



Adatfeldolgozás és elemzés

Bevezetés a kutatómódszertanba
WJLF SZM szak
SE addkonz posztgrad képzés
Pecze Mariann

AMIRŐL MA SZÓ LESZ

Az adatok előkészítése

- A változók típusától függően a mérési szintek meghatározása,
- Adatok rögzítése táblázatban, változók és attribútumainak meghatározása,
- kódolás, adattisztítás
- ADATBÁZIS kialakítása

Egyszerűbb elemzések

- Gyakoriság
- A középérték mérőszámai (átlag, módusz, medián)
- Szórás
- Keresztábra elemzés

Adatok grafikai ábrázolása

- Grafikus ábrázolási módszerek funkciói
- A diagram részei
- A grafikus ábrázolás alapelvei
- A legelterjedtebb ábrázolási formák
- az adatok grafikai ábrázolásának tipikus hibái

„puha” adatok elemzése

- Az interjú szövegének átírása,
- Adatok rendezése,
- Kódolás, kategóriák kialakítása,
- Elemzés, értelmezés szakasza



Az adatok előkészítése

1



1. Az adatok előkészítése

A) Változók -
Mérési szintek

B) Adatrögzítés,
kódolás

C) adattisztítás

A) VÁLTOZÓK ÉS MÉRÉSI SZINTEK

MÉRÉSI SZINTEK:



4 féle hozzárendelési szabály, mérési szint



Mérési szintek

Nominális mérési szint. „Minőségi” változók. Kategóriáikhoz technikai okból gyakran számok vannak rendelve. („nem” változó, 1: férfi, 2: nő) Ezek a számok tetszőlegesek, nem jelölnek mennyiségi, csak minőségi különbséget. Példák: párt-szimpatia, vallás, nemzetiségi hovatartozás.

Ordinális mérési szint. A változó sorba rendezhető kategóriáihoz a sorrendnek megfelelően számokat rendelünk. Pl. társadalmi osztályok, 1: munkásosztály, 2: középosztály, 3: felső osztály. A számok nagysága viszont kötetlen (1-2-3 helyett lehetne 1-3-8 is). Mondhatjuk, hogy a felső osztály magasabb pozíciót jelent, mint a középosztály, és hogy a középosztály is magasabb, mint a munkásosztály, de nem hasonlíthatjuk össze a két pár távolságát. Másik példa: tanya/falu/nagyközség/város/főváros.

Intervallumskála. A fenti példával szemben itt ismerjük a szomszédos értékek távolságát, de nincs „abszolút 0 pont” (szokásos példa a Celsius-fok); a zérus megválasztása esetleges (víz fagyáspont). Ilyen az IQ-pont is: 200 vs. 100 pont: nem kétszer okosabb.

Arányskála. Van „abszolút 0 pont” is (mint Kelvin-fok). Itt már van értelme az osztás műveletnek is. Ilyen pl. súly, magasság, jövedelem.

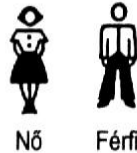
Példa: Mérési szintek és lehetséges összehasonlítások: iskolázottság

Nominális	Alapítványi középiskola / Állami középiskola	„különböző oktatásban részesültek”
Ordinális	Főiskolai / Egyetemi diploma	„magasabb iskolázottság”
Intervallum-arány	8 osztályt végzett / 16 osztályt végzett	„kétszer iskolázottabb”

Kategoriális változó
Csak tulajdonságot jelölnek

Numerikus (folytonos) változó
Tényleges számértéket mutat

Példa nominális szintű változóra:

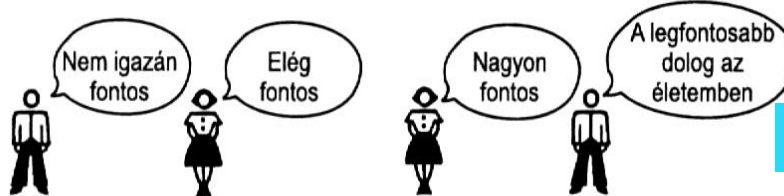


Nem, foglalkozás,
településnév, vallás,
nemzetiség...

kategória

medián

Példa ordinális mérési szintű változóra: vallásosság „Mennyire fontos az Ön számára a vallás?”



Településtípus

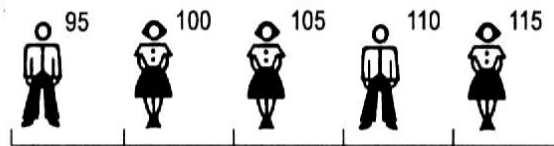
Kategória,
sorrend

módusz

Alacsony

Magas

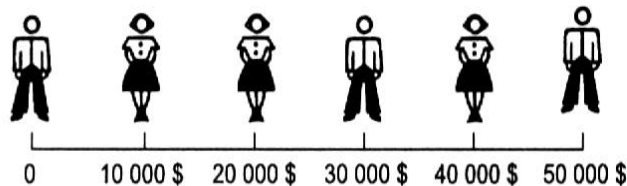
Példa intervallum mérési szintű változóra: IQ



hőmérséklet

Kategória,
sorrend,
különbség

Példa arányskálára: jövedelem



súly, magasság

Kategória,
sorrend,
különbség,
arány

B) ADATRÖGZÍTÉS, KÓDOLÁS

- Elektronikus adatok esetében egyszerűbb – adatkonvertálás

Papír alapú adatok kódolása és rögzítése számítógépen

	kor	nem	Isk.végz.	jövedelem	...
1. Mintaelem (kérdőív)	24	Nő	8 ált.	42.000	
2. Mintaelem (kérdőív)	20	ffi	Gimn.	80.000	
...					

Ilyen formában nem mérhető!

kódolás

● Kódutasítás és kódtáblázat

Kérdőív:

Mi az Ön neme?

Férfi

nő

Kódutasítás:

Mi az Ön neme? {NEM}

1. Férfi

2. nő

3. 99 nincs válasz



Kérdőív:

Van-e az Önök háztartásában:

Mikrosütő

Mosógép

Számítógép

...

Kódutasítás:

Van-e az Önök háztartásában:

{HMIK} Mikrosütő 0-nincs válasz

{HMOS} Mosógép 1-van

{HSZGEP} Számítógép 2-nincs

...



Nyitott kérdések kódolása

szöveges válasz is beírható (Excel, SPSS szöveges változói) – kategóriarendszer felállítása

- ne legyen túl sok kategória
- minden válasz beleférjen valamely kategóriába

C) ADATFILE ELŐKÉSZÍTÉSE, ADATTISZTÍTÁS

	Name	Typ	Wi	Dec	Label	Values	Missing
1	sorszam	Nu	8	0	sorszám	None	0
2	nem	Dot	8	0	Nem	{0, nincs vá	0
3	szulev	Nu	8	0	születési év	{0, nincs vá	0
4	iskvapa	Nu	8	0	apa iskolai végzettsége	{0, nincs vá	0
5	iskvanya	Nu	8	0	anya iskolai végzettsége	{0, nincs vá	0
6	foglapa	Nu	8	0	apa foglalkozása	{0, nincs vá	0
7	foglanya	Nu	8	0	anya foglalkozása	{0, nincs vá	0
8	tarthely	Nu	8	0	tartózkodási hely	{0, nincs vá	0
9	havift	Nu	6	0	havi gazdálkodási összeg	{0, nincs vá	0
10	havftott	Nu	1	0	otthoni támogatás aránya	{999, nincs	999
11	havftsjt	Nu	8	0	saját kereset aránya	{999, nincs	999
12	havftto	Nu	8	0	ösztöndíj aránya	{999, nincs	999
13	havfttam	Nu	8	0	támogatások aránya	{999, nincs	999
14	havfthit	Nu	8	0	diákhitel aránya	{999, nincs	999
15	havftegy	Nu	8	0	egyéb forrás aránya	{999, nincs	999
16	tapaszt	Nu	8	0	iskolaválasztás előtti tapasztalat	{0, nincs vá	0
17	szocprobl	Nu	8	0	szociális probléma a környezetben	{0, nincs vá	0



	sorszam	nem	szulev	iskvapa	iskvanya
1	101	2	89	3	4
2	102	2	83	3	2
3	103	2	86	5	4
4	104	2	88	3	4
5	105	2	88	6	6
6	106	2	89	3	6
7	107	2	87	3	3
8	108	2	89	5	5
9	109	2	85	3	4
10	110	1	87	4	6
11	111	2	86	4	3
12	112	2	88	5	4
13	113	1	86	2	2
14	114	1	85	1	1
15	115	1	88	6	5
16	116	1	85	5	6
17	117	1	87	5	5
18	118	2	89	6	6
19	119	2	77	2	2

Az összes változó definiálása (címkézés)

Az attribútumok megadása (kódutasításnak megfelelően)

Változó neve legyen egyértelmű (más számára is használható)

C) ADATTISZTÍTÁS

1. Kérdőívek ellenőrzése

kérdezőbiztos: át kell néznie kitöltés és leadás előtt
önkitöltős: ha lehet, mindet ellenőrizzük

lehetséges hibák:

Hiányzó válaszok

- a)Oka: szűrő kérdés miatt nincs adat (ez nem hiba!)
- b)Oka: kérdezőbiztos nem tette fel a kérdést vagy nem rögzítette a választ
- Esetenként pótolható az adathiány

Logikai ellentmondások

Nyitott kérdéseknél

Olvashatóság kérdése

Hiányzó kérdőívlapok

Leszakadt, szétesett...
pótlás vagy adathiány

C) ADATTISZTÍTÁS



2. Digitális adatfile tisztítása

Az adatbázisokban tárolt adatok minőségét javító, többnyire belső logikai ellenőrzésen alapuló eljárások összessége az adatfelvételi, a kódolási és a rögzítési hibák feltárása és kijavítása révén.

