

Elemi statisztika fizikusoknak

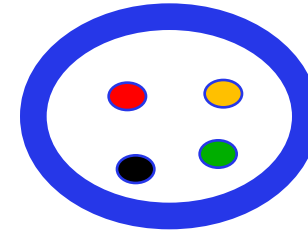
Pollner Péter
Biológiai Fizika Tanszék

pollner@elte.hu

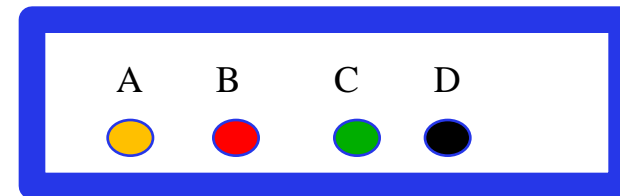
Összefoglalás - A mérések szintjei

2 . oldal

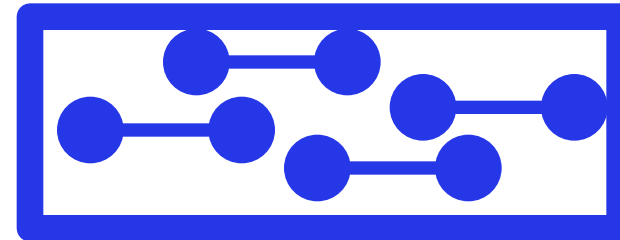
❖ **Nominális** – csak kategóriák



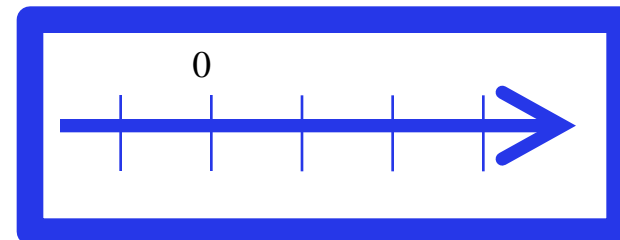
❖ **Ordinális** – kategóriák és
rendezhetőség



❖ **Intervallum** – különbségek
de nincs természetes 0 pont



❖ **Arány** – a különbségek és arányok értelmesek
és létezik természetes kezdőpont



Buktatók

3 . oldal

- ❖ Hibás minták
- ❖ Kicsi minták
- ❖ Félrevezető ábrák
- ❖ Becsapós ábrák
- ❖ Játék a százalékokkal
- ❖ Beugrató kérdések
- ❖ A kérdések sorrendje
- ❖ Válasz megtagadás
- ❖ Korreláció és kauzalitás
- ❖ Önérdekeltség a vizsgálatban
- ❖ Precíz számok
- ❖ Részleges képek
- ❖ Készakart hamisítás

1-4 fejezet

A kísérletek megtervezése

Fő pontok

5 . oldal

- ❖ Ha a mintát nem megfelelő módon gyűjtjük, akkor az annyira használhatatlan lesz, hogy semmiféle statisztikai manipulációval sem tudjuk megmenteni.
- ❖ **A véletlen** tipikusan kritikus szerepet játszik abban, hogy mely adatokat gyűjtsük össze.

Definíció: vizsgálati típusok

6 . oldal

❖ **Beavatkozásos vizsgálat (Experiment)**

valamilyen **kezelést** végzünk és azután megfigyeljük a hatásait a kísérlet tárgyan/alanyán

Pl.: klinikai gyógyszervizsgálat, részecske ütközések a CERN gyorsítójában

Definíció: vizsgálati típusok

7 . oldal

❖ Megfigyeléses vizsgálat (Observational study)

bizonyos jellemző tulajdonságok megfigyelése és mérése anélkül, hogy **megváltoztatnánk** a vizsgálat tárgyát/alanyát

pl.: közvéleménykutatás, csillagászati/asztrofizikai megfigyelések

Definíciók

8 . oldal

A megfigyeléses vizsgálat altípusai

❖ **Keresztmetszeti vizsgálat (Cross Sectional Study)**

Az adatokat egy időpontban mérjük, figyeljük meg és gyűjtjük be. (pl. lehallgatott telefon)

❖ **Utólagos vizsgálat (Retrospective, Case-control Study)**

Múltbéli adatokat használunk. (pl.: az autóbalesetben meghaltak és nem abban meghaltak összehasonlítása)

❖ **Előre tervezett**

(Prospective, Longitudinal, Cohort Study)

Az adatokat a jövőben gyűjtjük, olyan csoportokból, melyek valamilyen közös faktorban megegyeznek. (pl.: a mobil telefont használó és nem használó vezetők csoportjainak összehasonlítása)

❖ Zavar (confounding)

akkor lép fel egy kísérletben, ha a kísérletet végző nem tudja megkülönböztetni az egyes faktorokat

Pl.: Az Eötvös inga érzékeny-e a Nap helyzetére ahogy a Föld forog? Vagy a mérőhelyünk mellett nappal járnak, éjszaka nem járnak a villamosok? A két faktor nem különböztethető meg. Elektron szcintilláció és fogmosás.

Úgy kell a kísérletet megtervezni, hogy ne lépjen fel zavar!

A zavar elkerülése beavatkozásnál 1.

10 . oldal

❖ **Vak vizsgálat (Blinding), duplán vak vizsgálat**

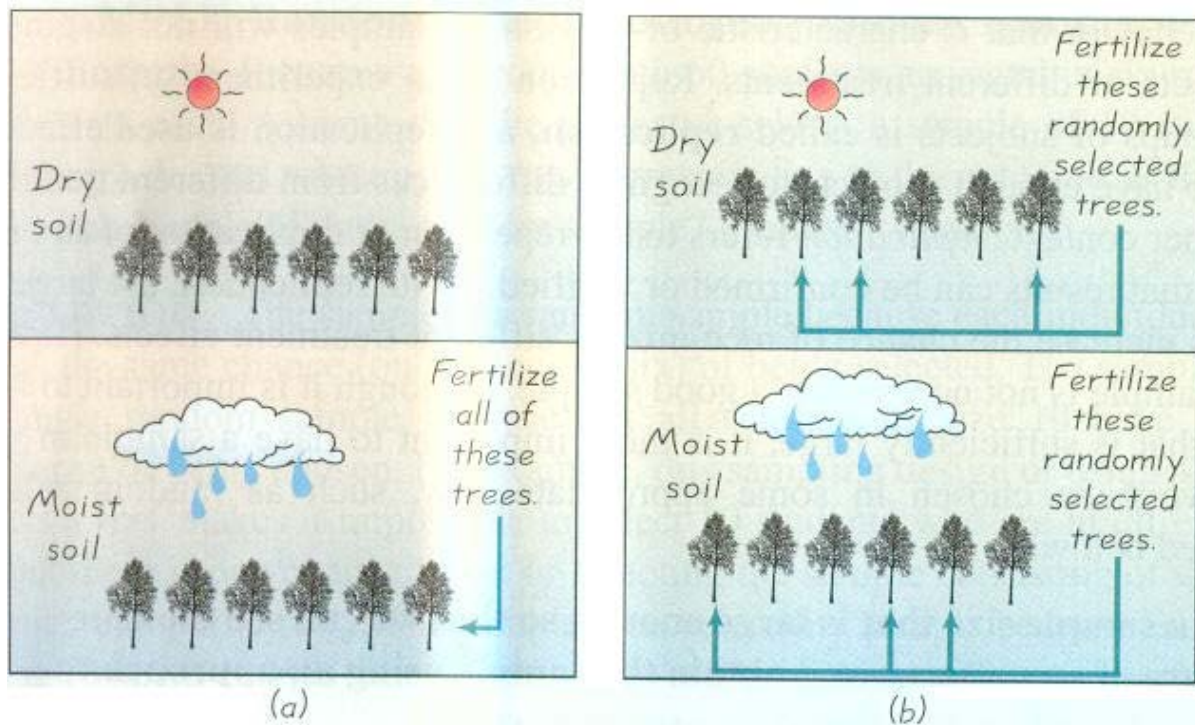
a vizsgálat alanya nem tudja, hogy kezelést kap-e vagy placebót, duplán vak, ha a kísérletező sem tudja (pl.: a gyermekbénulás Salk vakcina kipróbálása az USA-ban 1954-ben)

a kísérleti adatokat független statisztikus csoport értékeli ki

A zavar elkerülése beavatkozásnál 2.

❖ **Blokkosítás** — felosztjuk a populációt

olyan alcsoportokra amelyekben a kísérlet szempontjából fontos tulajdonságok megegyeznek . Mindegyik blokkban véletlenszerűen választjuk ki a kezeltet

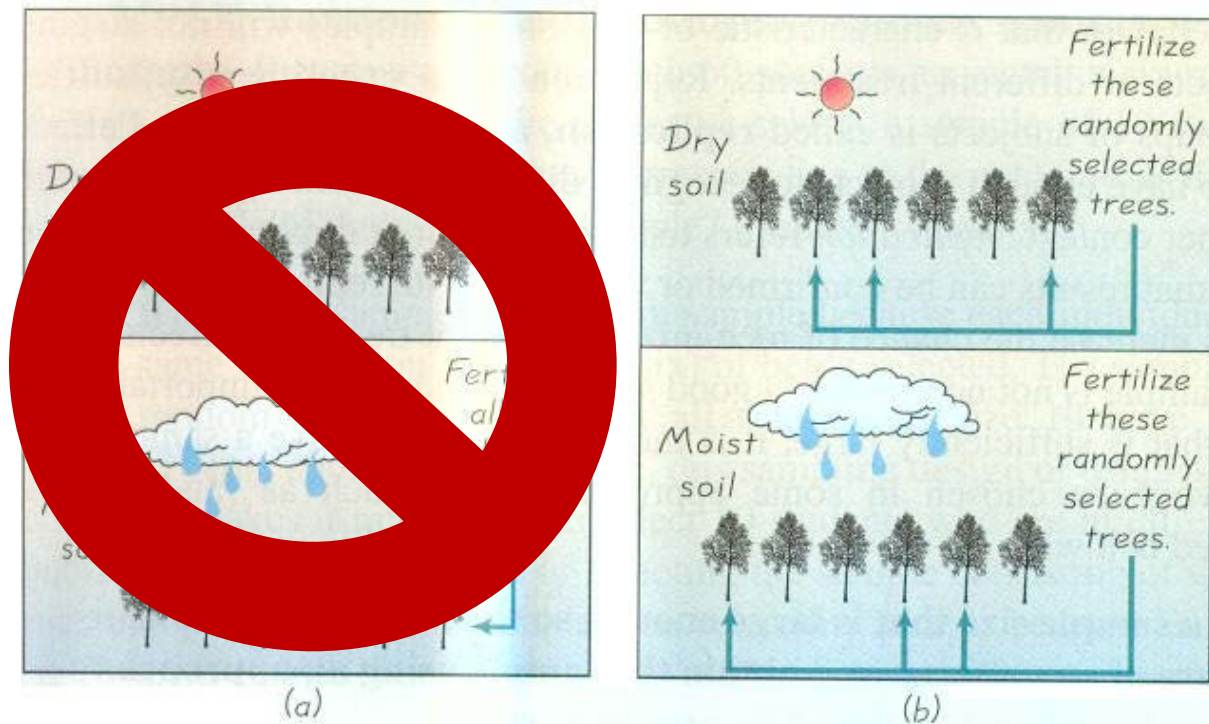


A zavar elkerülése beavatkozásnál 2.

