0,024	0,023
2,0	0,023	0,022	0,022	0,021	0,021	0,020	0,020	0,019	0,019	0,018
2,1	0,018	0,017	0,017	0,017	0,016	0,016	0,015	0,015	0,015	0,014
2,2	0,014	0,014	0,013	0,013	0,013	0,012	0,012	0,012	0,011	0,011
2,3	0,011	0,010	0,010	0,010	0,010	0,009	0,009	0,009	0,009	0,008
2,4	0,008	0,008	0,008	0,008	0,007	0,007	0,007	0,007	0,007	0,006
2,5	0,006	0,006	0,006	0,006	0,006	0,005	0,005	0,005	0,005	0,005
2,6	0,005	0,005	0,004	0,004	0,004	0,004	0,004	0,004	0,004	0,004
2,7	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003
2,8	0,003	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002
2,9	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,001	0,001	0,001
3,0	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001

2. Kritikus értékek különböző szignifikancia- szintek esetén

Szignifikancia-szint (α)						
Egyoldalú	0,1000	0,0500	0,0250	0,0225	0,0100	0,0050
Kétoldalú	0,2000	0,1000	0,0500	0,0450	0,0200	0,0100
z	1,280	1,645	1,960	2,000	2,330	2,587

3. Student féle t-eloszlás

Student féle t-eloszlás kritikus értékei különféle szignifikancia-szint mellett

Szabadság-fok	Szignifikancia-szint				
	0,1	0,05	0,025	0,01	0,005
1	3,078	6,314	12,706	31,821	63,656
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807
24	1,318	1,711	2,064	2,492	2,797
25	1,316	1,708	2,060	2,485	2,787
26	1,315	1,706	2,056	2,479	2,779
27	1,314	1,703	2,052	2,473	2,771
28	1,313	1,701	2,048	2,467	2,763
29	1,311	1,699	2,045	2,462	2,756
30	1,310	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
50	1,299	1,676	2,009	2,403	2,678
60	1,296	1,671	2,000	2,390	2,660
70	1,294	1,667	1,994	2,381	2,648
80	1,292	1,664	1,990	2,374	2,639
90	1,291	1,662	1,987	2,368	2,632
100	1,290	1,660	1,984	2,364	2,626
150	1,287	1,655	1,976	2,351	2,609
200	1,286	1,653	1,972	2,345	2,601

4. χ^2 -eloszlás

χ^2 -eloszlás kritikus értékei különféle szignifikancia-szintek mellett

Szabadságfok	Szignifikancia-szint					
	0,9900	0,9500	0,9000	0,1000	0,0500	0,0100
1	0,000	0,004	0,016	2,706	3,841	6,635
2	0,020	0,103	0,211	4,605	5,991	9,210
3	0,115	0,352	0,584	6,251	7,815	11,345
4	0,297	0,711	1,064	7,779	9,488	13,277
5	0,554	1,145	1,610	9,236	11,070	15,086
6	0,872	1,635	2,204	10,645	12,592	16,812
7	1,239	2,167	2,833	12,017	14,067	18,475
8	1,647	2,733	3,490	13,362	15,507	20,090
9	2,088	3,325	4,168	14,684	16,919	21,666
10	2,558	3,940	4,865	15,987	18,307	23,209
11	3,053	4,575	5,578	17,275	19,675	24,725
12	3,571	5,226	6,304	18,549	21,026	26,217
13	4,107	5,892	7,041	19,812	22,362	27,688
14	4,660	6,571	7,790	21,064	23,685	29,141
15	5,229	7,261	8,547	22,307	24,996	30,578
16	5,812	7,962	9,312	23,542	26,296	32,000
17	6,408	8,672	10,085	24,769	27,587	33,409
18	7,015	9,390	10,865	25,989	28,869	34,805
19	7,633	10,117	11,651	27,204	30,144	36,191
20	8,260	10,851	12,443	28,412	31,410	37,566
21	8,897	11,591	13,240	29,615	32,671	38,932
22	9,542	12,338	14,041	30,813	33,924	40,289
23	10,196	13,091	14,848	32,007	35,172	41,638
24	10,856	13,848	15,659	33,196	36,415	42,980
25	11,524	14,611	16,473	34,382	37,652	44,314
26	12,198	15,379	17,292	35,563	38,885	45,642
27	12,878	16,151	18,114	36,741	40,113	46,963
28	13,565	16,928	18,939	37,916	41,337	48,278
29	14,256	17,708	19,768	39,087	42,557	49,588
30	14,953	18,493	20,599	40,256	43,773	50,892
31	15,655	19,281	21,434	41,422	44,985	52,191
32	16,362	20,072	22,271	42,585	46,194	53,486
33	17,073	20,867	23,110	43,745	47,400	54,775
34	17,789	21,664	23,952	44,903	48,602	56,061
35	18,509	22,465	24,797	46,059	49,802	57,342
36	19,233	23,269	25,643	47,212	50,998	58,619
37	19,960	24,075	26,492	48,363	52,192	59,893
38	20,691	24,884	27,343	49,513	53,384	61,162
39	21,426	25,695	28,196	50,660	54,572	62,428
40	22,164	26,509	29,051	51,805	55,758	63,691
50	29,707	34,764	37,689	63,167	67,505	76,154
60	37,485	43,188	46,459	74,397	79,082	88,379
70	45,442	51,739	55,329	85,527	90,531	100,425
80	53,540	60,391	64,278	96,578	101,879	112,329
90	61,754	69,126	73,291	107,565	113,145	124,116
100	70,065	77,929	82,358	118,498	124,342	135,807
150	112,668	122,692	128,275	172,581	179,581	193,207
200	156,432	168,279	174,835	226,021	233,994	249,445
250	200,939	214,392	221,806	279,050	287,882	304,939