Az adatrögzítő hibáinak kiküszöbölése

elütések kiszűrése: egy gyakoriság (frequencies) lekérdezése után kijövő ellentmondások

Pl.: neme: 22 (valószínűleg elütés: 2-helyett 22)

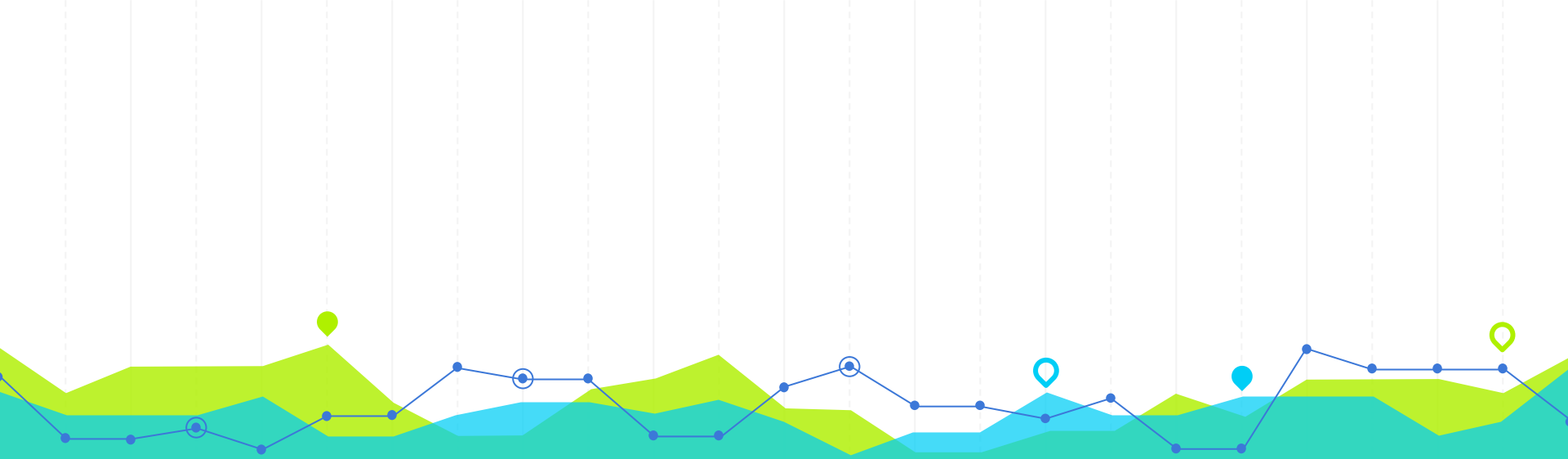
de: 125 000-es kódnál lehet elcsúszás az adatfile-ban!

korrekt: hiányzó értéknek jelölni (missing), így nem torzítja a statisztikát

Adattisztítás gyakoriság lekérdezésével

anya iskolai végzettsége

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	kevesebb, mint 8 általános	1	2,1	2,1	2,1
	8 általános	5	10,4	10,4	12,5
	szakmunkásképző	11	22,9	22,9	35,4
	szakközépiskola, gimnázium	14	29,2	29,2	64,6
	technikum	5	10,4	10,4	75,0
	főiskola, egyetem	12	25,0	25,0	100,0
	Total	48	100,0	100,0	



Egyszerűbb elemzések

Gyakoriság

A középérték mérőszámai (átlag, módusz, medián)

Szórás

Keresztábra elemzés

2



EGYSZERŰ FELMÉRÉSEK



gyakoriság

gyakoriság

ha sok a válaszlehetőség (pl. életkor), célszerű a kategóriákat képezni. (ne túl sokat!)

		születési év			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1951	1	1,2	1,2	
	1962	1	1,2	1,2	
	1965	2	2,4	2,4	
	1969	2	2,4	2,4	7,1
	1972	1	1,2	1,2	8,3
	1973	1	1,2	1,2	9,5
	1975	1	1,2	1,2	10,7
	1976	1	1,2	1,2	11,9
	1977	1	1,2	1,2	13,1
	1978	4	4,7	4,8	17,9
	1979	5	5,9	6,0	23,8
	1980	10	11,8	11,9	35,7
	1981	11	12,9	13,1	48,8
	1982	12	14,1	14,3	63,1
	1983	8	9,4	9,5	72,6
	1984	11	12,9	13,1	85,7
	1985	9	10,6	10,7	96,4
	1986	3	3,5	3,6	100,0
	Total	84	98,8	100,0	
Missing	nincs válasz	1	1,2		
	Total	85	100,0		

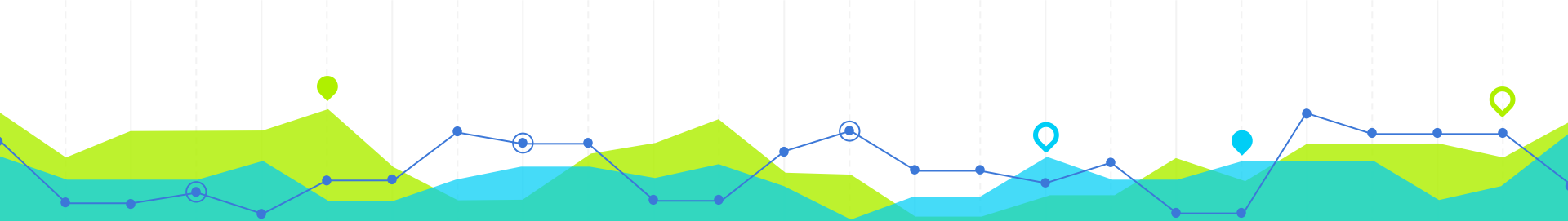
Képezzünk kategóriákat!

nézhetünk:

egyszerű gyakoriságot
(*frequency*; előfordulási
elemszám)

százalékot (*percent*)

a hiányzó értékektől
letisztított százalékot
(*valid percent*)



A középérték mérőszámai

a középérték mérőszámai

átlag

numerikus változónál lehet számolni

a változó összes értéke/adatok száma

- pl. 1 főre jutó jövedelem átlaga

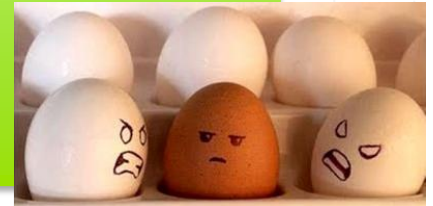
medián

az adatokat növekvő/csökkenő sorrendbe állítjuk, **a minta közepén megjelenő érték**; a minta adatainak fele felette, fele alatta lesz a szélső értékek befolyásolhatják az átlagot! – medián jövedelem: fölötté u.annyian helyezkednek el, mint alatta

módusz

a **leggyakrabban** előforduló érték

- az a kategória, amelybe a vizsgált populációból a legtöbben tartoznak



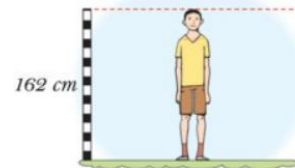
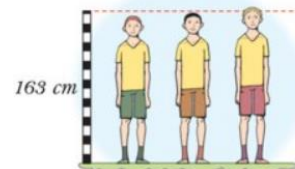
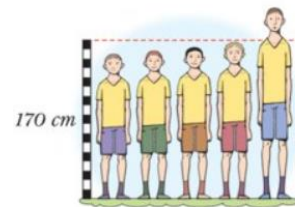
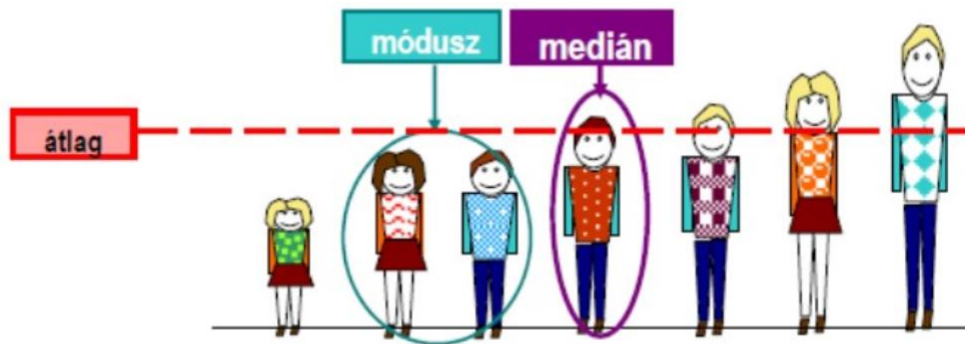
Átlag, medián, módusz

Az **átlag** nem mindig jellemzi jól az adatokat, mert az átlagot egy kiugró érték eltorzíthatja.

Ha az adatokat nagyság szerint sorba rendezzük, akkor az adathalmazt a középső adattal is jellemezhetjük, mert ugyanannyi adat kisebb nála, mint amennyi nagyobb. Ezért előfordul, hogy ez a „középtérték” pontosabb képet ad az adatokról, mint az átlag.

Ha páratlan számú adat van, akkor az adatok **mediánjának** nevezzük az adatok nagyság szerinti sorozatában a középső elemet. Páros számú adat esetén a medián a két középső adat átlaga.

Az adatok **móduszának** a leggyakoribb adatot nevezzük.



Hányan laknak egy háztartásban?

Válaszok 1-7-ig terjedtek

Mennyi lehet az átlagos létszám?

Négy?

Az egyes értékek gyakorisága nem azonos!

egyedül	112	Öten	127
Ketten	293	Hatan	28
Hárman	370	Heten	7
négyen	501	Össz.	1438

Mi lehet az átlag?

Számítás: $(1 \times 112 + 2 \times 293 + 3 \times 370 + 4 \times 501 + 5 \times 127 + 6 \times 28 + 7 \times 7) / 1438 = 3,24$

Mi lehet a medián (középérték)?

A minta közepét megjelenítő érték – a mintában szereplő adatok fele (az első 719 válaszadó) a medián alatt, fele (719) a medián felett található. E változónál ez az érték: 3

Mi lehet a módusz (leggyakrabban előforduló érték)?

A válaszadók közül legtöbben 4-en laknak egy háztartásban.

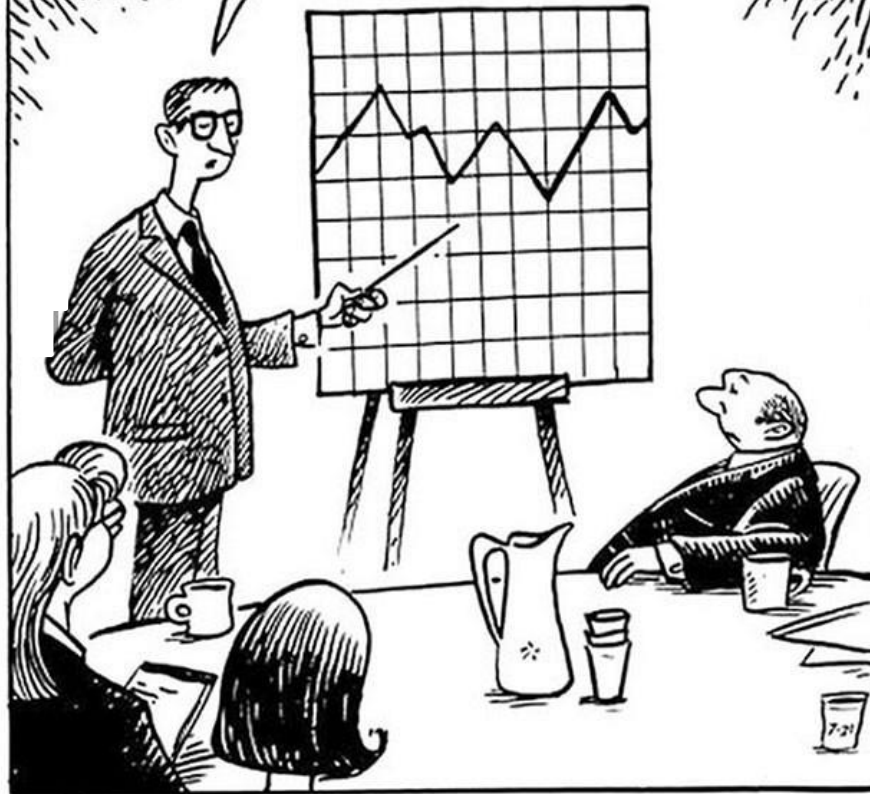
Átlag, medián és módusz értékei egy másik konkrét felmérésben

Statistics

havi gazdálkodási összeg

N	Valid	48
	Missing	0
Mean		49235,35
Median		40000,00
Mode		20000

...Ez a grafikon pedig megmutatja,
hogy mit láthatunk, ha egy teniszütőn
keresztül szemlélünk egy hegyláncot.



