



CORVINUS

UNIVERSITY of
B U D A P E S T

TÖBBVÁLTOZÓS ADATELEMZÉSI MODELLEK

CORVINUS

Elérhetőségek



Kovács László

laszlo.kovacs2@uni-corvinus.hu

E épület 111. (oktató)



Gombos Andrea

statisztika.tanszek@uni-corvinus.hu

titkárság, E.101.

Tananyag

- Moodle anyagok
 - R jegyzet HTML formátumban
 - Prezentációk
 - Órán használt adatbázisok
- Tankönyv
 - Kerékgyártó Györgyné - L.Balogh Irén - Sugár András - Szarvas Beatrix: Statisztikai módszerek és alkalmazásuk a gazdasági és társadalmi elemzésekben, AULA 2008
Elérhető:
https://mersz.hu/hivatkozas/dj270smea_book1#dj270smea_book1
 - Hanck, C., Arnold, M., Gerber, A., & Schmelzer, M. (2019). Introduction to Econometrics with R.
Elérhető: <https://www.econometrics-with-r.org/ITER.pdf>

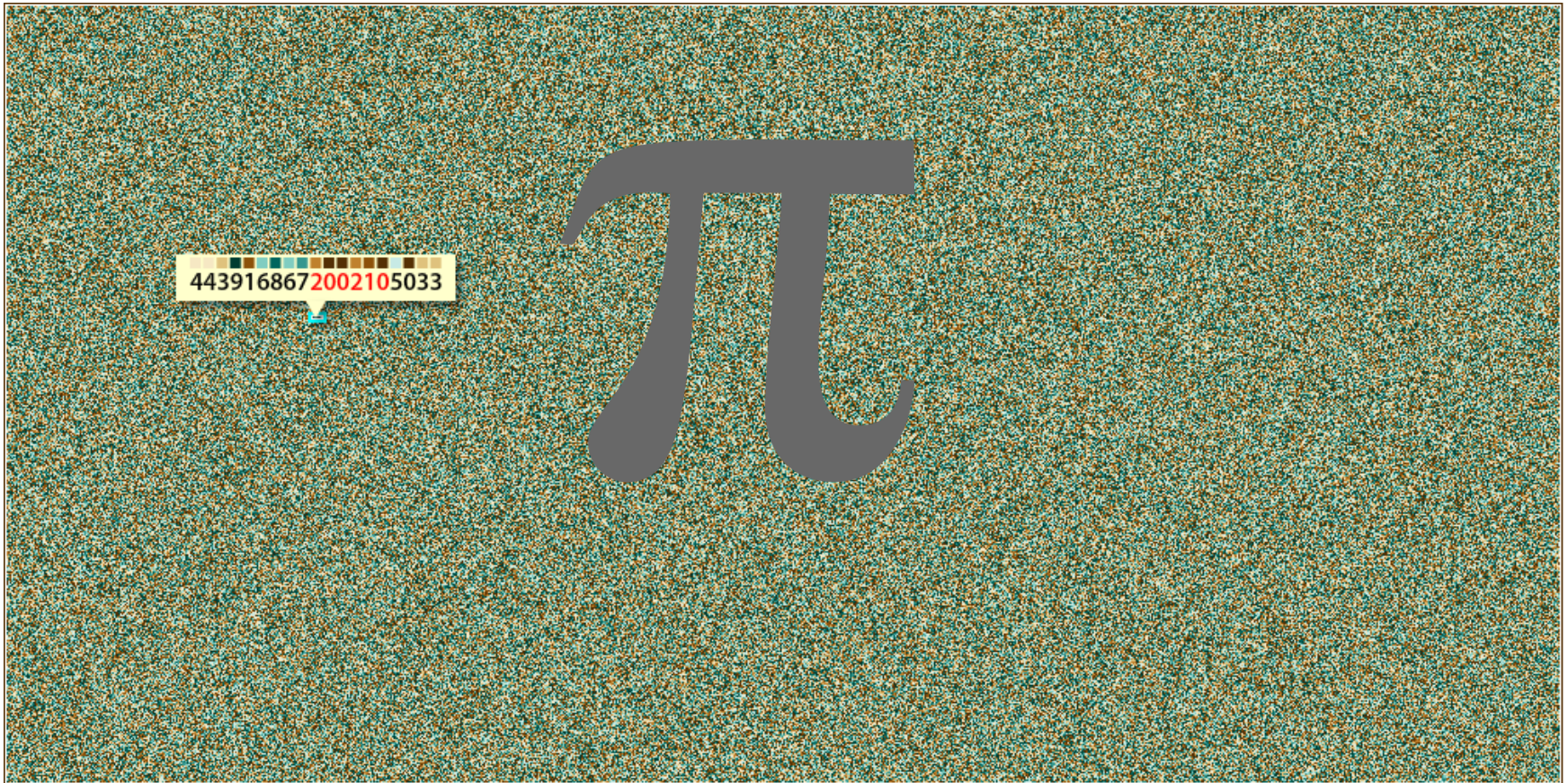
Számonkérés

- A félévi jegy összetevői:
 - Házi dolgozat: 60 pont
 - Prezentáció a vizsgaidőszakban: 40 pont
 - Összesen: 100 pont
- Ponthatárok:

