

Generatív AI rendszerek használata bevezető statisztika kurzusokon

Peter Kovacs, Klara Kazar, Eva Kuruczleki, Regina Bodo,
Vivien Kardos, Tunde Szanto, Tamas Racz, Bence Kovacs

University of Szeged





GAISE ajánlások

1. Teach **statistical thinking**.
 - Teach statistics as an investigative process of problem-solving and decision making.
 - Give students experience with multivariable thinking.
2. Focus on **conceptual understanding**.
3. Integrate **real data** with a context and purpose.
4. Foster **active learning**.
5. Use **technology** to explore concepts and analyze data.
6. Use **assessments** to improve and evaluate student learning.

Source: Guidelines for Assessment and Instruction in Statistics

My Principles in Stat Ed

- Use IT: software, databases
- Don't teach platform, but teach critical thinking, statistical way of thinking
- Combine PP and IT.
- Use active teaching methods also
- Develop Stat Lit
- Attitudes development
- Video support





Why GenAI?

- Különböző példák az oktatásban: tanítás, tanulás, értékelés, oktatás támogatása
- Tipikus statisztikai alkalmazások az irodalomban:
 - R vagy Python kód létrehozása
 - Gépi tanulási technikák alkalmazása
- Használt rendszerek: ChatGPT, Gemini, Claude
- Hogyan használhatjuk ezeket más módon a Statisztika oktatásban?





OKTATÁS



Kurzusleírások készítése

- Promptok fontossága
- Kurzus célja, óraszáma
- ILO (Intended Learning Outcomes): tudás, készség, attitűd, felelősség, autonómia
- <https://claude.ai/chat/bbeb59b9-4946-4280-93a5-8a63085200fe>





▼ Egy bevezető Excel kurzus az alábbi tematikát követi: 1. Mire jó az Excel? 2. Az Excel menürendszere, felépítése (cellák, sorok, oszlopok, munkalapok, munkafüzet) 3. Adatbevitel, adatok importálása, cellaformátumok, rendezés, szűrők 4. Képletek 5. Excel munkakörnyezet (nézet, oldalbeállítás, élőfej, élőláb) 6. Alapvető függvények (szum, min, max, átlag, darab, darab2) használata 7. Abszolút, relatív, vegyes cellahivatkozás 8. Egyszerűbb diagramok készítése (vonal, oszlop, kör, pontdiagram) Sorolj fel a tananyaggal kialakítandó konkrét tanulási eredmények A tananyag elsajátítását követően a hallgató milyen tudás, képesség, attitűd, autonómia és felelősség elemekre tesz szert? Sorolj fel 10-10 darabot!

Köszönöm a kérdést! Szívesen segítek összeállítani egy listát a tanulási eredményekről az Excel bevezető kurzus tematikája alapján. Íme 10-10 konkrét elem a tudás, képesség, attitűd, valamint autonómia és felelősség kategóriákban:

- Tudás:
1. Ismeri az Excel alapvető funkcióit és felhasználási területeit.
 2. Érti az Excel felhasználói felületének felépítését és a menürendszer logikáját.
 3. Felismeri a különböző cellaformátumokat és azok alkalmazási területeit.
 4. Ismeri az alapvető Excel függvények (SZUM, MIN, MAX, ÁTLAG, DARAB, DARAB2) szintaxisát és működését.
 5. Megérti a képletek felépítését és logikáját.

8. Open to knowledge sharing with colleagues regarding Excel usage.

9. Strives to create user-friendly and easily understandable worksheets.
10. Values the importance of visual data representation in decision-making.

Autonomy and Responsibility (10 elements):

1. Independently designs and creates complex Excel workbooks.
2. Takes responsibility for the accuracy and reliability of their Excel analyses.
3. Initiates the introduction of advanced Excel functions into workflows.
4. Independently selects the most appropriate Excel techniques for solving given tasks.
5. Responsibly handles and protects Excel files containing confidential data.
6. Provides assistance to colleagues in using advanced Excel functions.
7. Independently develops Excel skills using online resources and literature.
8. Takes responsibility for the proper functioning of macros and automations they create.
9. Autonomously optimizes and develops existing Excel-based workflows.
10. Makes responsible decisions about using Excel or other software for efficient task solving.