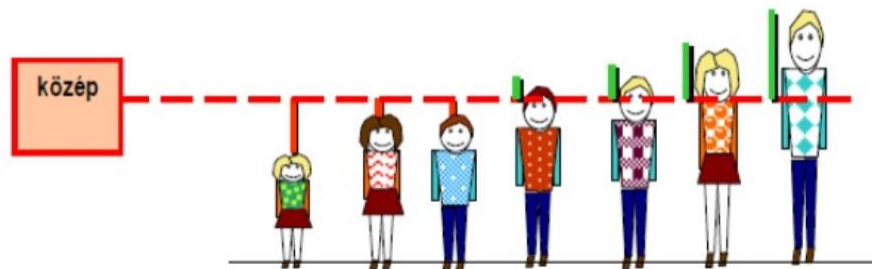


ne maradjunk meg a középérték
elemzésénél, nézzük meg a szóródást is!

szórás

szórás

Annak megállapítása, hogy az egyes adatok mennyire térnek el a középértéktől



Minél inkább közelít a nullához a szórás, annál jellemzőbbek az átlag körüli értékek.

szórás

Pl.: jövedelmi viszonyok megítélése:

- a sokaság nagy része az átlag közelében van,
- vagy sokan messze az átlag alatt és/vagy felett helyezkednek el?

Mérésére leggyakrabban használt statisztika a *standard deviáció* (SPSS-ben: *Std. Deviation*)

Minél inkább közelít a nullához a szórás, annál jellemzőbbek az átlag körüli értékek.

Pl.: megkérdezettek jövedelmének átlaga: 97660 Ft
Lehetne ez úgy is, hogy mindenki ilyen jövedelmű, de úgy is, ha egyik fele 1 Ft-ot mond, a másik fele 195319 Ft-ot. Így a két esetben más a szórás! 0 és 97692,37 a két esetben az érték



Keresztábra - elemzés

keresztábla-elemzés (crosstabulation)

két változó összevetése egy táblázatban
abszolút számok mellett: százalék sorok és oszlopok szerint is

Kizárható-e, hogy a talált összefüggés nem csak a mi mintánkra jellemző?



statisztikai hipotézisvizsgálat (khi-négyzet próba-Pearson) –
szignifikanciaszint meghatározása: mekkora a valószínűsége annak, hogy
az összefüggést mintavételi hiba okozta? (határérték szokásjog alapján
0,05, ez alatt az eredmény szignifikáns)

Háromdimenziós keresztáblák esetén legyen elegendő elemszám!

Keresztábra példa

Nem * iskolaválasztás előtti tapasztalat Crosstabulation

			iskolaválasztás előtti tapasztalat		Total
			nem volt	volt	
Nem	férfi	Count	10	0	10
		% within Nem	100,0%	,0%	100,0%
	% within iskolaválasztás előtti tapasztalat	23,3%	,0%	20,8%	
	% of Total	20,8%	,0%	20,8%	
nő	Count	33	5	38	
		% within Nem	86,8%	13,2%	100,0%
	% within iskolaválasztás előtti tapasztalat	76,7%	100,0%	79,2%	
	% of Total	68,8%	10,4%	79,2%	
Total	Count	43	5	48	
		% within Nem	89,6%	10,4%	100,0%
	% within iskolaválasztás előtti tapasztalat	100,0%	100,0%	100,0%	
	% of Total	89,6%	10,4%	100,0%	

			nagyobb értékű lakás	kisebb értékű lakás	összesen
összesen	alacsony foglalkozás	presztízsű	23%	77%	100%
	magas foglalkozás	presztízsű	65%	35%	100%
alacsony jövedelem	alacsony foglalkozás	presztízsű	33%	67%	100%
	magas foglalkozás	presztízsű	32%	68%	100%
magas jövedelem	alacsony foglalkozás	presztízsű	72%	28%	100%
	magas foglalkozás	presztízsű	69%	31%	100%

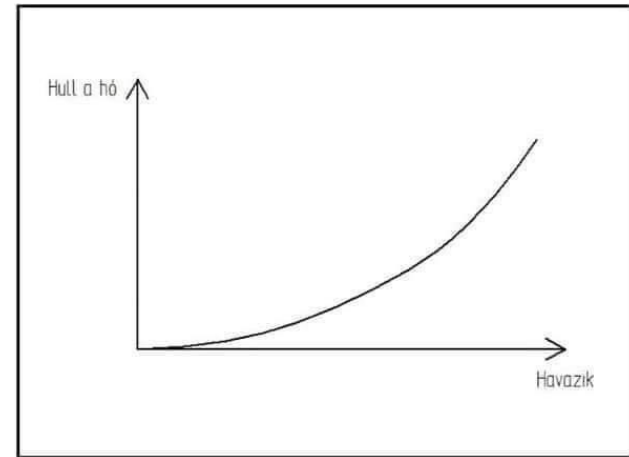
A keresztábra (táblázás, tabulation) változók közötti kapcsolat (asszociáció) jellemzésére alkalmas adattábla. A mátrixban többnyire két (nominális vagy ordinális) változó értékeinek együttes eloszlása ábrázolható, azaz a változókhoz tartozó értékek kereszt-kombinációit jeleníti meg.

Nincs elegendő elemszám, következtetés levonása elhamarkodott lenne!

Elemzés?



- Égbolt
- Piramis napos oldal
- Piramis árnyékos oldal



**Minél inkább hull a hó,
annál inkább havazik**

Összefoglalva:

Az elemzéshez használt változók mérési szintjére azért kell odafigyelni, hogy a szociológus ne alkalmazzon adatai elemzésére olyan módszereket, amelyek az adott mérési szinten nem alkalmazhatók, csupán a magasabbakon.

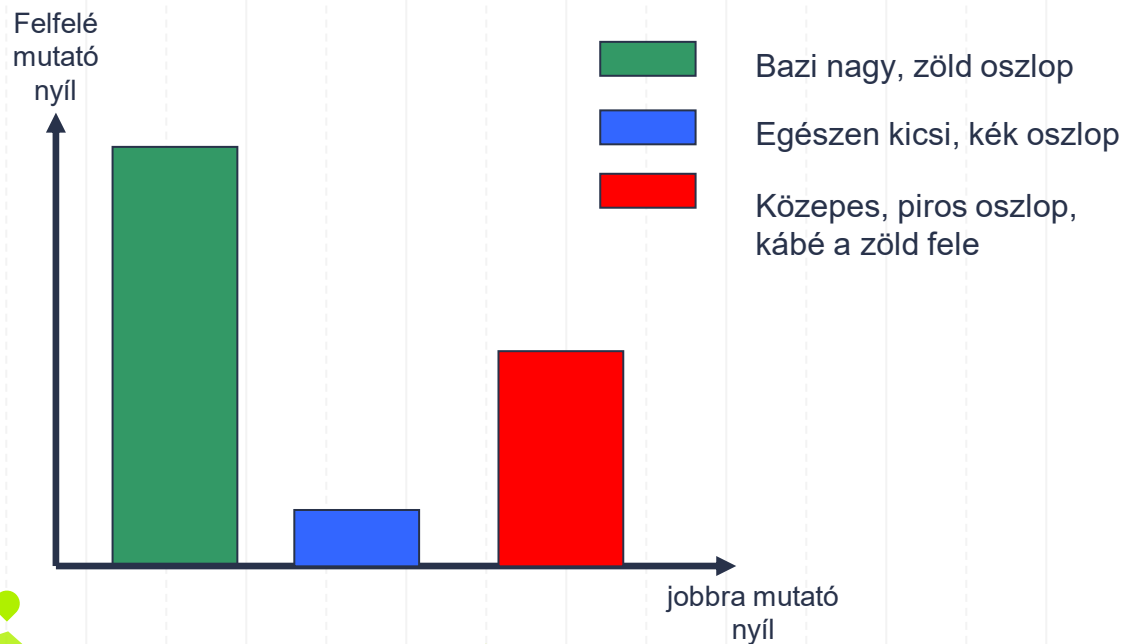
Például két változónk közötti kapcsolat erősségének mérésére nominális szinten mért változók esetén a **chi-négyzetet** lehet használni, ordinális szintű változók esetén a **rangkorrelációs együtthatót**, intervallum- és aránymérő szinten mért változók esetén pedig a **Pearson-féle korrelációs együtthatót**.

Az adatok elemzésének első lépése rendszerint az, hogy megnézzük minden egyes változónk gyakorisági eloszlását, tehát azt, hogy a különböző kódkategóriákba hány megkérdezett személy került. Ezután végiggondoljuk, hogy mely változópárok kombinációját volna érdemes megvizsgálni, és ennek megfelelően kétdimenziós kereszttáblákat készítünk.

Először többnyire „vizszintesen és függőlegesen” „kiszázelékoljuk” a táblázatot, vagyis kiszámítjuk a százalékos meg-oszlásokat, és azokat összehasonlítjuk. Ha a százalékos megoszlások különbségeket mutatnak, akkor a nominális szinten felhasználható statisztikák egyikével (chi-négyzet stb.) ellenőrizhető, hogy a kimutatott különbségek szignifikánsak-e (tehát nem lehetséges-e, hogy csak a véletlen okozta őket) és hogy milyen erősek.

Kellőképpen nagy minták esetében három vagy négy változót kombinálva, tehát három- vagy négydimenziós táblázatokat is lehet készíteni, ezek „olvasása” azonban a dimenziók számának növekedésével egyre bonyolultabbá válik.

Elemzés?



Az adatokat **táblázatba** rendezhetjük
vagy **diagramon** ábrázolhatjuk.