	60	elégtelen
61	70	elégséges
71	80	közepes
81	90	jó
91		jeles

1. Alapfogalmak

Mi látható a képen?



two-n.com/pi

A STATISZTIKA TÁRGYA

- A *statisztika* a valóság tényeit **tömören, számszerűen** jellemezni, modellezni törekvő kettős tevékenység:
 - **gyakorlati tevékenység**
a tömegesen előforduló jelenségek egyedeire vonatkozó információk gyűjtése, feldolgozása, elemzése, közzététele
 - **tudományos módszertan**
a statisztikai tevékenység elmélete, az adatelemzés és következtetés tudományos módszertana

Negatív előítéletek a statisztikáról

- „Csak abban a statisztikában hiszek, amit én magam hamisítok.” (Churchill)
- „Háromfajta hazugság van: hazugság, szemérmetlen hazugság, és statisztika.” (Disraeli)
- „Háromféleképpen lehet hazudni: igennel, nemmel és statisztikával.” (Napóleon)
- „A statisztika olyan, mint a bikini: sok mindent megmutat, de a lényegét eltakarja.” (Levenstein)
- Okok:
 - statisztikát kevesen tanulnak, de az információkat használják
 - változó folyamatokról csak pillanatképet kapunk
 - bonyolult jelenségek \leftrightarrow leegyszerűsített modellek
 - az elemzésekkel eleve együtt járnak bizonyos (pl. mérési) hibák
 - etikai tényezők (torzított értelmezés)

Miért tanuljuk akkor mégis?

- „Míg az egyes ember megfejthetetlen rejtély, sokaságban matematikai bizonyossággá válik. Sosem tudjuk például előre megmondani, hogy egy ember mire lesz képes, de egész pontosan meg lehet mondani, átlagosan hány ember lesz képes rá. Az egyének különböznek, de a százalékok állandók. Így mondja a statisztikus.”
(Sir Arthur Conan Doyle: Sherlock Holmes - A négyek jele)
- „Statisztikával hazudni könnyű, de statisztika nélkül még könnyebb.” (Charles Frederick Mosteller)
- „Amit nem lehet számokban kifejezni, az nem tudomány, hanem vélemény.” (Robert A. Heinlein - The Notebooks of Lazarus Long)

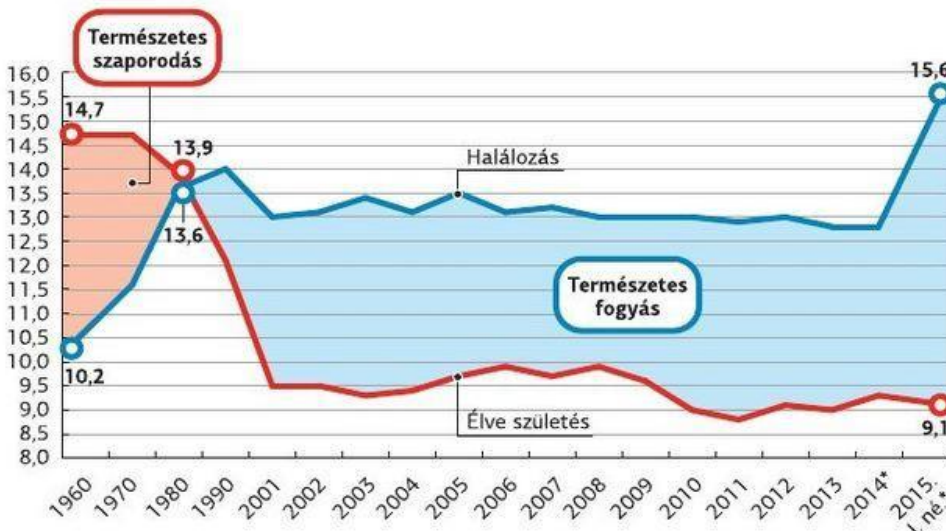
Gondoljuk végig az alábbi újságcikk tartalmát!

Döbbenetesen megugrott a magyarországi halálesetek száma

A szakemberek egyelőre csak találgatják az okait az első negyedéves brutális halálozási adatoknak...