12 . oldal

❖ **Blokkosítás** — felosztjuk a populációt

olyan alcsoportokra amelyekben a kísérlet szempontjából fontos tulajdonságok megegyeznek . Mindegyik blokkban véletlenszerűen választjuk ki a kezeltet

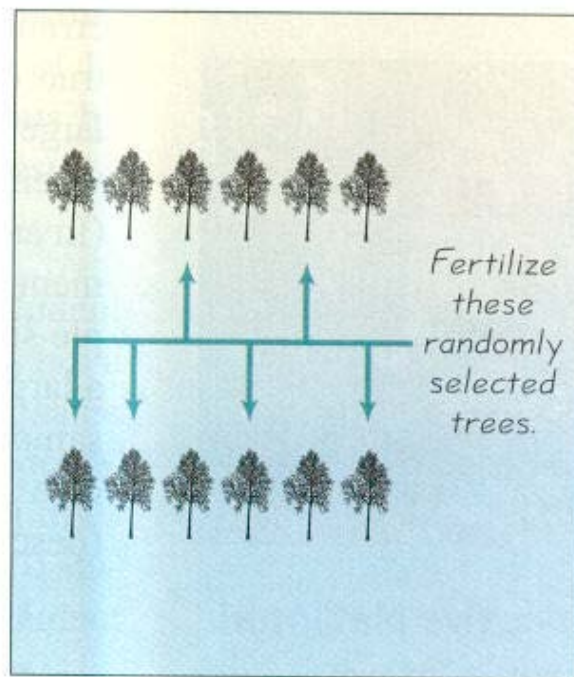


A zavar elkerülése beavatkozásnál 3.

13 . oldal

❖ **Teljesen randomizált (véletlenszerűsített) kísérleti elrendezés**

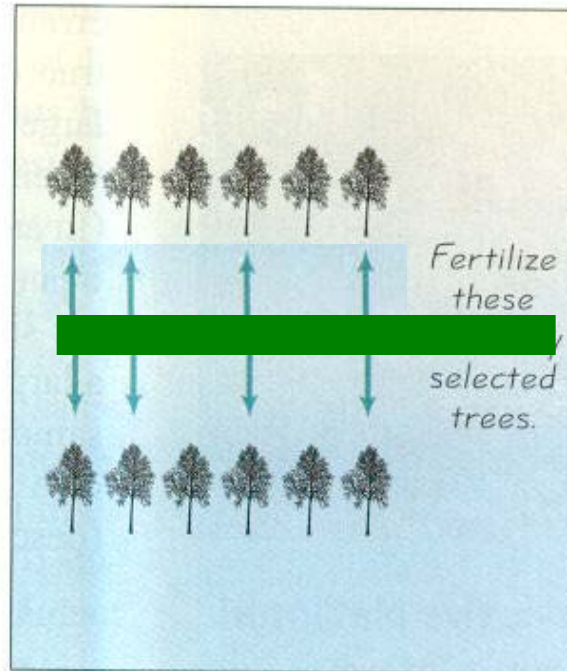
véletlen kiválasztással választjuk ki azokat, akik kezelést kapnak



(c)

A zavar elkerülése beavatkozásnál 4.

14 . oldal



❖ **Szigorúan kontrollált elrendezés**

nagyon körültekintően kiválasztott egyedek

pl, : a fákat koruk, betegségeik, termésátlaguk stb alapján párosítjuk, és ezeket a párokat vizsgáljuk

❖ Véletlen mintavétel

a populáció minden tagjának **ugyanakkora esélye** van arra, hogy a mintába bekerüljön

❖ Egyszerű véletlen mintavétel

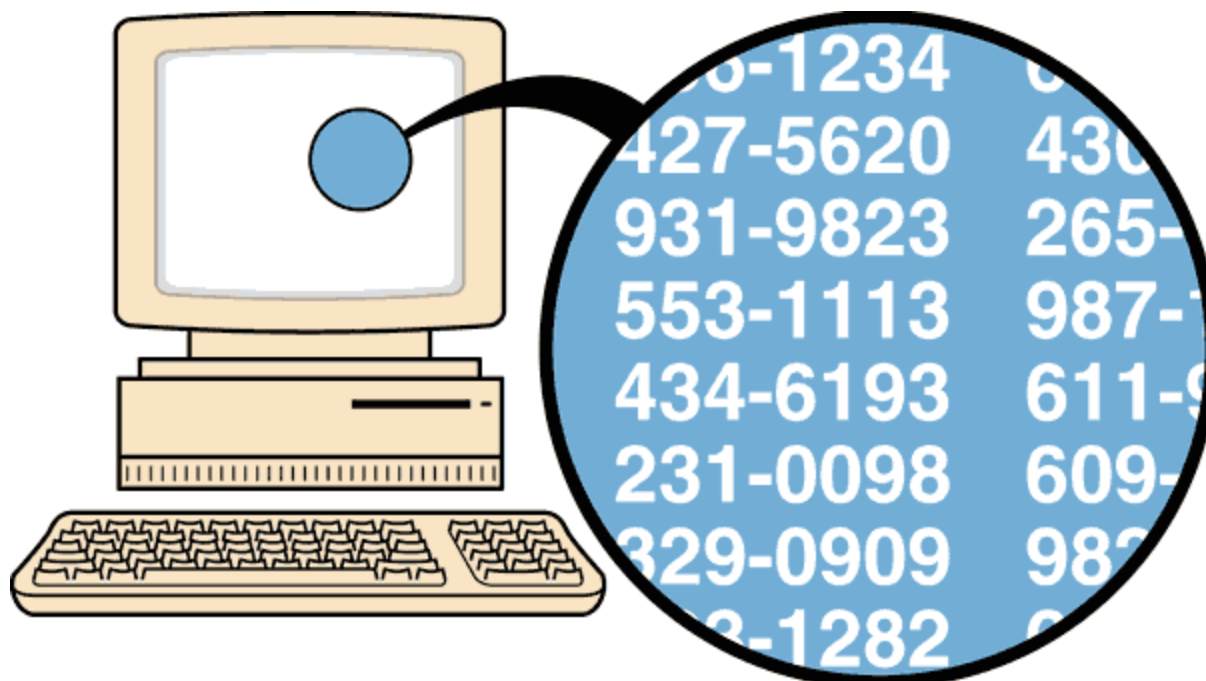
a minta tagjait úgy választjuk ki, hogy bármelyik n hosszúságú mintának ugyanakkora a kiválasztási esélye

❖ Valószínűségi mintavétel

a populáció bármelyik tagja esélyes, de eltérő, előre meghatározott súly szerint

Véletlen számok generálása

16 . oldal



Majdnem véletlen mintavételek

Szisztematikus mintavétel

18. oldal

Valamilyen kezdőponttól indulva kiválasztjuk minden K -adik elemet a populációból