These learning outcomes provide a comprehensive picture of the knowledge, skills, attitudes, and autonomy and responsibility that students can acquire after completing the advanced Excel course. The list can be further refined or expanded according to specific educational goals and the needs of the target audience.



Feladatok generálása

- Tesztek
- Vizsgafeladatok
- Feladatsorok
- Módoak: csak prompt/fájlok alapján/Mintafeladatsor alapján
 - Pld. Súlyozott átlag”





Tapasztalatok

- Válaszalternatívák és témakörök, módosító promptok
- Csak egy helyes válasz?
- Túl sablon alapú feltöltött mintafeladatsorok használatakor
- Kreatív szövegek
- Mértékegységek hiánya
- Alternatívák függetlensége vs. tipikus tanulói hibák
- Hiányos szöveg, kevesebb generált kérdés
- Néha helyesen, néha nem
- Megoldások során alkalmazott módszerek, tesztek



- <https://gemini.google.com/app/1526259d3e3a7986>
- <https://gemini.google.com/app/cb539b098d0ad77f>
- <https://gemini.google.com/app/c1d9ca04afb587a4>



Példák

Which factor affects the width of a confidence interval?

- A) Sample size
- B) Population size
- C) Number of observations above the mean
- D) Range of the sample

Given a sample mean of 50, a standard error of 5, and a z-score of 1.96 for a 95% confidence interval, what is the margin of error?

- A) 5
- B) 7.5
- C) 9.8
- D) 10





10. In a dataset where all values are positive integers, if the mean is 3.5 and the median is 3, what is a possible value for the mode?

- a) 2
- b) 3
- c) 4
- d) 5

1. Understanding the Given Information:

- The dataset consists of positive integers.
- The mean is 3.5.
- The median is 3.

2. Implications of the Median:

- Since the median is 3, we know that when the numbers are arranged in ascending order, the middle value(s) must result in 3.
- This implies that the dataset must contain the number 3 at least once and that the data is not perfectly symmetrical.

3. Implications of the Mean:

- The mean is 3.5, which is a non-integer. This tells us that the sum of the numbers in the dataset is not evenly divisible by the number of values.
- Because the mean is higher than the median, we know that the data is skewed to the right.

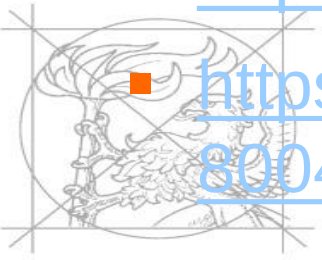
4. Finding a Possible Dataset:





Adatbázisok generálása

- Változók típusai
- Nem valós adatfájlok
- Adattábla összeállítása
- Adatbázis összeállítása
- <https://claude.ai/chat/99ec632c-5000-4afb-87f8-b86b53964ce6>
- <https://chatgpt.com/g/g-R6VqLNHFM-excel-ai/c/67db108c-ff28-8004-a79c-37a3c64b1588>





▼ gyűjtse össze a valaha létező 100 legjobbra értékelt filmet, foglalja táblázatba a sorszámat, a megjelenés évét, a műfaját és a hosszát is

Összeállítok egy táblázatot a valaha létezett 100 legjobbra értékelt filmről. Megjegyzem, hogy különböző források és értékelő platformok eltérő listákat készítenek, ezért ez a válogatás több népszerű kritikai forrás (IMDb, Rotten Tomatoes, Metacritic) és filmtörténeti jelentőség alapján készült.

 **A 100 Legjobbra Értékelt Film**
Click to open document

Elkészítettem a táblázatot a 100 legjobbra értékelt filmről, amely tartalmazza a sorszámat, a film címét (magyar és eredeti címmel), a megjelenés évét, a műfaját és a hosszát percben megadva.

A listában igyekeztem egyensúlyt teremteni a klasszikus filmtörténeti alkotások és a modern filmművészet kiemelkedő darabjai között. A válogatás több forrás értékelései alapján készült, figyelembe véve különböző kritikai szempontokat és a filmek kulturális jelentőségét.