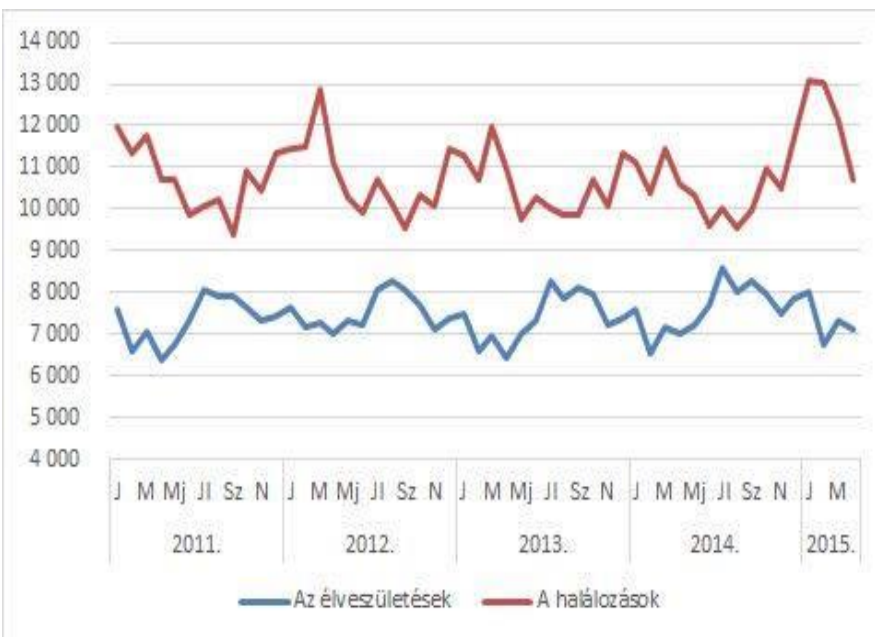
Az élve születések és a halálozások száma

EZRELÉK FORRÁS: KSH / NÉPSZABADSÁG-GRAFIKA



* Előzetes, részben becült adat

- Az élveszületések és a halálozások száma havi bontásban 2011-2015 között:



ALAPFOGALMAK

- Statisztika
 - **tömeges adatokból** jellemzés, elemzés, következtetés
- Alapfogalmak
 - mit vizsgálunk? → **sokaság (populáció)**
 - mi alapján vizsgáljuk? → **ismérv/változó**
 - hogyan mérjük? → **mérési skála**

SOKASÁG, ISMÉRV/VÁLTOZÓ

- **Sokaság:** a vizsgálat tárgyát képező statisztikai egységek összessége
🧠 pl. foglalkoztatottak, részvények, hallgatók
- **Ismérv/Változó:** a sokaság egységeinek jellemzője
🧠 pl. foglalkoztatottak életkora, részvény árfolyama, hallgatók neme, életkora
- **Változó értékkészlete:** a változó lehetséges kimenetelei
🧠 pl. hallgatók neme: nő, férfi

Változófajták

- Közös \leftrightarrow megkülönböztető változók
 - közös változók : a sokaság egyedei ezen jellemzők szerint egyformák \rightarrow ezek definiálják a sokaságot
 - megkülönböztető változók: eltérések \rightarrow ezeket vizsgáljuk a sokaság statisztikai jellemzésekor
- A változók fajtái:
 - **minőségi (kvalitatív)**: verbális jellemzés (pl. digitális fényképezőgép márkája)
 - **területi**: térbeli elhelyezés (pl. cég székhelye)
 - **numerikus (kvantitatív)**: számszerű jellemzés (pl. lakás alapterülete)
 - **időbeli**: időbeli elhelyezés (pl. születési év)

MÉRÉSI SKÁLÁK

- A nem numerikus változók lehetséges változatai számmá kódolhatók → minden megfigyelés mérésnek tekinthető
- Az így kialakult számértékekkel csak azok a műveletek végezhetők, amelyek az eredeti változó-értékeken is elvégezhetők
- *Mérési skálák*
 - névleges (nominális) skála
 - sorrendi (ordinális) skála
 - különbségi (intervallum) skála
 - arányskála

A mérési skálák fajtái 1.

- *Névleges (nominális)* mérési skála
 - a szimbólumok, számértékek csak egy megkülönböztető címke szerepét játsszák → azonosítást szolgálnak
 - csak az a tulajdonság értékelhető, hogy megegyeznek-e
 - 🤔 pl: kódszám (férfi:1, nő:0)
- *Sorrendi (ordinális)* mérési skála
 - a skálaértékek sorrendje hordoz információt, a skála egyes pontjai között hierarchia van, de nem tudjuk a távolságot közöttük; egy nagyságrendi viszony beszámozása
 - 🤔 pl.: helyezések

A mérési skálák fajtái 2.