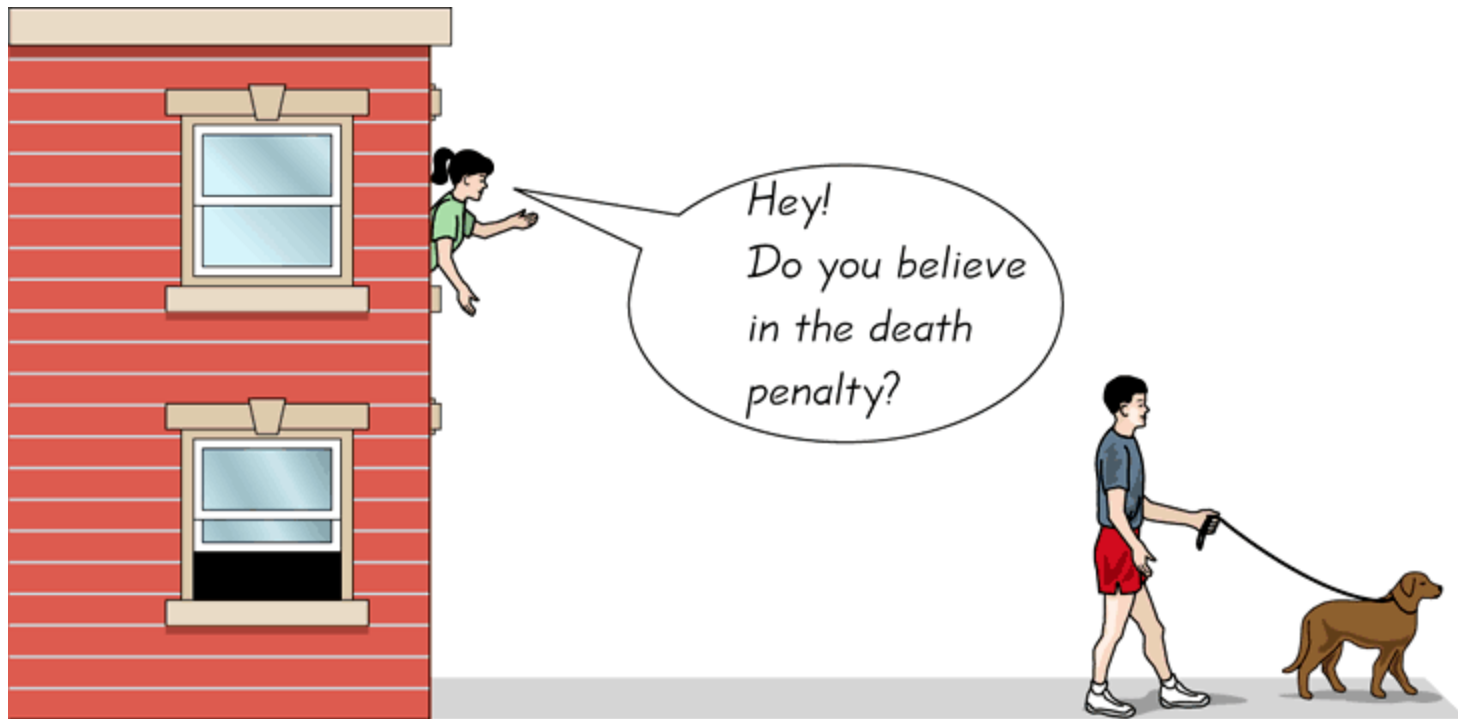


problémás lehet, ha a populáció is szisztematikusan van rendezve

Kényelmes mintavétel

19. oldal

használjuk azt a mintát, amit a legkönnyebb beszerezni



Rétegzett mintavétel

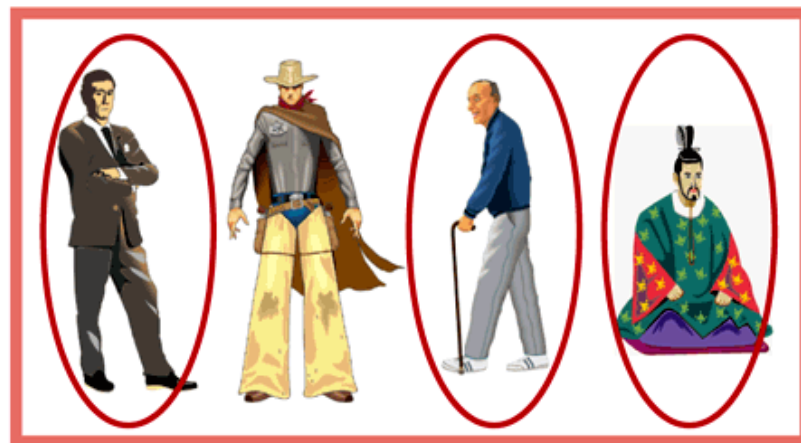
20 . oldal

oszd fel a populációt kettő vagy több csoportra (rétegre), melyeken belül bizonyos (a kísérlet szempontjából fontos) tulajdonságok azonosak vagy hasonlóak, majd vegyünk mintát mindegyik rétegből

Women



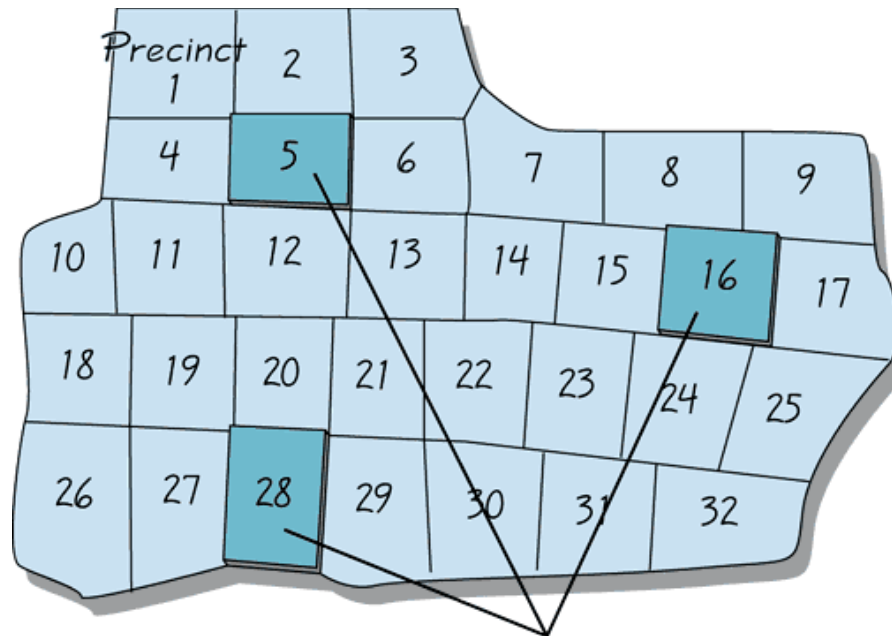
Men



Klaszter mintavétel

21 . oldal

oszd a populációt valamilyen természetes módon klaszterekre; véletlenül válassz közülük, használd **az összes** tagot



Interview all voters in shaded precincts.

A mintavételezés módszerei

22 . oldal

- ❖ Véletlen
- ❖ Szisztematikus
- ❖ Kényelmi
- ❖ Rétegzett
- ❖ Klaszter

- ❖ **Mintavételi hiba (Sampling error)**
a minta és a populáció eredménye közti eltérés, ami a minták fluktuációjából származik
- ❖ **Nem mintavételi hiba (Non-sampling error)**
olyan eltérés, ami az inkorrekt adatgyűjtésből, adat felvitelből vagy analízisből ered

Hibacsökkentés: a minta mérete

24 . oldal

❖ **Minta mérete**

akkora mintát kell használni, ami elég nagy ahhoz, hogy kimutathassuk benne az effektust

❖ **Méret növelés: replikázás**

a kísérlet megismétlése, amikor van elegendő alany ahhoz, hogy észrevehessük a különböző kezelések közti eltéréseket

Ebben a fejezetben:

- ❖ **A vizsgálatok és mérések típusait**
- ❖ **A változók hatásának kontrollálását**
- ❖ **Randomizációt**
- ❖ **A mintavételezés típusait**
- ❖ **A minta hibáit**

tekintettük át.

Az adatok leírása, megismerése és összehasonlítása

- 2-1 Áttekintés
- 2-2 Gyakoriság eloszlások
- 2-3 Adatok vizualizációja
- 2-4 A középpont mérőszámai

Cotinine

27 . oldal

From Wikipedia, the free encyclopedia

Cotinine is a [metabolite](#) of [nicotine](#). Cotinine typically remains in the [blood](#) between 48 and 96 hours. The level of cotinine in the blood is proportionate to the amount of exposure to tobacco smoke, so it is a valuable indicator of tobacco smoke exposure, including secondary smoke. Women who smoke [menthol](#) cigarettes retain cotinine in the blood for a longer period.^[1] Race may also play a role, as blacks routinely register higher blood cotinine levels than whites. Several variable factors, such as menthol cigarette preference and puff size, suggest that the explanation for this difference may be more complex than gender or race.

Drug tests can detect cotinine in the [blood](#), [urine](#), or [saliva](#).

The word 'cotinine' is an [anagram](#) of 'nicotine'.

Chemical Name: (S)-1-methyl-5-(3-pyridinyl)-2-Pyrrolidinone

Synonymes: Cotinine; (-)-Cotinine; 1-Methyl-5-(3-pyridinyl)-2-pyrrolidinone;

Chemical Formula: C₁₀H₁₂N₂O

Molar mass: 176.22 g/mol

There is some research being done on the effects of cotinine on memory and cognition. Some studies have suggested that cotinine (as well as nicotine) improves memory and prevents [neuron](#) death. For this reason it has been studied for effectiveness in treating [schizophrenia](#), [Alzheimer's](#) and [Parkinson's](#) diseases ^[2]. There is research, however, which also suggests that nicotine and cotinine contribute to Alzheimer's disease in other ways which counter and maybe even negate the possible positive effects they might have ^[3]

Szérum cotinine szint (ng/ml) 120 vizsgált alanyánál, egész számra kerekítve

Table 2-1 Measured Cotinine Levels in Three Groups

Smoker: The subjects report tobacco use.

ETS: (Environmental Tobacco Smoke) Subjects are nonsmokers who are exposed to environmental tobacco smoke ("secondhand smoke") at home or work.

NOETS: (No Environmental Tobacco Smoke) Subjects are nonsmokers who are not exposed to environmental tobacco smoke at home or work. That is, the subjects do not smoke and are not exposed to secondhand smoke.

Smoker:	1	0	131	173	265	210	44	277	32	3
	35	112	477	289	227	103	222	149	313	491
	130	234	164	198	17	253	87	121	266	290
	123	167	250	245	48	86	284	1	208	173
ETS:	384	0	69	19	1	0	178	2	13	1
	4	0	543	17	1	0	51	0	197	3
	0	3	1	45	13	3	1	1	1	0
	0	551	2	1	1	1	0	74	1	241
NOETS:	0	0	0	0	0	0	0	0	0	0
	0	9	0	0	0	0	0	0	244	0
	1	0	0	0	90	1	0	309	0	0
	0	0	0	0	0	0	0	0	0	0

2-1 fejezet

Áttekintés

❖ **Leíró statisztika**

a legfontosabb tulajdonságok **összegzése**
vagy **leírása** a populáció egy ismert
részhalmazán

❖ **Következtető statisztika**

a minta felhasználásával
következtetések (vagy általánosítások) a
populációról

Az adatok legfontosabb leíró tulajdonságai

31 . oldal

1. **Eloszlás (distribution)**: Az adatok eloszlásának természete, alakja (pl.: harang görbe, egyenletes, ferde)
2. **Középpont (center)**: Egy reprezentáns vagy átlag érték, ami megmutatja, hogy hol van a közepe az adathalmaznak
3. **Szóródás**: Annak a mértéke, hogy mennyire szóródnak az adatok
4. **Kiugró adatok (outliers)**: Olyan mintaértékek, melyek a mintaértékek döntő többségétől messze helyezkednek el
5. **Idő**: Az adatok időben változó tulajdonságai

2-2 fejezet

Gyakoriság eloszlás

❖ **Gyakoriság eloszlás (frequency distribution)**

az adat értékek listája (vagy egyenként vagy mint intervallum csoportok), együtt a gyakoriságukkal vagy számukkal

A dohányosok cotinine szintjeinek eloszlása

34 . oldal

Table 2-2

Frequency Distribution
of Cotinine Levels
of Smokers

<u>Cotinine</u>	<u>Frequency</u>
0–99	11
100–199	12
200–299	14
300–399	1
400–499	2