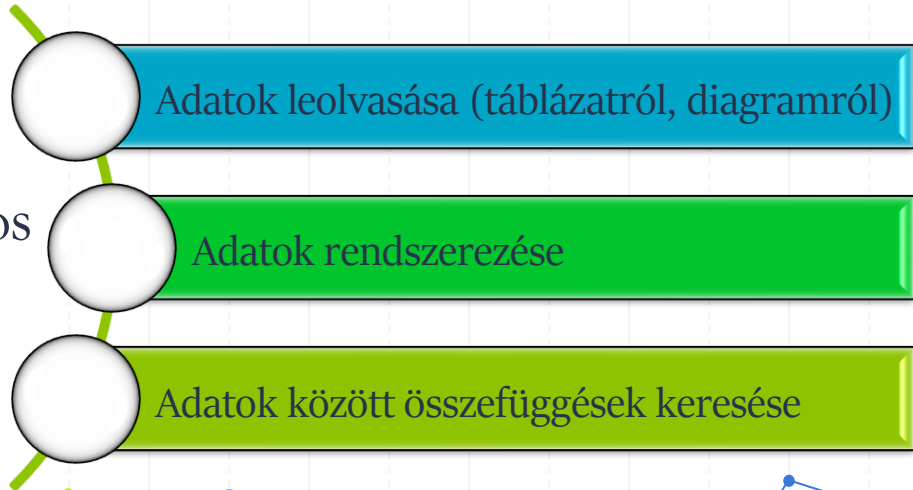
Főbb grafikus ábrázolási módszerek

3



*Az összegyűjtött adatok rendszerezését és a következtetést **adatfeldolgozás**nak nevezzük.*

Az adatokkal kapcsolatos
tevékenységek:



Grafikus ábrázolási módszerek típusai, funkciói

TÍPUSAI

Általános
statisztikai grafikus
módszerek;

Térképészeti
eljárások

FUNKCIÓI

Kutatási munkában
elemzési eszköz

Dolgozatban,
prezentációban
illusztrációs cél

Jelmagyarázat

az adatsorok nevét tartalmazza (milyen adatokat ábrázolunk?). Forrása a táblázat első oszlopa.

Kategóriatengely

a különböző adatsorok kategóriáit tartalmazza. A tengelyfeliratok jelzik azokat a csoportokat, amelyekhez az értékeket ábrázoljuk. Forrása a táblázat első sora.

Értéktengely

az értékek megjelenítésére szolgál, a pontos értékek leolvasását teszi lehetővé.

Adatpont

a táblázat egyetlen cellájában tárolt adat, amely megjelenik a diagramon.

Adatsor

a táblázat egy sorában található adatpontok. Vonaldiagramoknál ezeket vonalak kötik össze

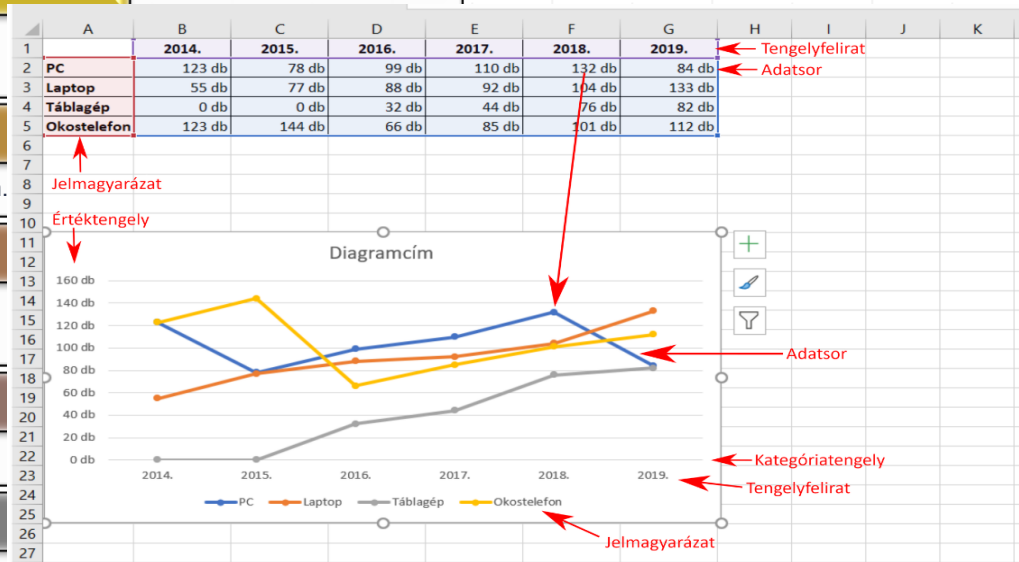
Diagramcím

a diagram címe

Diagramterület

a diagram háttere

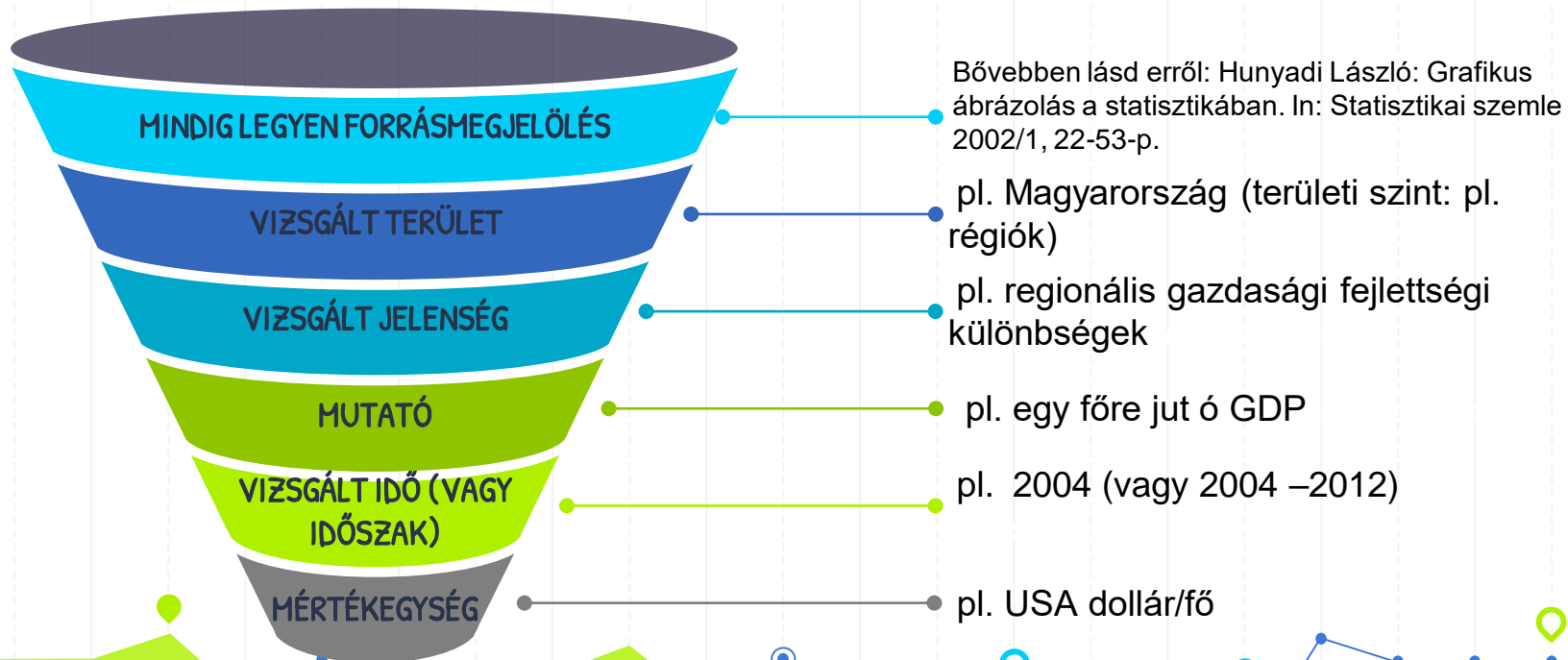
A diagram részei



Forrás: OH (és jogelődjei) / Gál Tamás

A grafikus ábrázolás alapelvei

Minden lényeges információ rajta legyen (ha lehet, a címben)



Ritkán szerepel a cím magán az Excel ábrán (ha mégis, az sem hiba)

A grafikus ábrázolás alapelvei

Az ábra legyen áttekinthető, azt és csak azt mutassa, amire szolgál.

Mondjunk nemet a szép, de értelmetlen formákra, feleslegesen kevert színekre és árnyalatokra, a tárgy szempontjából lényegtelen, sokszor zavaró formai megoldásokra. Ügyelni kell arra is, hogy a színes gépi ábrák fekete-fehér nyomtatásban olyan árnyalatokat kaphatnak, amelyek az egyébként jól megkülönböztethető színeket összemoszák.

Az ábra legyen célorientált és homogén, lehetőleg egy jelenséget ábrázoljon.

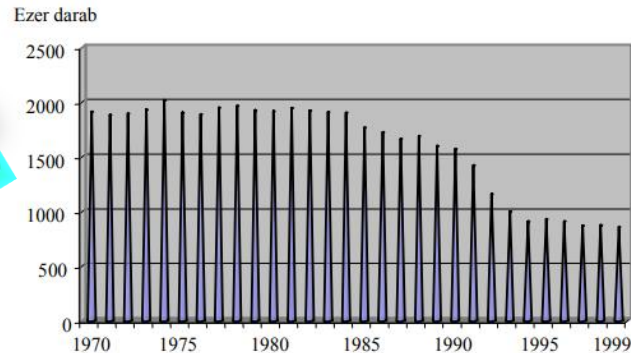
Minimális követelmény:

• a pontos cím, a tengelyek mértékegységének és skálájának egyértelmű megjelölése, a forrás megadása, valamint a tömör, de a szöveges rész alapján érthető jelmagyarázat

szöveg nélkül is megállja a helyét, de: nem helyettesítheti az elemzést! (minden ábrához legyen szöveg)

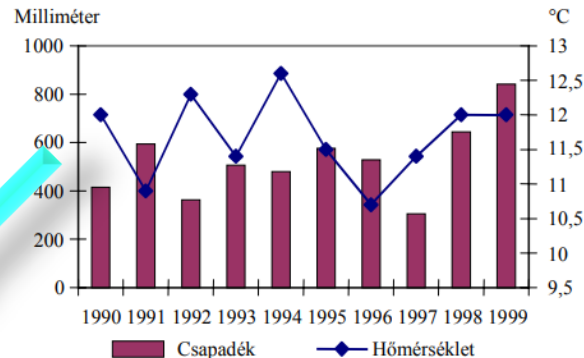
- Az ábra túl sok információt hordoz ezért áttekinthetetlen.
- A két tengely eltérő skálája nem segíti a jelenség jobb, gyorsabb megértését.

1. ábra. Magyarország szarvasmarha-állománya, 1970–1999



Adatforrás: Magyar statisztikai évkönyv, 1999 (2000). Központi Statisztikai Hivatal, Budapest.

2. ábra. Az éves átlaghőmérséklet és az évi csapadék mennyisége Budapesten



Adatforrás: Magyar statisztikai évkönyv, 1999 (2000). Központi Statisztikai Hivatal, Budapest.