- *Különbségi (intervallum) mérési skála*
 - a skála kezdőpontja önkényesen adott, vagy konvención alapszik → a skálának a mértékegység is szerves tartozékát képezi
 - a skálaértékek különbségei is információt hordoznak
 - 🤔 pl. hőmérséklet
- *Arány mérési skála*
 - bármely két skálaérték egymáshoz viszonyított aránya is egyértelműen meghatározható, és információt hordoz
 - 🤔 pl. termelési érték

TESZT: Alapfogalmak



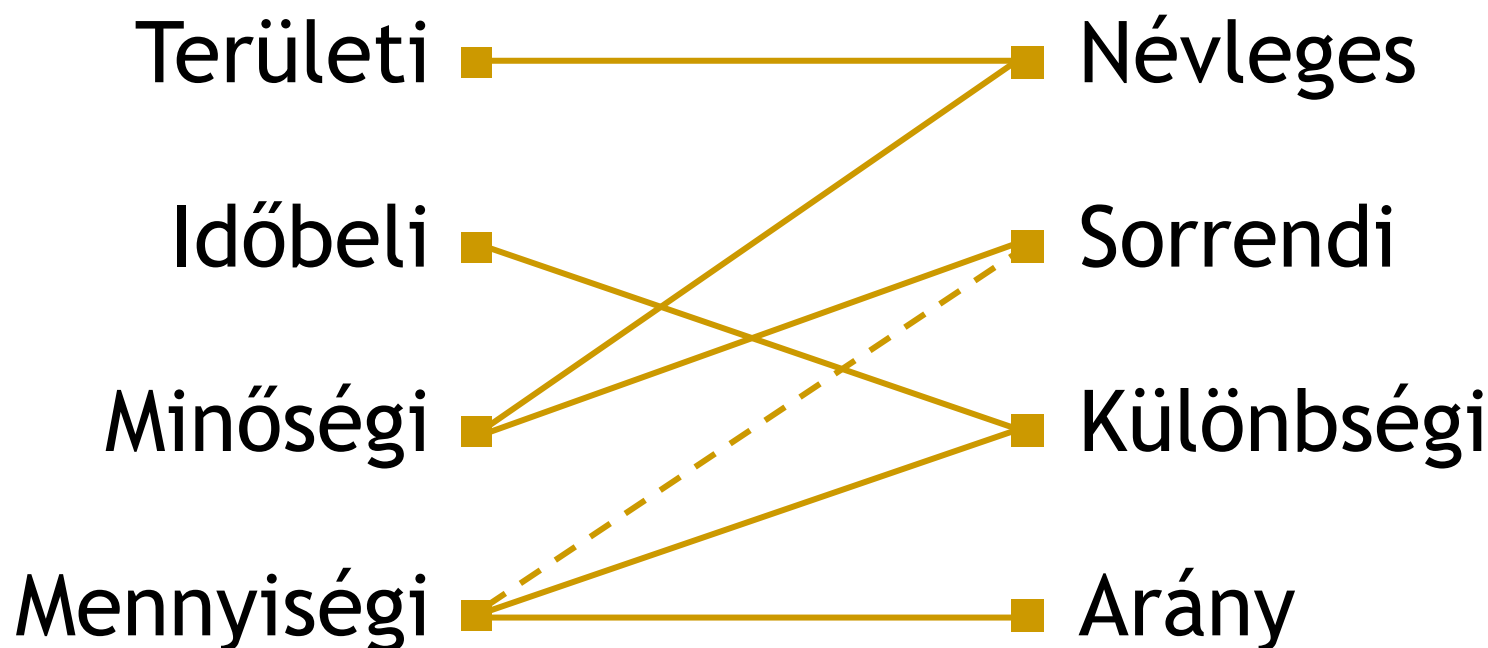
Mely mérési skálán mérhető a valamely tárgyból elért érdemjegy?

- a. nominális
- ☒ b. sorrendi
- c. különbségi
- d. arány

Ismérvfajták és mérési skálák

VÁLTOZÓFAJTA

MÉRÉSI SKÁLA



Egy kórház sürgősségi osztályán a tegnapi napon megjelent betegek adatait tartalmazza a következő táblázat. Mutassuk be ezen a tanult fogalmakat!

Név	Láz	Sürgősségi kat.	Szül. idő	Életkor	Lakhely ir.szám
Bete Gergő	36,6°C	2	1974.10.12.	43	1031
Influ Elza	38,5°C	5	1951.12.23.	66	1135
...	

↑
egyed

↑
különbségi
mérési
skála

↑
sorrendi
mérési
skála

↑
különbségi
mérési
skála

↑
arány
mérési
skála

↑
névleges
mérési
skála

🤔 Sokaság: a tegnapi napon megjelent betegek