Definíciók

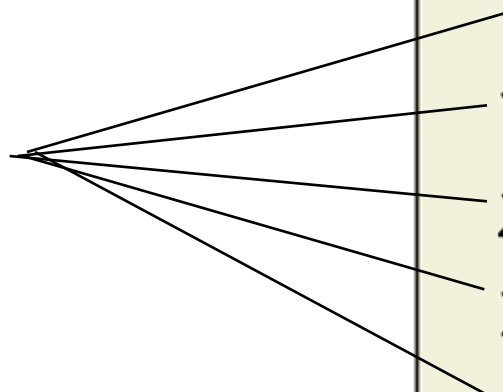
Az alsó osztálykorlát

36 . oldal

a legkisebb szám, ami még az adott osztályba kerülhet

**Alsó
osztálykorlát**

<u>Cotinine</u>	<u>Frequency</u>
0–99	11
100–199	12
200–299	14
300–399	1
400–499	2



Felső osztálykorlát

37 . oldal

a legnagyobb szám, ami egy osztályba bekerülhet

**Felső
osztálykorlátok**

<u>Cotinine</u>	<u>Frequency</u>
0-99	11
100-199	12
200-299	14
300-399	1
400-499	2

Osztályhatárok

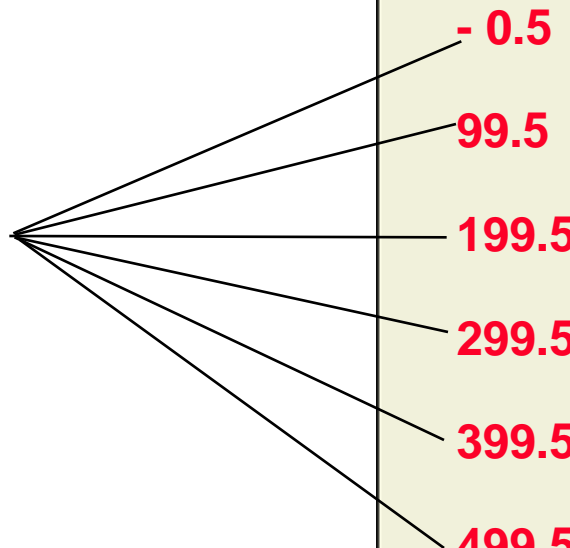
38 . oldal

azok a számok, amelyek elválasztják az osztályokat egymástól, de az osztálykorlátok által kreált hézagok nélkül

Osztályhatár

az osztályokat elválasztó számok

**Osztály-
határok**



	<u>Cotinine</u>	<u>Frequency</u>
- 0.5	0–99	11
99.5	100–199	12
199.5	200–299	14
299.5	300–399	1
399.5	400–499	2
499.5		

Osztály felezőpont

40 . oldal

az osztályok felezőpontjai

az osztály felezőpontot a felső és alsó osztálykorlátok összege elosztva kettővel adja meg

Osztály felezőpontok

41 . oldal

az osztályok felezőpontjai

Osztály
felezőpontok

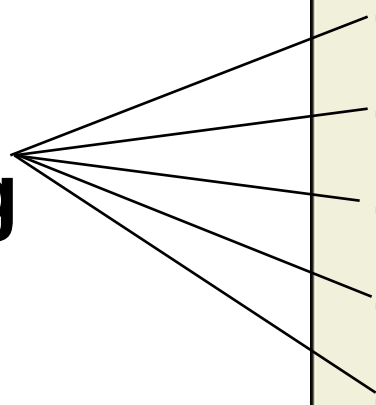
Cotinine	Frequency
0 - 49.5	11
100 - 149.5	12
200 - 249.5	14
300 - 349.5	1
400 - 449.5	2

Osztályszélesség

két egymás utáni alsó osztálykorlát vagy osztályhatár különbsége

**Osztály-
szélesség**

	<u>Cotinine</u>	<u>Frequency</u>
100	0–99	11
100	100–199	12
100	200–299	14
100	300–399	1
100	400–499	2



A gyakoriság eloszlások készítésének okai

- 1. Összegezzük nagy adathalmazokat.**
- 2. Betekintést nyerünk az adatok tulajdonságaiba.**
- 3. Grafikonok készítésének alapjául szolgálnak.**

Gyakorisági tábla készítése

44 . oldal

1. Döntsd el hány osztályt szeretnél (tipikusan 5 és 20 között, de lehet sokkal több is) .

2. Számold ki az adatokra (kerekítve).

$$\text{osztályszélesség} \approx \frac{\text{(legnagyobb)} - \text{(legkisebb)}}{\text{osztályok száma}}$$

3. Kezdőpont: Először az első osztály alsó korlátját határozd meg.

4. Kezdve az első osztály alsó korlátjával osztályszélesség lépésekben listázd ki az összes osztály alsó korlátjait.

5. Listázd függőleges oszlopba az alsó korlátokat, és írd melléjük a felső korlátokat.

6. Menj végig az adatokon, és egyenként és sorold be őket a helyes osztályba, számold meg, hogy egy osztályba hány adat került.

A relatív gyakoriság eloszlás

45 . oldal

$$\text{relatív gyakoriság} = \frac{\text{az osztály gyakorisága}}{\text{az összes gyakoriság összege}}$$

Relatív gyakoriság eloszlás

46 . oldal

Cotinine	Frequency
0–99	11
100–199	12
200–299	14
300–399	1
400–499	2

Összes gyakoriság = 40

Table 2-3

Relative Frequency
Distribution of Cotinine
Levels in Smokers

Cotinine	Relative Frequency
0–99	28%
100–199	30%
200–299	35%
300–399	3%
400–499	5%

$11/40 = 28\%$

$12/40 = 40\%$

stb.

Kumulatív gyakoriság eloszlás

47 . oldal

Cotinine	Frequency
0–99	11
100–199	12
200–299	14
300–399	1
400–499	2

Table 2-4
Cumulative Frequency Distribution
of Cotinine Levels in Smokers

Cotinine	Cumulative Frequency
Less than 100	11
Less than 200	23
Less than 300	37
Less than 400	38
Less than 500	40

**Kumulált
gyakoriságok**

Gyakoriság táblák (ismétlés)

48 . oldal

Table 2-2

Frequency Distribution
of Cotinine Levels
of Smokers

<u>Cotinine</u>	<u>Frequency</u>
0–99	11
100–199	12
200–299	14
300–399	1
400–499	2

Table 2-3

Relative Frequency
Distribution of Cotinine
Levels in Smokers

<u>Cotinine</u>	<u>Relative Frequency</u>
0–99	28%
100–199	30%
200–299	35%
300–399	3%
400–499	5%

Table 2-4

Cumulative Frequency Distribution
of Cotinine Levels in Smokers

<u>Cotinine</u>	<u>Cumulative Frequency</u>
Less than 100	11
Less than 200	23
Less than 300	37
Less than 400	38
Less than 500	40

Ebben a fejezetben szó volt:

- ❖ **Az adatok néhány fontos tulajdonságáról**
- ❖ **A gyakoriság eloszlásról**
- ❖ **A gyakoriság eloszlás készítésének procedúrájáról**
- ❖ **A relatív gyakoriság eloszlásról**
- ❖ **A kumulatív gyakoriság eloszlásról**

2-3 fejezet

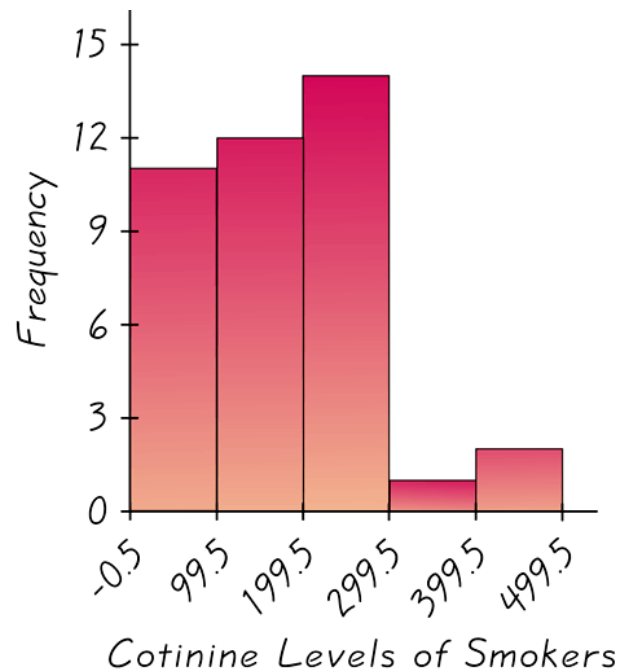
Az adatok vizualizációja

**Ábrázoljuk az adatok eloszlását
vagy az eloszlásának jellegét**

Hisztogram

Egy oszlopdiagram, melyben a vízszintes skála az adatok osztályait, a függőleges a gyakoriságokat reprezentálja.

Cotinine	Frequency
0–99	11
100–199	12
200–299	14
300–399	1
400–499	2

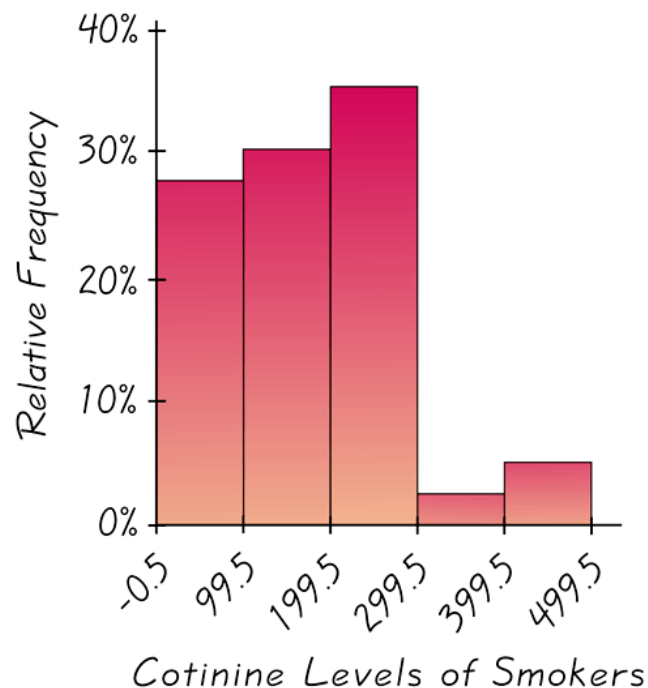


2-1 ábra

Relatív gyakoriság hisztogram

Oszlopdiaagram, amin a vízszintes tengelyen az osztályok, a függőlegesen a relatív gyakoriságok vannak ábrázolva.