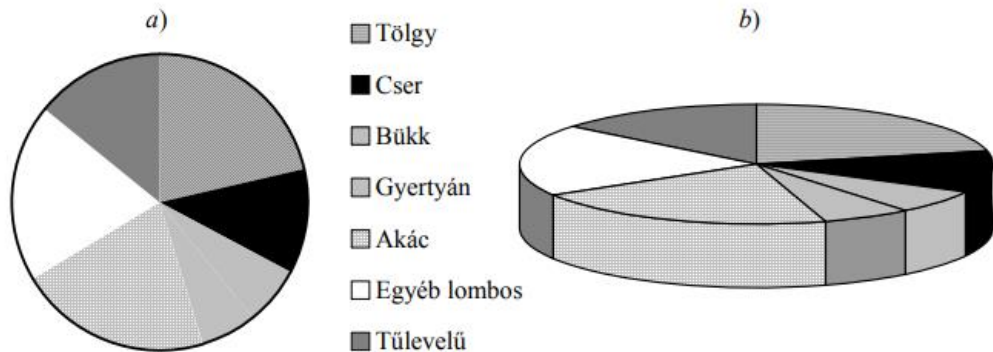
A grafikus ábrázolás alapelvei

Az ábra legyen a lehető legegyszerűbb.

Napjainkban elterjedtek a háromdimenziós (3D) ábrák, melyek szebben mutatnak, mint kétdimenziós párjaik.

Ha azonban a harmadik dimenziónak semmiféle funkciója nincs, ez az ábrázolás csak bonyolítja az áttekintést, elvonja a figyelmet, akadályozza a jelenség gyors felismerését, ezért nemcsak hogy felesleges, de kifejezetten káros is.

3. ábra. A magyarországi erdőállomány fajta szerinti megoszlása 1999-ben



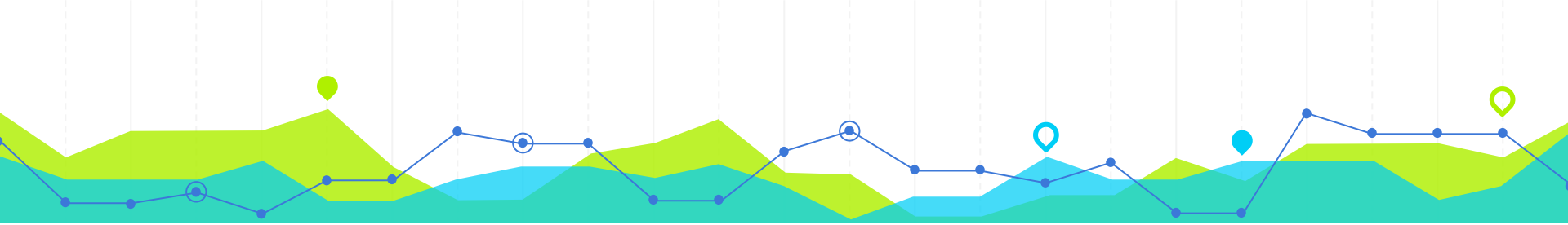
a b) kördiagram feleslegesen tartalmazza a harmadik dimenziót (a korong magasságát), és ezzel megosztja a figyelmet

kifejezetten zavaró az, hogy nem szimmetrikus helyzeténél fogva az akác arányát látszólag nagyobbak, az egyéb lombos és a cser arányát kisebbnek tünteti fel a valóságnál

•térhatás kiemelésére hivatott árnyékolás alkalmazat ad az egyes fajták jelölésének összekeverésére

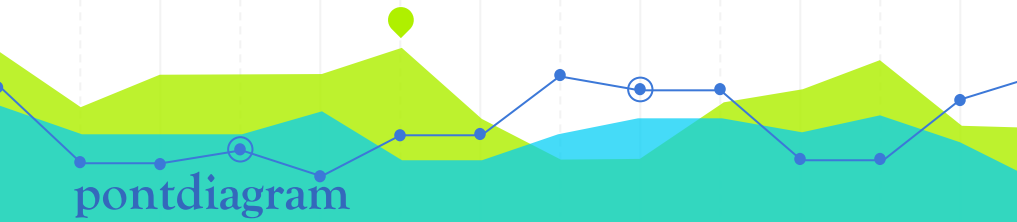
Adatforrás: Magyar statisztikai évkönyv, 1999 (2000). Központi Statisztikai Hivatal, Budapest.

Az a) kördiagramon jól látható a fő fajták (tölgy, akác, egyéb lombos) túlsúlya, és egymáshoz viszonyított arányuk is világosan leolvasható az ábráról.



Melyik típusú diagramot válasszam?

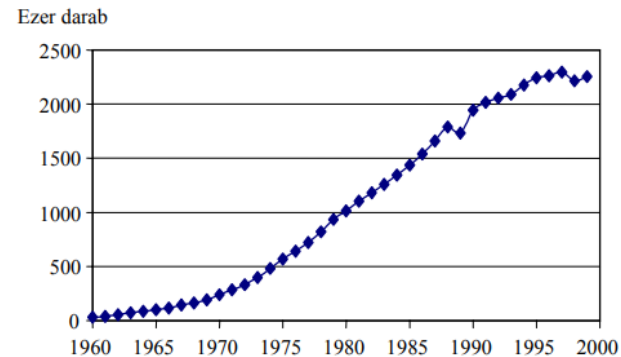
Diagram típusok	Mikor használjuk
Vonaldiagram	Mennyiségi változók esetében, főként trendek, időbeli változások vizsgálatára alkalmazzuk.
Pontdiagram	Mennyiségi változók esetében, főként, amikor két változó lineáris kapcsolatát vizsgáljuk.
Oszlopdiagram	Minőségi változók esetében, főként gyakoriságok ábrázolására használjuk, amikor két változót szeretnénk összehasonlítani.
Kördiagram	Minőségi változók esetében, főként nominális mérési szintű változókra alkalmazzuk, amikor gyakoriságot, illetve a változó összetételét szeretnénk ábrázolni.



az egyes pontokat egyenesekkel össze is lehet kötni. Ez egyrészt utalhat arra, hogy a jelenség a megfigyelt időpontok közt is létezik, de szolgálhat egyszerűen arra is, hogy az idősorban meglévő tendencia jobban láthatóvá váljék.

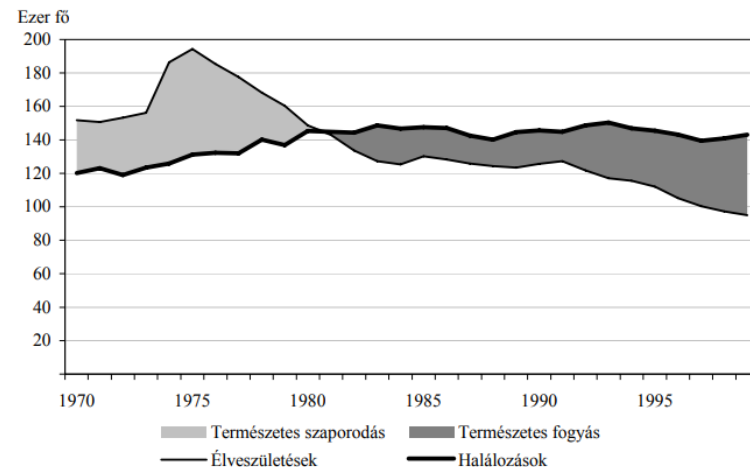
Két vagy több idősor esetén egyes esetekben szemléletesen lehet ábrázolni kapcsolatukat, mint ahogy azt a 8. ábra is mutatja. Az ábrán a két idősor (az élveszületések és a halálozások száma) egyenlegeként előálló árnyékolt terület igen szemléletesen mutatja a népesség szaporodásának, illetve fogyásának fő összetevőit.

5. ábra. A magyarországi személygépkocsi-állomány alakulása, 1960–1999 (év végi állomány)



Adatforrás: Magyar statisztikai évkönyv, 1999 (2000). Központi Statisztikai Hivatal, Budapest.

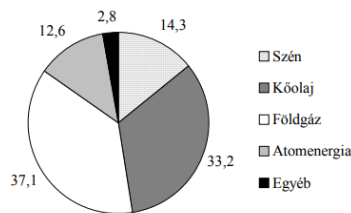
8. ábra. Élveszületések, halálozások, szaporodás és fogyás, 1970–1999



Kördiagram (tortadiagram)

Minőségi ismérvek szerinti megoszlások (például pártpreferenciák, foglalkozások, ágazati megoszlások stb.) ábrázolására legcélszerűbb a kördiagram használata. százalékos megoszlást mutat egyes adathalmazok esetén.

9. ábra. Az egyes energiahordozók százalékos aránya az összes felhasználásban, 1999

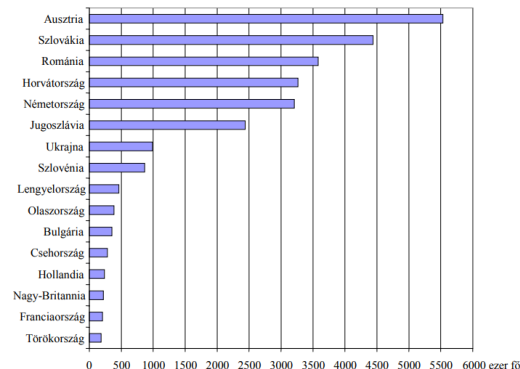


Adatforrás: Magyar statisztikai évkönyv, 1999 (2000). Központi Statisztikai Hivatal, Budapest.

szalagdiagram

Ha az ismérvváltozatok száma nagy, a kördiagram áttekinthetetlenül sok szeletből áll; ilyen esetekben elkerülhetetlen az oszlopos ábrázolás. Ekkor azonban – inkább a fekvő oszlopos ábrázolás (szalagdiagram) javallott

10. ábra. A Magyarországra néhány európai országból érkezet látogatók száma 1999-ben

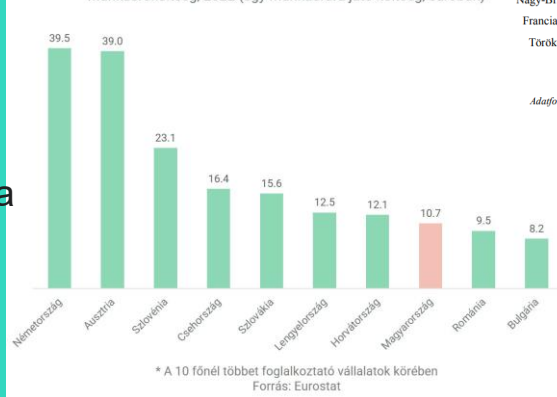


Adatforrás: Magyar statisztikai évkönyv, 1999 (2000). Központi Statisztikai Hivatal, Budapest.

oszlopdiaqram

Leggyakrabban a kvalitatív változó ábrázolására használt diagram a leíró statisztikában. A figyelmet a kategóriák gyakoriságára irányítja, ahol az oszlopok magassága vagy hosszúsága jelöli a gyakoriságot.