5. F-eloszlás

F-eloszlás kritikus értékei 5%-os egyoldalú (10%-os kétoldalú) szignifikancia-szint mellett

Nevező szf.	Számológó szabadságfoka															
	1	2	3	4	5	6	7	8	9	10	15	20	25	30	50	100
2	18,5	19,0	19,1	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5
3	10,1	9,5	9,2	9,1	9,0	8,9	8,8	8,8	8,8	8,7	8,7	8,6	8,6	8,6	8,5	8,5
4	7,7	6,9	6,5	6,3	6,2	6,1	6,0	6,0	6,0	5,9	5,8	5,8	5,7	5,7	5,7	5,6
5	6,6	5,7	5,4	5,1	5,0	4,9	4,8	4,8	4,7	4,7	4,6	4,5	4,5	4,5	4,4	4,4
6	5,9	5,1	4,7	4,5	4,3	4,2	4,2	4,1	4,1	4,0	3,9	3,8	3,8	3,8	3,7	3,7
7	5,5	4,7	4,3	4,1	3,9	3,8	3,7	3,7	3,6	3,6	3,5	3,4	3,4	3,3	3,3	3,2
8	5,3	4,4	4,0	3,8	3,6	3,5	3,5	3,4	3,3	3,3	3,2	3,1	3,1	3,0	3,0	2,9
9	5,1	4,2	3,8	3,6	3,4	3,3	3,2	3,2	3,1	3,1	3,0	2,9	2,8	2,8	2,8	2,7
10	4,9	4,1	3,7	3,4	3,3	3,2	3,1	3,0	3,0	2,9	2,8	2,7	2,7	2,7	2,6	2,5
11	4,8	3,9	3,5	3,3	3,2	3,0	3,0	2,9	2,9	2,8	2,7	2,6	2,6	2,5	2,5	2,4
12	4,7	3,8	3,4	3,2	3,1	3,0	2,9	2,8	2,8	2,7	2,6	2,5	2,5	2,4	2,4	2,3
13	4,6	3,8	3,4	3,1	3,0	2,9	2,8	2,7	2,7	2,6	2,5	2,4	2,4	2,3	2,3	2,2
14	4,6	3,7	3,3	3,1	2,9	2,8	2,7	2,7	2,6	2,6	2,4	2,3	2,3	2,3	2,2	2,1
15	4,5	3,6	3,2	3,0	2,9	2,7	2,7	2,6	2,5	2,5	2,4	2,3	2,2	2,2	2,1	2,1
16	4,4	3,6	3,2	3,0	2,8	2,7	2,6	2,5	2,5	2,4	2,3	2,2	2,2	2,1	2,1	2,0

6. F-eloszlás kritikus értékei 2,5%-os egyoldali (5%-os kétoldali) szignifikancia-szint mellett

Nevező szf.	Számoló szabadságfoka															
	1	2	3	4	5	6	7	8	9	10	15	20	25	30	50	100
2	38,5	39,0	39,2	39,3	39,3	39,3	39,4	39,4	39,4	39,4	39,4	39,5	39,5	39,5	39,5	39,5
3	17,4	16,0	15,4	15,1	14,9	14,7	14,6	14,5	14,5	14,4	14,4	14,4	14,4	14,4	14,4	14,4
4	12,2	10,7	10,0	9,6	9,3	9,2	9,0	8,9	8,9	8,8	8,6	8,5	8,5	8,4	8,3	8,3
5	10,0	8,4	7,7	7,3	7,1	6,9	6,8	6,7	6,6	6,6	6,4	6,3	6,2	6,2	6,1	6,0
6	8,8	7,2	6,6	6,2	5,9	5,8	5,7	5,6	5,5	5,4	5,2	5,1	5,1	5,0	4,9	4,9
7	8,0	6,5	5,8	5,5	5,2	5,1	4,9	4,9	4,8	4,7	4,5	4,4	4,4	4,3	4,2	4,2
8	7,5	6,0	5,4	5,0	4,8	4,6	4,5	4,4	4,3	4,3	4,1	4,0	3,9	3,8	3,8	3,7
9	7,2	5,7	5,0	4,7	4,4	4,3	4,2	4,1	4,0	3,9	3,7	3,6	3,6	3,5	3,4	3,4
10	6,9	5,4	4,8	4,4	4,2	4,0	3,9	3,8	3,7	3,7	3,5	3,4	3,3	3,3	3,2	3,1
11	6,7	5,2	4,6	4,2	4,0	3,8	3,7	3,6	3,5	3,5	3,3	3,2	3,1	3,1	3,0	2,9
12	6,5	5,1	4,4	4,1	3,8	3,7	3,6	3,5	3,4	3,3	3,1	3,0	3,0	2,9	2,8	2,8
13	6,4	4,9	4,3	4,0	3,7	3,6	3,4	3,3	3,3	3,2	3,0	2,9	2,8	2,8	2,7	2,6
14	6,3	4,8	4,2	3,8	3,6	3,5	3,3	3,2	3,2	3,1	2,9	2,8	2,7	2,7	2,6	2,5
15	6,2	4,7	4,1	3,8	3,5	3,4	3,2	3,2	3,1	3,0	2,8	2,7	2,6	2,6	2,5	2,4
16	6,1	4,6	4,0	3,7	3,5	3,3	3,2	3,1	3,0	2,9	2,7	2,6	2,6	2,5	2,4	2,4



01

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

92%

02

$$R = x_{\max} - x_{\min}$$

93%

03

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

95%

04

$$TQ = Q_3 - Q_1$$

95%

05

$$Y = \frac{f_{11} \times f_{22} - f_{12} \times f_{21}}{f_{11} \times f_{22} + f_{12} \times f_{21}}$$

97%