Cotinine	Relative Frequency
0–99	28%
100–199	30%
200–299	35%
300–399	3%
400–499	5%

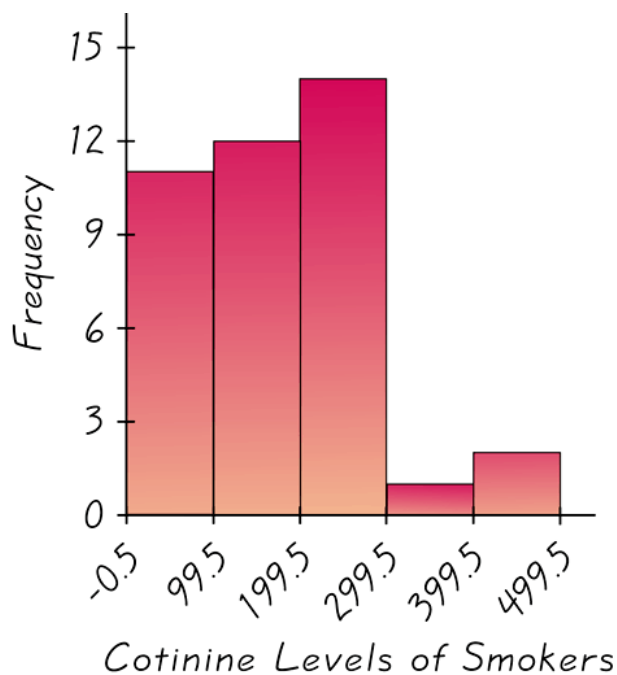


2-2 ábra

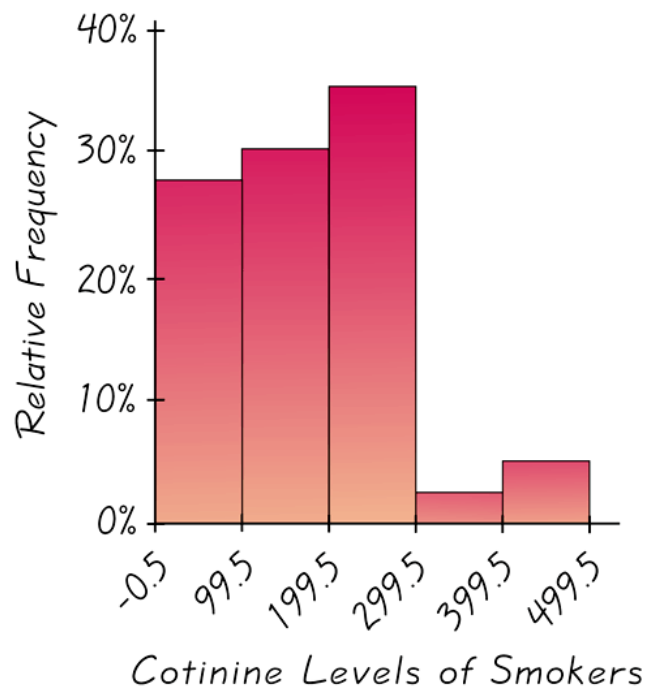
Hisztogram

és

Relatív gyakoriság hisztogram



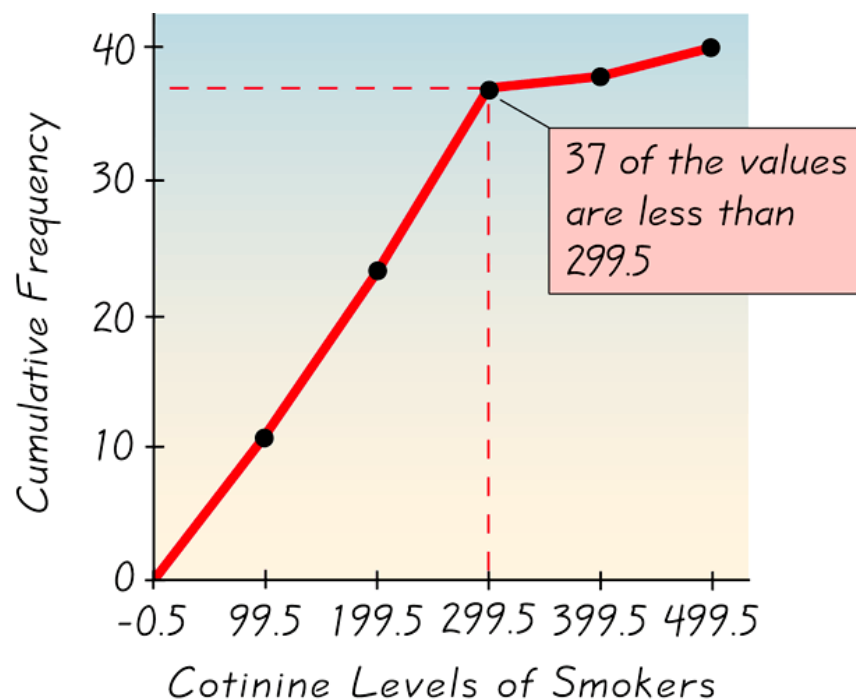
2-1 ábra



2-2 ábra

Ogiva (ogive)

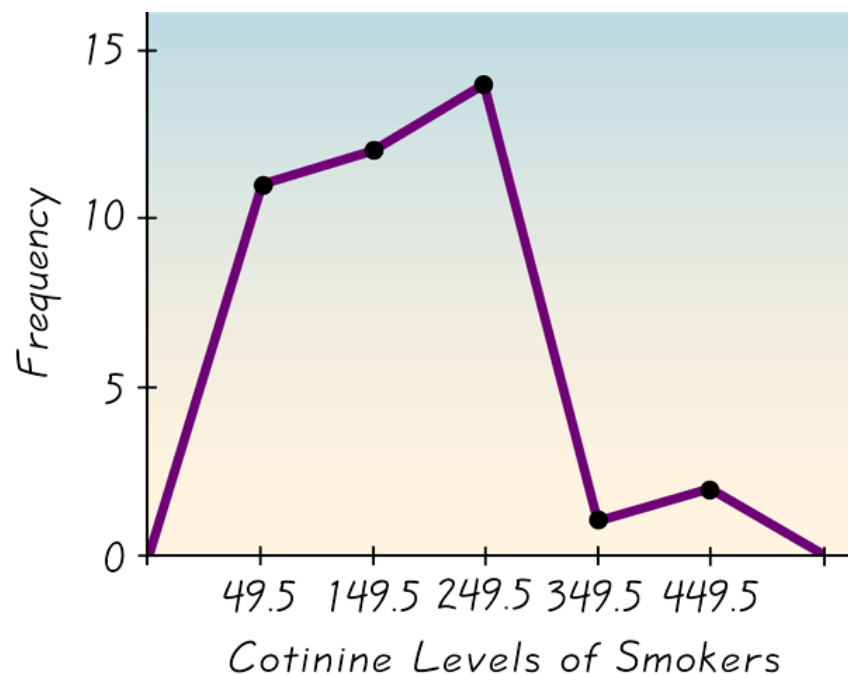
Vonal diagram ami a **kumulatív** gyakoriságokat mutatja



2-4 ábra

Gyakorisági poligon

Egyenes szakaszokkal köti össze az osztály felezőpontok feletti hisztogram értékeket.