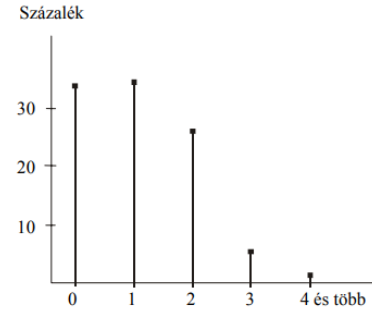
Munkaerőköltség, 2022 (egy munkaóra jutó költség, euróban)*



pálcikadiagram

Mennyiségi sorok esetén leggyakoribb a nagyság szerinti megoszlás ábrázolása. Amennyiben a csoportosító ismérv diszkrét és kevés változata van, célszerű forma a pálcikadiagram (13. ábra)

13. ábra. A családok gyermekszám szerinti megoszlása, 1996

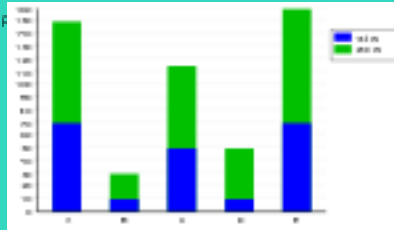
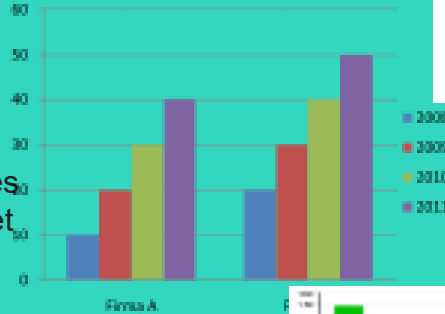


Adatforrás: Magyar statisztikai évkönyv, 1999 (2000). Központi Statisztikai Hivatal, Budapest.

Oszlopdiaagram- változatok

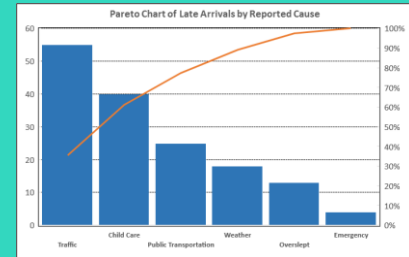
Csoportosított oszlopdiaagram

Az egyes kategóriákat összetevőikre bontjuk és kategóriánként csoportosítjuk. Az összetevőket különféle színezéssel különböztethetjük meg.



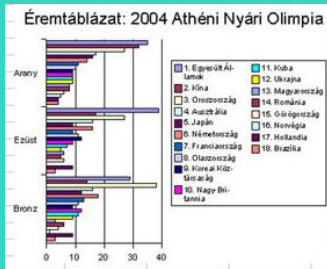
Pareto diagram

a változók gyakoriság szerint vannak rendezve, bal oldalt a leghosszabb oszloppal.



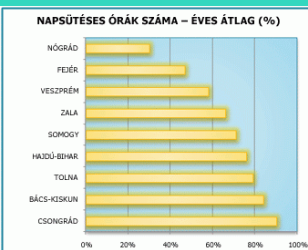
Rétegelt (halmozott) oszlopdiaagram

a kategóriákat összetevőikre bontjuk, de ez esetben nem egymás mellé hanem egymás tetejére helyezük az oszlopokat.



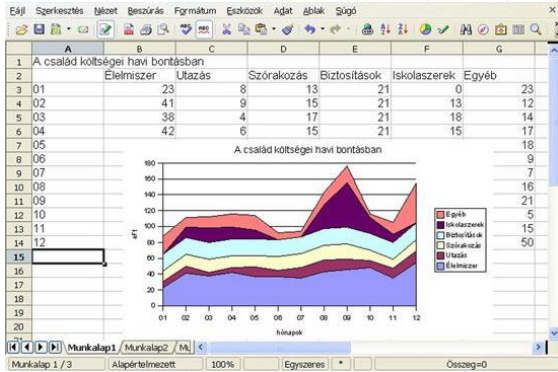
Sávdiaagram

Függőlegesen megrajzolt oszlopdiaagram



Egyéb diagram típusok

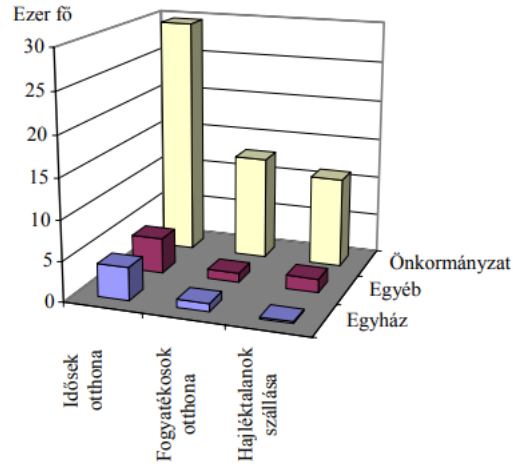
Területdiagram



Az értékeket egy a kategóriatengelyen álló sokszög felső pontjaiként ábrázolja. A területdiagram két egymást követő kategória közötti változás vagy eloszlás kiemelésére alkalmas.

3D hisztogram

22. ábra. A szociális intézményekben gondozottak száma intézménytípus és fenntartó szerint, 1999



Adatforrás: Magyar statisztikai évkönyv, 1999 (2000). Központi Statisztikai Hivatal, Budapest.

piktogram (piktográf)

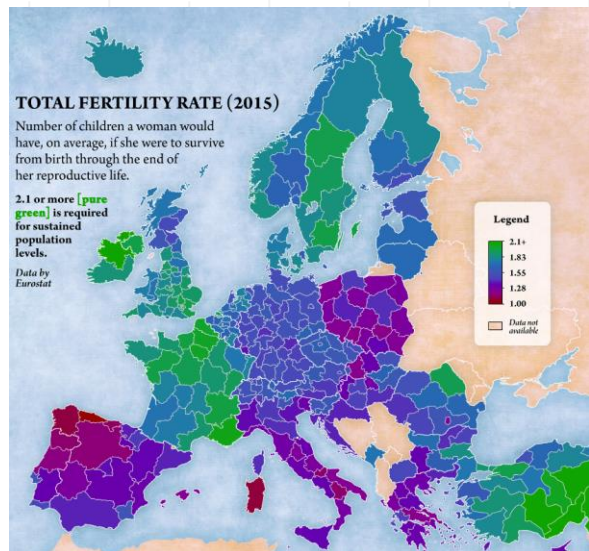
A piktogram a legszélesebb közvélemény számára közvetíthet egyszerű statisztikai információkat valamilyen figurális ábrázolás révén



a legkevésbé rugalmas, és a legkevésbé alkalmas a jelenségek árnyalt ábrázolására.

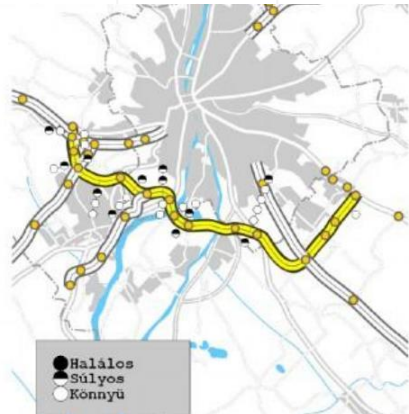
Kartogram

a színezés jelzi az adott mutató nagyságát



Ponttérkép

egyfajta sűrűségi mutató, a pontok mennyisége jelzi a mutató nagyságát



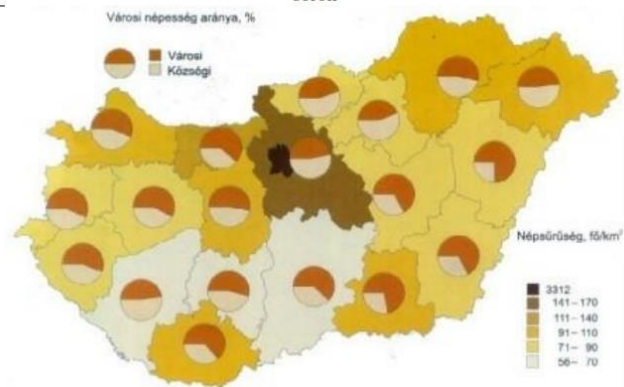
3. ábra: M0 autóbaleseti ponttérkép 2007. I-VIII.hó Forrás: www.police.hu

Térképen alapuló ábrázolás

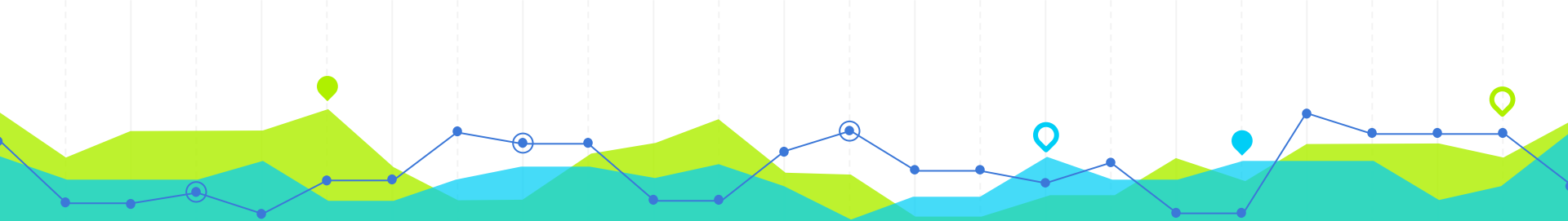
Kartodiagram

A térkép különböző pontjain diagramokat helyeznek el, ezzel szemléltetve a mutató nagyságát

A statisztika alapfogalmai, feladatai, táblázatok, grafikus ábrázolás. Leíró statisztika, osztályozás, gyakorisági sorok



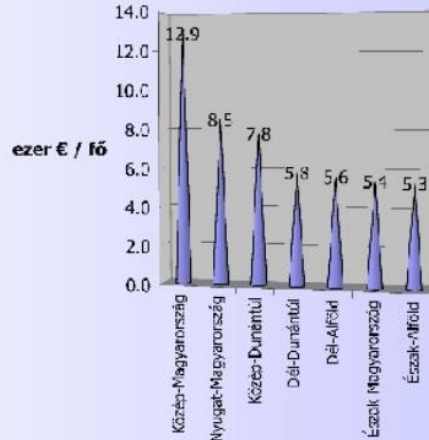
2. ábra: Népsűrűség és városi népesség aránya 2002.január 1. Forrás: sdt.sulinet.hu



az adatok grafikai ábrázolásának tipikus hibái

- Kerüljük a színes vagy mintás háttérrel
- Elegánsabb a fehér háttér
- Régi Excel: alapbeállításban szereplő szürke háttér előtt kevésbé látszanak a szürkés-kék jelölők
- Nyomtatásnál felesleges „festékpazarlás”
- Kúpok helyett oszlopok
- 3D helyett 2D

Regionális gazdasági fejlettségi különbségek Magyarországon 2004-ben (NUTS2-es szinten, egy főre jutó GDP alapján)

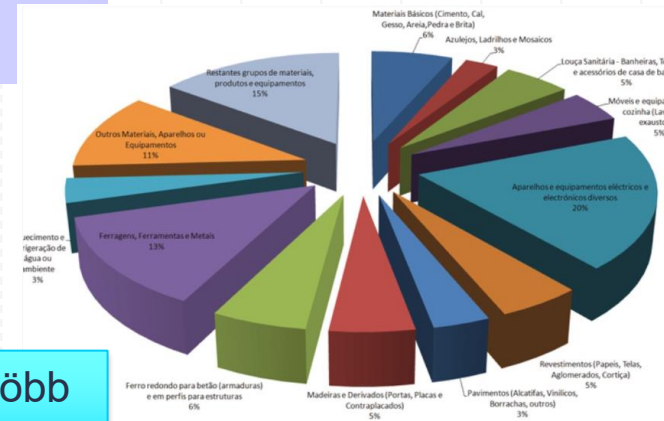


Hibás ábra!

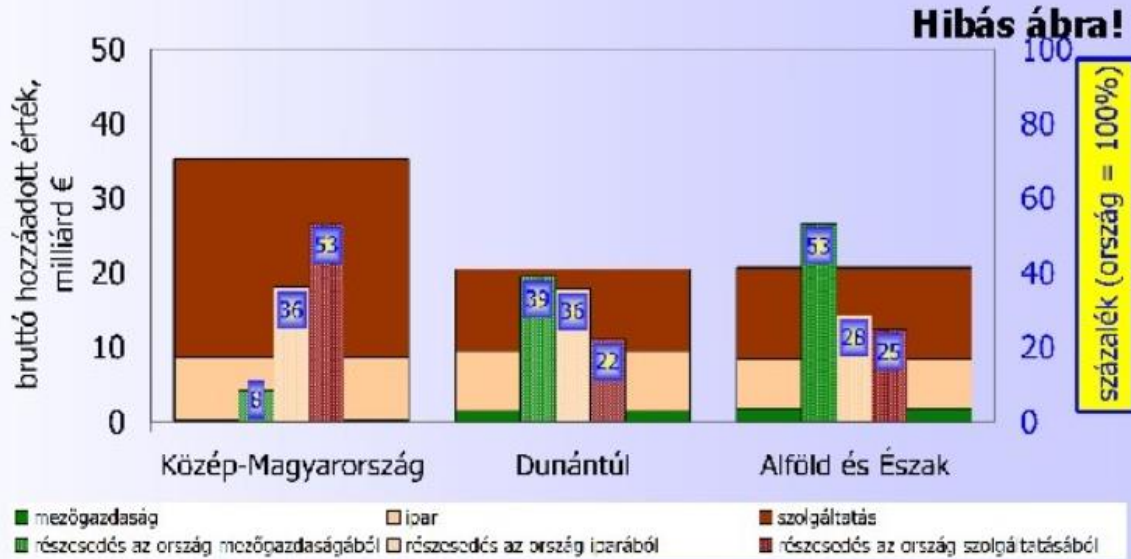
Forrás: EUROSTAT

Optikailag semleges legyen, kerüljük a térhatást (3D -t)!

Rossz kördiagram (ne legyen 3-4-nél több körcikk)



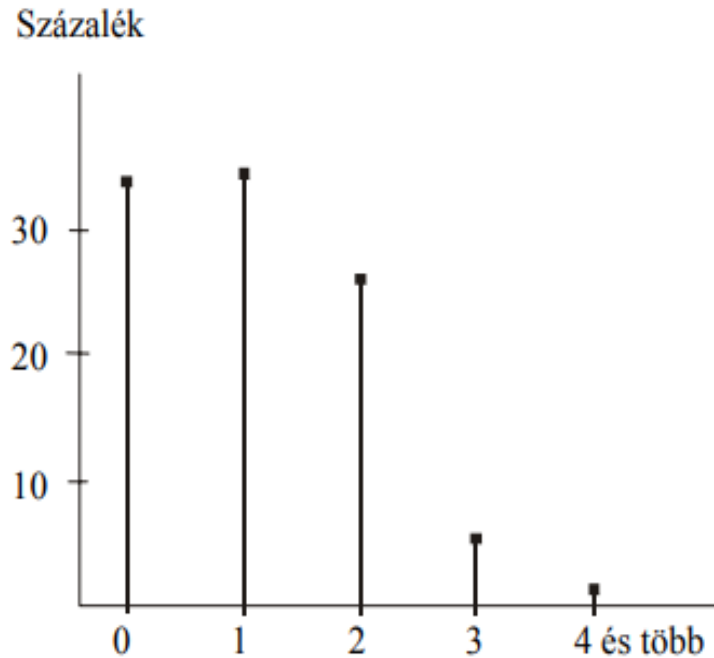
A magyar régiók gazdasága ágazonként, illetve a régiók aránya az ágazatokból (NUTS1-es szinten, bruttó hozzáadott érték alapján), 2004-ben



Forrás: EUROSTAT

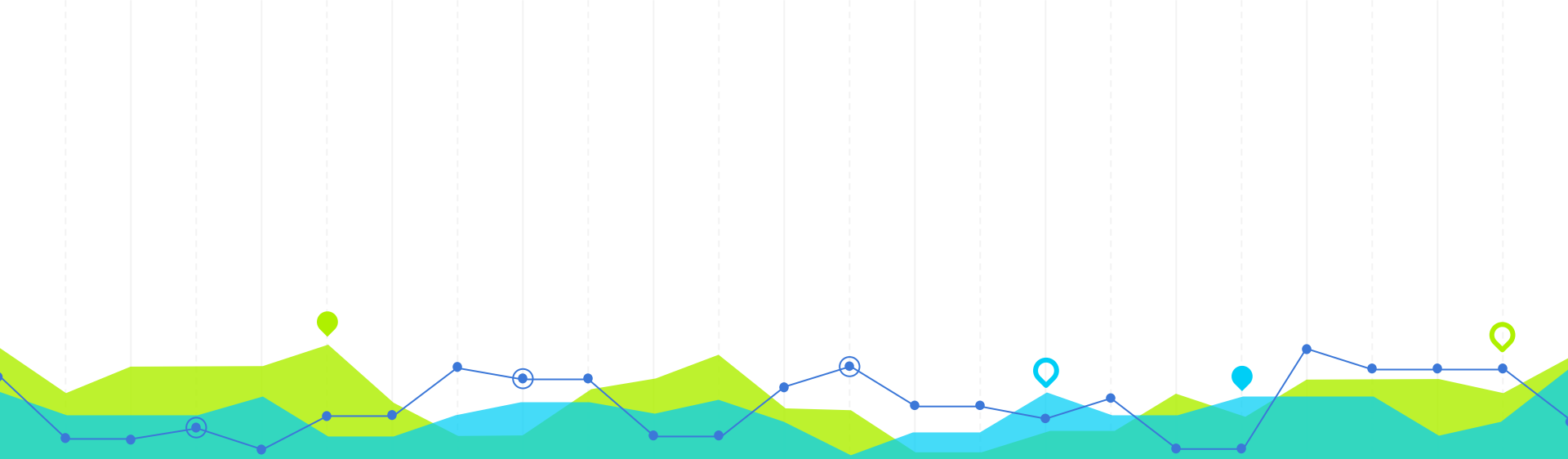
Célorientált, áttekinthető legyen, ne túlságosan összetett (inkább külön diagramokon)!

13. ábra. A családok gyermekszám szerinti megoszlása, 1996



Egyszerű
legyen, ne
túldizájnolt!

Adatforrás: Magyar statisztikai évkönyv, 1999 (2000). Központi Statisztikai Hivatal, Budapest.



„Puha adatok” elemzése

4

Adatgyűjtés szakaszában célszerű hozzálátni az adatok tisztításához, előkészítéséhez, elemzéséhez – változtatás lehetősége!

3.1. az interjú leírása és átírása

ha lehet, a kutató maga írja le!

3 formája:

1. vázlat az interjú szövegéből

2. interjú szó szerinti szövege

3. további információk közlése (idő, megjegyzések, gesztusok, utalások más interjúkra...)

1. Vázlat az interjú szövegéből

- „1. Üdvözlés, bemelegítés
 2. A téma ismertetése
 2. 1. Az interjúalany bemutatja az iskolát, a helyzetet
 2. 2. Rátér arra, hogy a roma tanulókkal mi a céljuk
 2. 3. Elmondja, hogy az iskola feladata a társadalmi integráció segítése
 2. 4. Rávilágít, hogy azért gyengék a cigány származású diákok a tanulásban, mert sok kudarcot éltek át
 3. Az iskola elhelyezkedése a város oktatásában
 3. 1. Földrajzi elhelyezkedés
 3. 2. Elmeséli annak történetét, hogyan lett az intézmény „kiszegítő” iskola
 3. 3. Kiderül, hogy komoly problémát jelent a busz, a gyerekek nehezen érik el az iskolát
 4. A továbbtanulási lehetőségekről
 4. 1. Ebből az iskolából legfeljebb speciális szakmunkásképzőbe lehet eljutni
 4. 2. Fontos az európai uniós program, mert a gyerekek ennek segítségével sokfelé eljutnak
 5. A gyógypedagógus szakmájáról, munkájáról
 5. 1. Ők maguk éppúgy kiszolgáltatottak, mint a diákjaik
 5. 2. Kényszerből került a pályára, de időközben megszerette, és kihívásnak tekinti
 5. 3. Fontos, hogy szeretettel kell közelíteni a gyerekekhez
 6. A szülők és az iskola
 6. 1. Gondoltak arra, hogy a szülőkkel együttműködjenek, például a költözéskor hívták be őket
 6. 2. Szerinte a gyerekek által kellene a szülőket is nevelni”
- (Az interjú egy Hajdú-Bihar megyei, igen szegény általános iskola igazgatójával készült 2003-ban.)