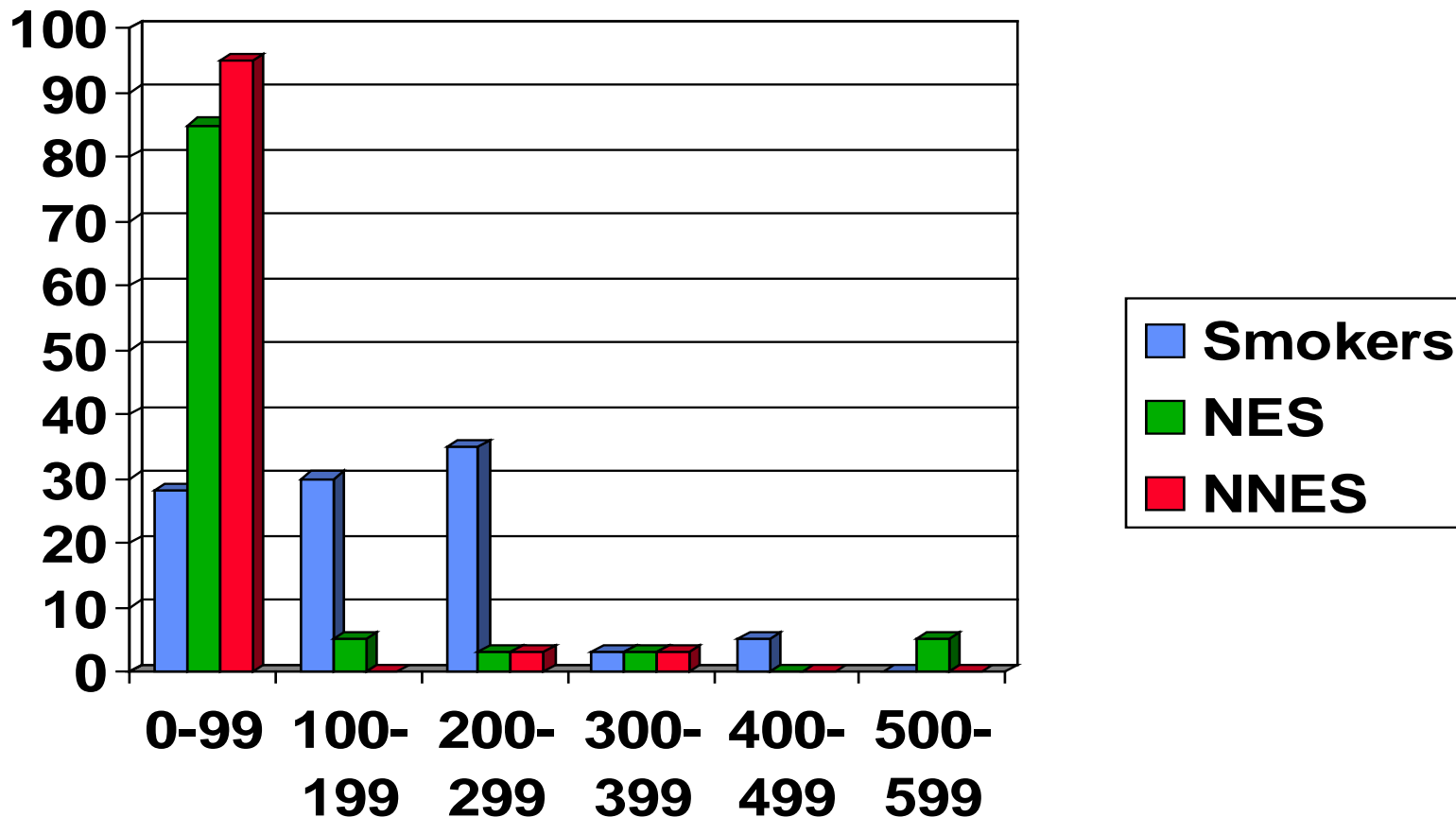
2-3 ábra

Kritikus gondolkodás

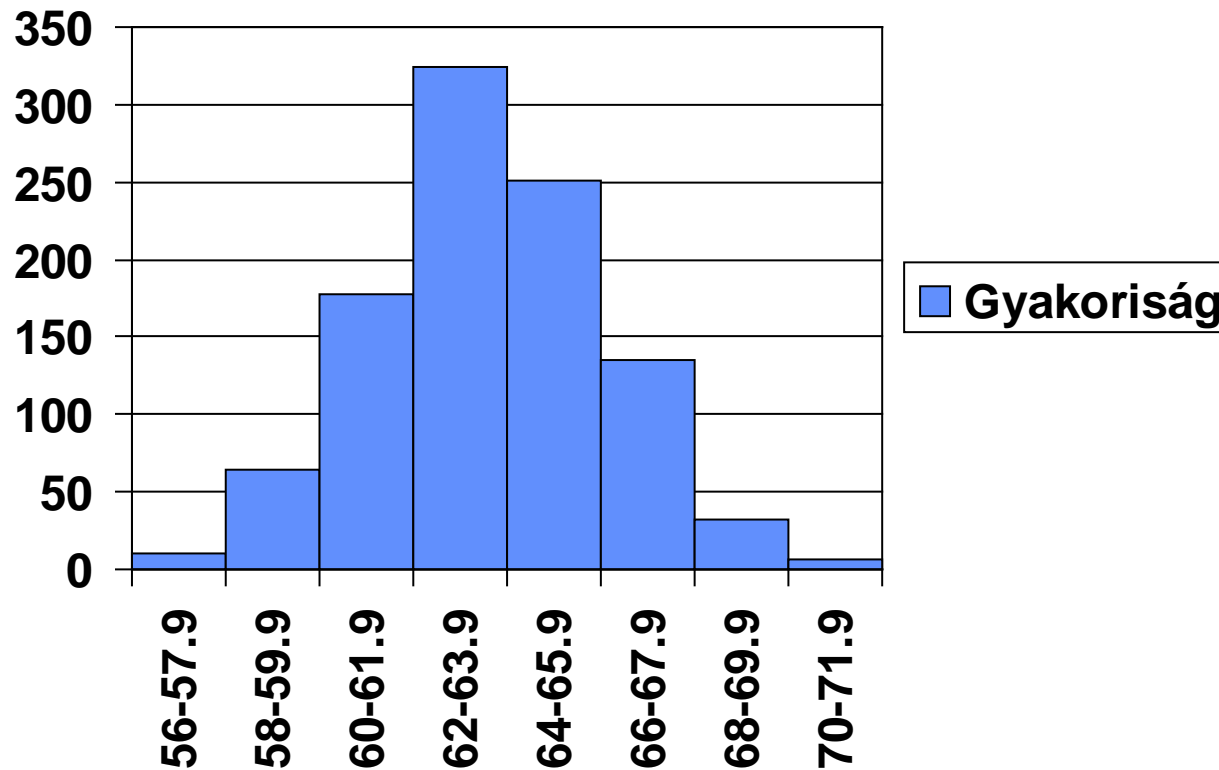
57 . oldal

- **Mit tudhatunk meg a hisztogramokból?**
- **Összehasonlíthatunk különböző adatcsoportokat.**
- **Észlelhetünk valami érdekeset az alakjában.**

Cotinine szint a három csoportban



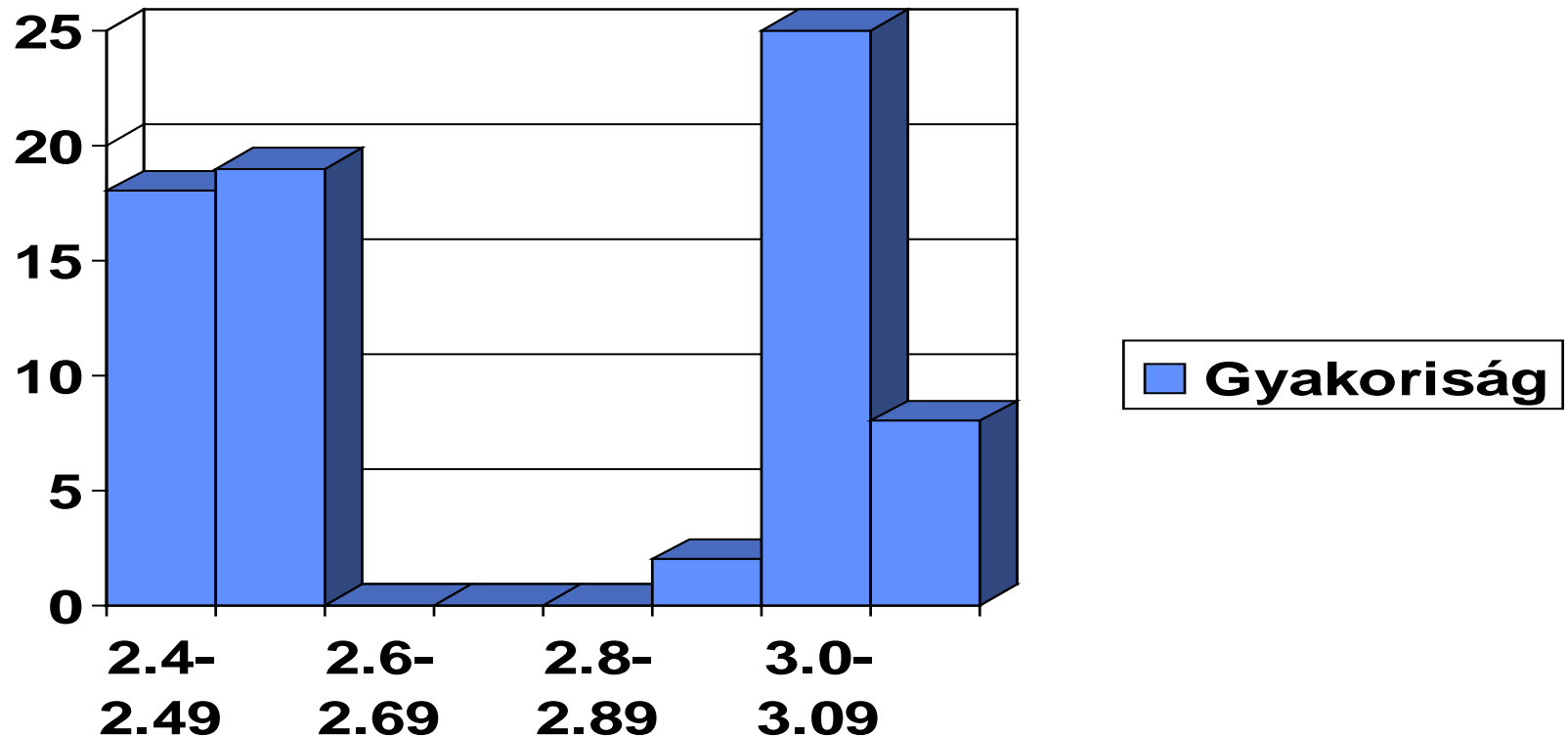
1000 nő magasságának gyakorisága



Haranggörbe – „normális” eloszlás

Véletlenül kiválasztott pennyk

60 . oldal



Mit jelent a hézag? 97% réz 3% cink (1983 előtt) 3% réz 97% cink (1983 után)

Telefon szolgáltatókkal kapcsolatos reklamációk

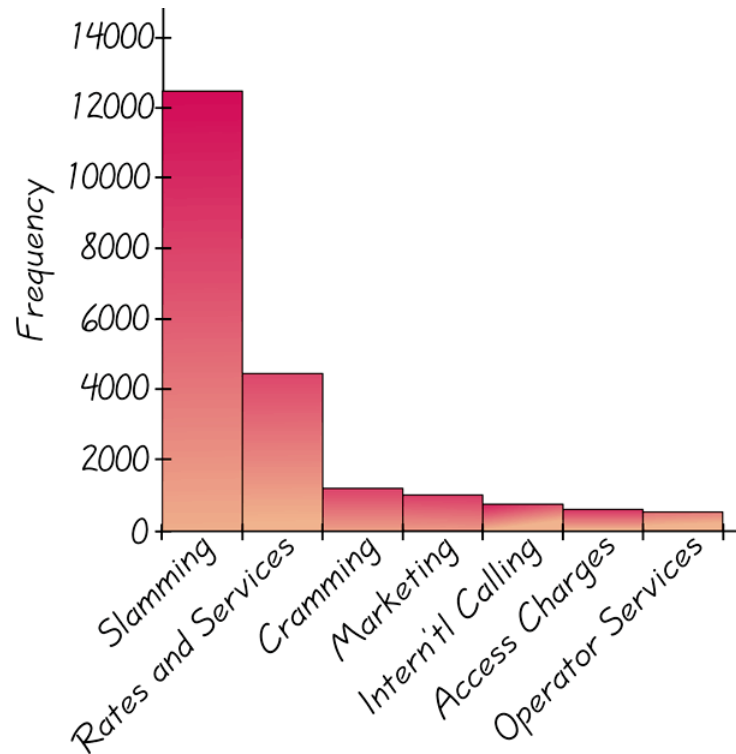
61 . oldal

Probléma	Száma
Rates and services (Árak és szolgáltatások)	4473
Marketing	1007
International Calling (Nemzetközi hívások)	766
Access Charges (Előfizetési díjak)	614
Operator services (Operátor szolgálat)	534
Cramming (nem kap vonalat)	1214
Slamming (átverés)	12478

Pareto grafikon (Pareto Chart)

62 . oldal

Kvalitatív adatokból képzett oszlopdiaagram, melynél az oszlopokat nagyság szerint rendezzük.

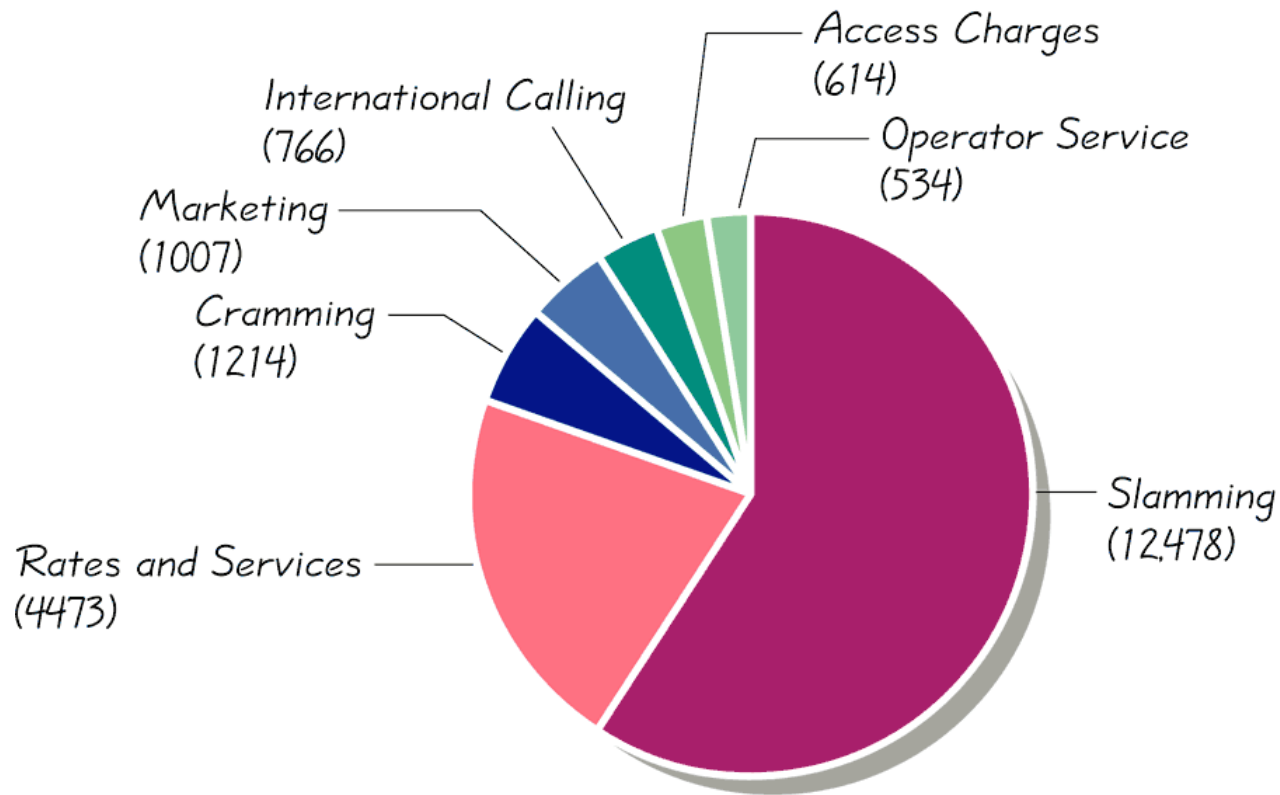


2-6 ábra

Kördiagram (Pie Chart)

63 . oldal

Kvalitatív adatokat mint egy torta szeleteit mutatja be.

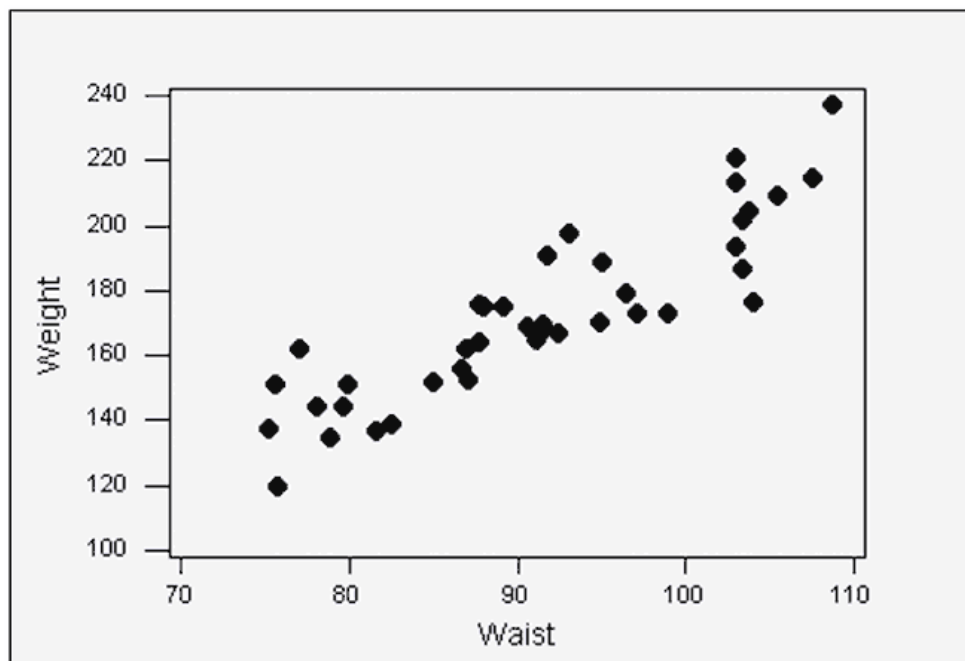


2-7 ábra

Szórásdiagram (Scatter Diagram)

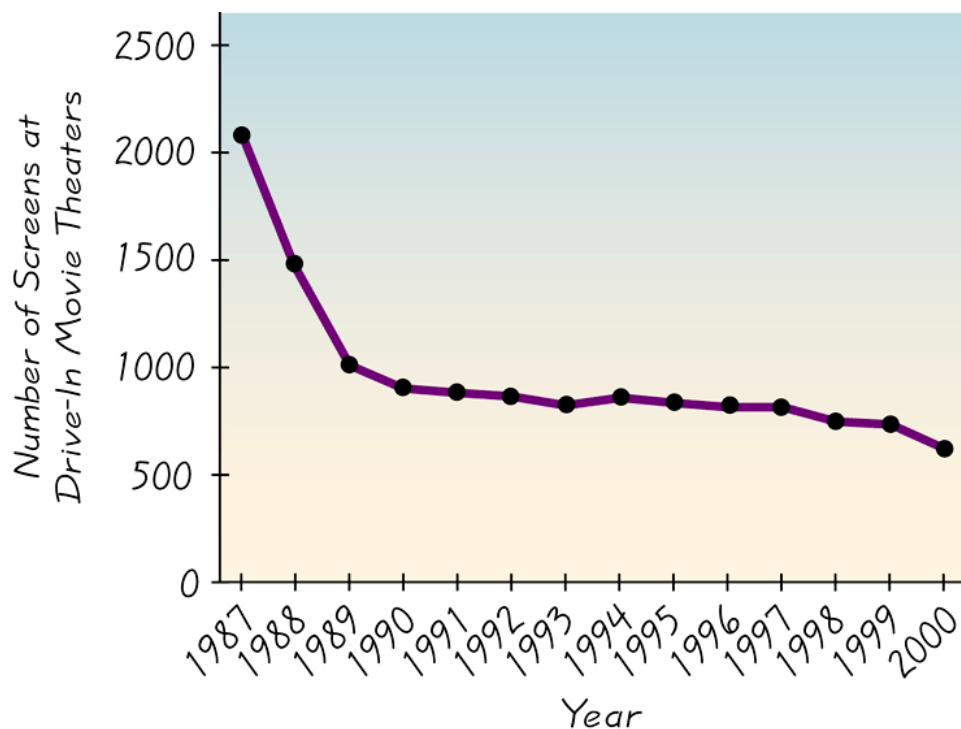
64 . oldal

Adatpárokba rendezett (x,y) adatok ábrázolása, ahol a vízszintes tengelyen az x, a függőlegesen az y értékek vannak feltüntetve