2. Interjú szó szerint

„K.: Gondoltak-e arra, hogy behívják azokat a szülőket, akiknek a gyerekeik ide fognak járni, hogy vegyenek részt itt a...
V.: Hát végül is a költözésben, a fizikai... a szekrények, a padok... azt cigány szülők, közmunkások, ők költöztettek el bennünket. A meszelésben meg a takarításban nem szóltunk nekik, és utána mondták. Nem sokan, de mondták, hogy miért nem szóltunk. Szóval érzik ők azt, hogy ha jó szándékkal közelítünk, sokszor kell velük csatázni, mert azért, eleve más a szemléletük, ezek az emberek nem dolgoznak, meg érdemes lenne azt végignézni a B. utcai szociális épületeket, meg miegyebek. Igazából én azt gondolom, hogy mindenről lehet beszélni fehér asztal mellett, de menjenek ki, és nézzék meg, hogy élnek azok az emberek. Én mindig azt mondom, hogy ha egyáltalán csak betámo-lyognak ebbe az iskolába, az is egy csoda. Ahogy ők élnek, ahhoz képest. Ott kéne először is változtatni, de az a költséges oldala ennek. Szóval ennek az egész romakérdésnek egy borzalmas rákfeneje, hogy ezek az emberek kies-tek a munkából, összezsúfolva élnek, szegények mit tudnak csinálni, egymással a hűlyeséget erősíteni. Unalmában, mert hát olvasni sem nagyon tud, a tévén, a zenén meg az egymás ökörködésén, meg egymást hogy verjék át, meg kártyáznak. Mert akkor is azt mondom, hogy a szocializmusban ilyen szempontból másabbul adta magát az ő beilleszkedésük, mert mindegy, hogy az egy töltelék munkahely volt, és egy seprűvel próbálták a főnök előtt egész nap elkeverni valahogy, de akkor is közösségben volt, „mosdjál már meg, mert büdös vagy!”, próbálták azért úgy idomulni ahhoz a dologhoz. Meg lehet figyelni például, amikor én még idekerültem az iskolába, a tanévzáró is majdnem ugyanolyan nap volt, mint a többi, mert nem öltöztek fel a gyerekek, szedett-vedettül, hogy egy virágot hozzanak, mert nőnap van vagy pedagógusnap van? Nem. Most igaz, hogy egy műanyag virágot, de a tanévzáróra mindig elhozzák a virágot, a legszegényebb... szóval azért én mondom, hogy nagyon szívesen meg hosszú időn át azért lehet őket... meg ilyen, hogy szülői értekezlet: nem járnak, de például most a Phare program keretében járt a tánc-tanár, táncruhát kölcsönöztünk olyan igazít, olyan szépet, pirosat, meg rózsásat, meg izé, és arra bejöttek megnézni. Én azt mondom, hogy a gyerekek által kellene a szülőket is valamelyest nevelni.

3. Részletes leírás

69:13

K: Gondoltak-e arra, hogy behívják azokat a szülőket, akiknek a gyerekeik ide fognak járni, hogy vegyenek részt itt a...
V: Hát végül is a költözésben, a fizikai... a szekrények, a padok... azt cigány szülők, közmunkások, ők költöztettek el bennünket. A meszelésben meg a takarításban nem szóltunk nekik és utána mondták. Nem sokan, de mondták, hogy miért nem szóltunk. Szóval érzik ők azt, hogy ha jó szándékkal közelítünk, sokszor kell velük csatázni, mert azért, eleve más a szemléletük, ezek az emberek nem dolgoznak, meg érdemes lenne azt végignézni a B. utcai szociális épületeket, meg miegyebek.

Kicsit meg-
áll, nagy le-
vegőt vesz

Igazából én azt gondolom, hogy mindenről lehet beszélni fehér asztal mellett, de menjenek ki, és nézzék meg, hogy élnek azok az emberek. Én mindig azt mondom, hogy ha egyáltalán csak betámo-lyognak ebbe az iskolába, az is egy csoda. Ahogy ők élnek, ahhoz képest. Ott kéne először is változtatni, de az a költséges oldala ennek.

Szóval ennek az egész romakérdésnek egy borzalmas rákfeneje, hogy ezek az emberek kies-tek a munkából, összezsúfolva élnek, szegények mit tudnak csinálni, egy-mással a hűlyeséget erősíteni. Unalmában, mert hát olvasni sem nagyon tud, a tévén, a zenén meg az egymás ökörködésén, meg egymást hogy verjék át, meg kártyáznak. Mert akkor is azt mondom, hogy a szocializ-musban ilyen szempontból másabbul adta magát az ő beilleszkedésük, mert mindegy, hogy az egy töltelék munkahely volt, és egy seprűvel próbálták a főnök előtt egész nap elkeverni valahogy, de akkor is közösségben volt, „mosdjál már meg, mert büdös vagy!”, próbálták azért úgy idomulni ahhoz a dologhoz. Meg lehet figyelni például, amikor én még idekerültem az iskolába, a tanévzáró is majdnem ugyanolyan nap volt, mint a többi, mert nem öltöztek fel a gyerekek, szedett-vedettül, hogy egy virágot hozzanak, mert nőnap van vagy pedagógusnap van? Nem.

Felgyorsul
a beszéde

Most igaz, hogy egy műanyag virágot, de a tanévzáróra mindig elhozzák a virágot, a legszegényebb... szóval azért én mondom, hogy nagyon szívesen meg hosszú időn át azért lehet őket... meg ilyen, hogy szülői érte-kezlet: nem járnak, de például most a Phare program ke-retteben járt a tánc-tanár, táncruhát kölcsönöztünk olyan igazít, olyan szépet, pirosat, meg rózsásat, meg izé, és arra bejöttek megnézni. Én azt mondom, hogy a gyere-kek által kellene a szülőket is valamelyest nevelni.

A közoktatási törvény szerint a szülőnek ki-emelten joga van az is-kolával való kapcsolat-tartáshoz. Ennek milyen a módja ebben az intéz-ményben? Itt sincsenek bevonva, ahogyan példá-ul D. településen sem.

Ősszeköti az iskolai problémákat a családi problémákkal, noha nem mentes az ítéletek-től, de megkísérli meg-érteni a szülőket.

„Romakérdés” – egyben mondja, kicsit hasonlít a hajdani *Kékfény* tévémű-sor megfogalmazásá-hoz.

Saját munkája sikerének is tekinti a virágot: hivat-kozik a Phare program-ra, utána kell nézni, mit mond róla, hogyan zaj-lott az iskolában a prog-ram.

3. A „puha” adatok elemzése

3.2. adatok rendezése

Összegyűlt kutatási nyersanyagok rendszerezése

adatok formátuma

Legyen egységes! (kazetta, feljegyzések... legyen gépen rögzítve, filenév egyértelmű...)

adatok csoportosítása

minden adatot fel kell címkézni! (interjú: kódszám vagy monogram; helyszín, más szempont szerint csoportosítani)

adattisztítás

a későbbi elemezhetőség állapotába juttassuk az adatokat! Más is tudja esetleg elemezni

biztonsági intézkedések

Mentés készítése, más helyen tárolás!!!

3. A „puha” adatok elemzése

1 a témára vonatkozó utalások kiszűrése

2 adott szövegrészek kigyűjtése

3 azok csoportosítása

4 táblázat elkészítése a csoportok és válaszaik szerint

Interjúk alapján tipológia készíthető!

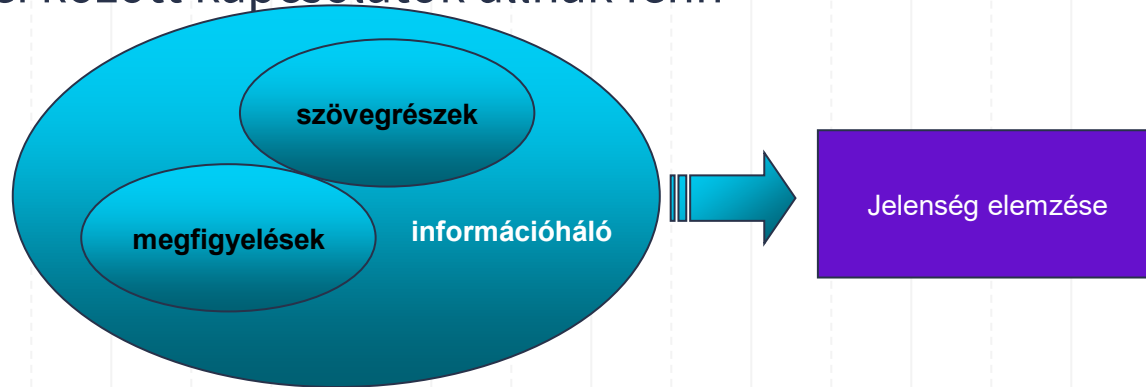
3. A „puha” adatok elemzése

3.4. esettanulmány

Célravezetőbb, ha több interjút készítünk

Egymás mellé helyezve ezek szöveghálót alkotnak

A háló elemei között kapcsolatok állnak fenn



Esettanulmány példája: Fellegi Borbála – Ligeti György (2003): Hátrányos helyzetűek a közoktatásban; Kutatási beszámoló Kurt Lewin Alapítvány (elérhető: <http://www.kla.hu/referenciak/kutatas/>)

A tartalomelemzés szakaszai



1

A kódolás szakasza

- **Ebben a fázisban a szöveg egyes részeit (pl. a fókuszcsoport vagy az egyéni mélyinterjú szöveges átiratának „kódolási egységeit”) előre megállapított kategóriákhoz soroljuk.**
- **Olyan szavak, mondatok jelentése kódolható, amelyek a szövegben ténylegesen megjelennek.**

Kódolás és kategóriák

- A kódolás az a folyamat, amelynek révén a nyers szövegeket módszeresen nagyobb egységekbe soroljuk, s így lehetővé tesszük a tartalom valóban lényeges leírását és összefüggéseinek megragadását.
- Például a „*haragszom*”, „*nagyszerű örömben volt részem*”, „*bosszant*”, „*szomorúságot tükröz*”, „*egyszer nagyon be fogok gurulni*”, „*örültem neki*”, „*mindig dühös leszek, ha ilyet látok*”, „*bepöccenek, ha azt látom, hogy*”, „*félelmet keltenek bennem*” szóösszetételeket az „*érzelemnyilvánítás*” nagyobb egységébe sorolhatjuk, s e címszóból alkotott betűkombinációval kódolhatjuk (é. ny. = érzelemnyilvánítás).
- Két típus:
 - az előre meghatározott kategóriák
 - a vizsgálat során kibontakozó kategóriák

2

Az elemzés szakasza

A kódolt tartalmakat dolgozzuk fel

Vizsgáljuk az egyes előfordulási gyakoriságok számát

A kódok együttes előfordulása

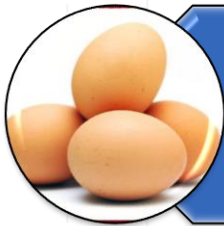
A nem jelenlévő fogalmak is felszínre kerülnek (látens tartalom)

Két-három kódnak az együttes előfordulása „jelentéstöbbletet” hoz létre

tartalmi mutatóvá válhat valaminek a hiánya

3

Az értelmezés szakasza



A tendenciaszerű együttes előfordulások a szövegben törvényszerűségekre engednek következtetni, amelyek értelmezhetővé válnak.



Ha egy dokumentumban nem szerepel egy olyan kód (által jelzett fogalom), amelyet ott elvárnánk – mert pl. korábbi hasonló dokumentumokban mindig szerepelt –, akkor egy nem jelenlevő tartalomra, rejtett, látens mozzanatra bukkantunk, amely esetleg éppen „sokatmondó hiánya” miatt lehet meghatározó.



*Minél izébb valami, annál jobban
akármilyen bármi más, ami
amúgy nem is lenne nagyon
semmilyen se.*



SE feedback QR kóð