Férfiak derékbősége és súlya

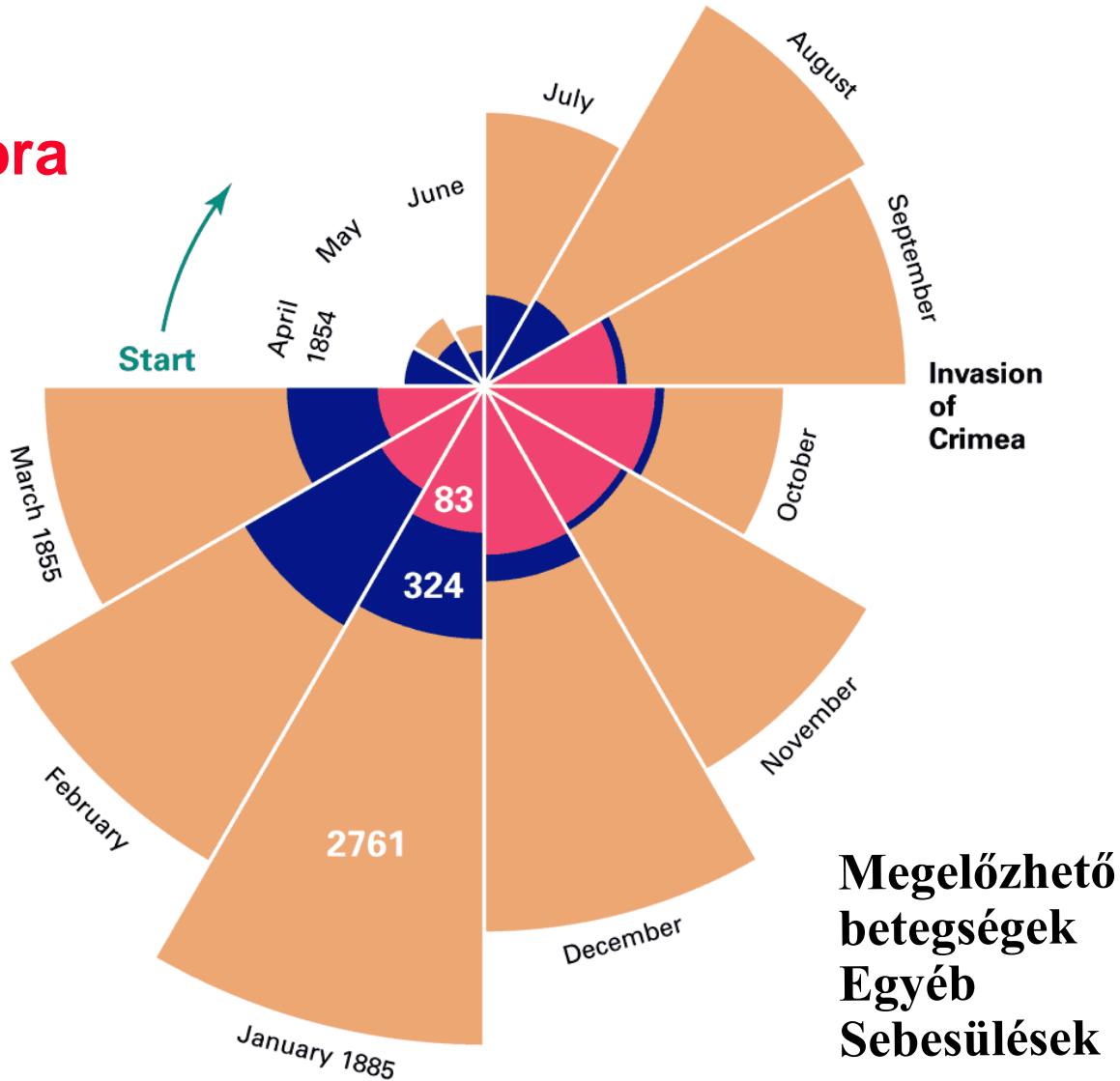
Különböző időpontokban gyűjtött adatok ábrázolására



2-8. ábra (Autós mozik száma)

Egyebek

2-9. ábra



A Krími háborúban Brit katonai kórházakban meghalt katonák

Ebben a fejezetben eloszlásokat bemutató grafikonokról volt szó.

A lényeg nem csak az, hogy hogyan rajzoljunk ilyeneket, hanem főleg az, hogy tanuljunk valamit belőle magukról az adatokról, megismerjük az eloszlás jellegét.

2-4 fejezet

A centrum mérőszámai

❖ **A centrum mérőszámai**

Az az érték, ami a közepén van az adathalmaznak

Definíció

70 . oldal

Számtani közép (Átlag)

a centrumot az adatok összeadása és annak az adatok számával való elosztása (átlaga) útján kapjuk

Jelölések

71 . oldal

Σ jelöli az értékek halmazának **összeadását**

x a **változó** amivel általában jelölni szoktuk az egyes adat értékeket

n jelöli az **értékek számát a mintában**

N jelöli az **értékek számát a populációban**

Jelölés

72 . oldal

\bar{x} 'x felülvonás' ('x-bar') jelöli a **mintabeli értékek átlagát**

$$\bar{x} = \frac{\sum x}{n}$$

μ 'mú' jelöli a **populáció** minden tagjára vett átlagot

$$\mu = \frac{\sum x}{N}$$

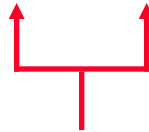
❖ Medián

a középső érték, ha az adatok növekvő vagy csökkenő sorrendbe vannak rendezve

❖ gyakran jelölik \tilde{x} ('x-hullám', 'x-tilde')

❖ nincs rá hatással egy-egy kiugró érték

5.40	1.10	0.42	0.73	0.48	1.10
0.42	0.48	0.73	1.10	1.10	5.40



(páros számú adat – nincs pontos közepe)

$$\frac{0.73 + 1.10}{2}$$

0.915 a MEDIÁN

5.40	1.10	0.42	0.73	0.48	1.10	0.66
0.42	0.48	0.66	0.73	1.10	1.10	5.40

(rendezve -

páratlan számú adat)

pontos közép **0.73 a MEDIÁN**

Definíciók

75 . oldal

❖ **Módusz, modális érték**

az az érték, ami a leggyakrabban fordul elő

A módusz nem mindig egyértékű. Az adathalmaz lehet:

Bimodális

Multimodális

Nincs modális érték

❖ **M –mel jelölik**

centrális tendencia egyetlen olyan mértéke,
amit **nominális** adatokra is lehet értelmezni

Példák

a. 5.40 1.10 0.42 0.73 0.48 1.10

← **1.10 a módusz**

b. 27 27 27 55 55 55 88 88 99

← **Bimodális - 27 & 55**

c. 1 2 3 6 7 8 9 10

← **Nincs módusz**

❖ Tartományközép (Midrange)

a legnagyobb és a legkisebb érték között félúton van

$$\text{Tartományközép} = \frac{\text{legnagyobb} + \text{legkisebb}}{2}$$

Súlyozott átlag

78 . oldal

Bizonyos esetekben az adatok nem egyformán fontosak, és adhatunk nekik eltérő súlyt

$$\bar{x} = \frac{\sum (w \cdot x)}{\sum w}$$

Pl.: év végi jegyet átlagolunk

A centrum legjobb jellemzése

Table 2-10 Comparison of Mean, Median, Mode, and Midrange

Measure of Center	Definition	How Common?	Existence	Takes Every Value into Account?	Affected by Extreme Values?	Advantages and Disadvantages
Mean	$\bar{x} = \frac{\sum x}{n}$	most familiar "average"	always exists	yes	yes	used throughout this book; works well with many statistical methods
Median	middle value	commonly used	always exists	no	no	often a good choice if there are some extreme values
Mode	most frequent data value	sometimes used	might not exist; may be more than one mode	no	no	appropriate for data at the nominal level
Midrange	$\frac{\text{high} + \text{low}}{2}$	rarely used	always exists	no	yes	very sensitive to extreme values

General comments:

- For a data collection that is approximately symmetric with one mode, the mean, median, mode, and midrange tend to be about the same.
- For a data collection that is obviously asymmetric, it would be good to report both the mean and median.
- The mean is relatively *reliable*. That is, when samples are drawn from the same population, the sample means tend to be more consistent than the other measures of center (consistent in the sense that the means of samples drawn from the same population don't vary as much as the other measures of center).

❖ Szimmetrikus

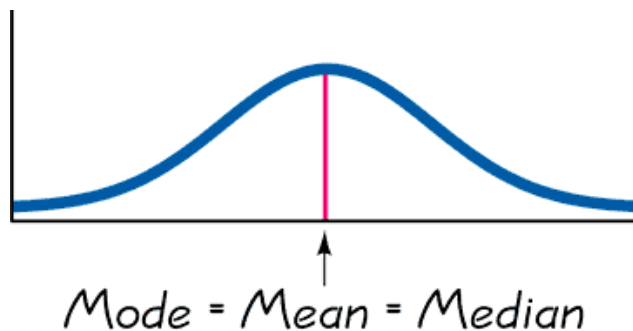
Az adatok szimmetrikusak, ha a hisztogram baloldala közelítőleg a tükörképe a jobboldalának.

❖ Ferde (aszimmetrikus)

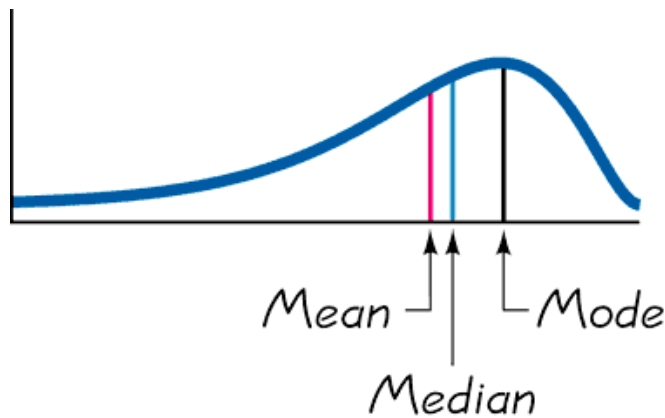
Az adatok ferdek vagy aszimmetrikusak, ha az egyik oldal felé jobban húznak mint a másik oldalra.

Ferdeség

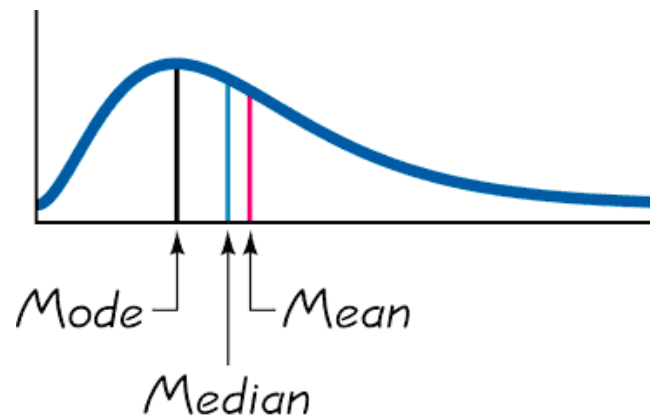
2-11. ábra



(b) Symmetric



(a) Skewed to the Left
(Negatively)



(c) Skewed to the Right
(Positively)

Összefoglalás

82 . oldal

Ebben a fejezetben szó volt:

- ❖ **A centrumot jellemző számok**

 - Átlag

 - Medián

 - Módusz

- ❖ **Az átlag kiszámítása a gyakoriság eloszlásból**

- ❖ **Súlyozott átlagok**

- ❖ **A centrum legjobb jellemzése**

- ❖ **Ferdeség**