A táblázatban szerepelnek olyan ikonikus alkotások, mint "A remény rabjai", "A keresztapa", "Schindler listája", valamint újabb remekművek, mint a "Parasite", "Interstellar" vagy "Soul – Lelki ismeretek".

A 100 Legjobbra Értékelt Film

Sorszám	Film címe	Megjelenés éve	Műfaj	Hossz (perc)
1	A remény rabjai (The Shawshank Redemption)	1994	Dráma	142
2	A keresztapa (The Godfather)	1972	Bűnügyi, Dráma	175
3	A keresztapa II (The Godfather: Part II)	1974	Bűnügyi, Dráma	202
4	A sötét lovag (The Dark Knight)	2008	Akción, Bűnügyi, Dráma	152
5	12 dühös ember (12 Angry Men)	1957	Dráma	96
6	Schindler listája (Schindler's List)	1993	Életrajzi, Dráma, Történelmi	195
7	A Gyűrűk Ura: A király visszatér (The Lord of the Rings: The Return of the King)	2003	Akción, Kaland, Dráma	201
8	Ponyvaregény (Pulp Fiction)	1994	Bűnügyi, Dráma	154

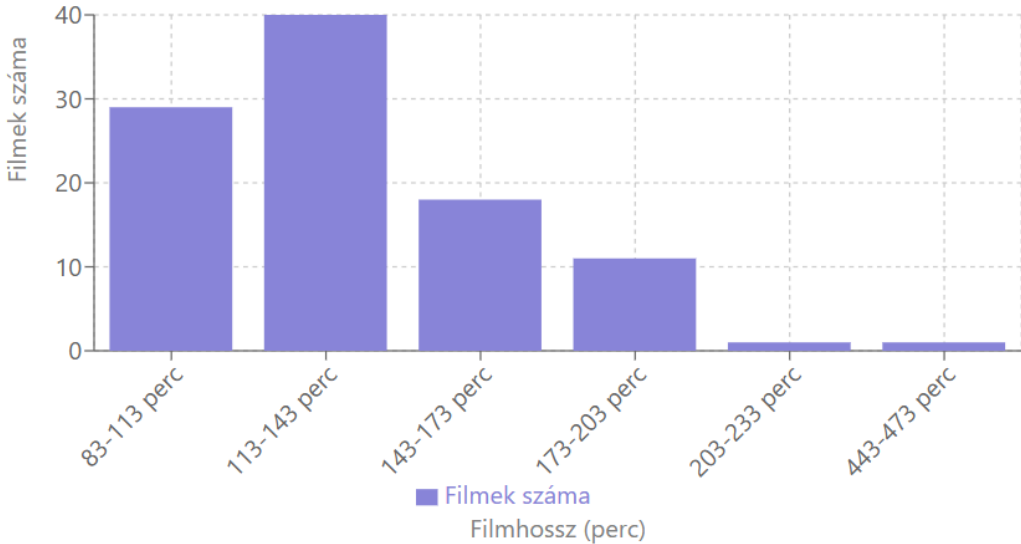
- 113-143 minutes: 40 films (the most common range)
- 143-173 minutes: 18 films
- 173-203 minutes: 11 films
- 203-233 minutes: 1 film
- 233-443 minutes: 0 films
- 443-473 minutes: 1 film (Sátántangó)

Summary

The statistical analysis reveals that most top-rated films are approximately 2-2.5 hours long. The fact that the mean (134.44 minutes) is greater than the median (127.5 minutes) confirms the positive skewness of the distribution. The large standard deviation and high skewness value demonstrate significant variation in film lengths, with a concentration in the 2-hour range but with some notable outliers.

The bar chart visually demonstrates this distribution, highlighting both the concentration of films in the 2-2.5 hour range and the impact of extreme values like Sátántangó's 450-minute runtime.

A 100 legjobb film hossz szerinti megoszlása (percben)





◆ **Mode (Módusz): 118 perc**

→ The most frequently occurring film length in the dataset.

◆ **Skewness (Ferdesség): 4.16**

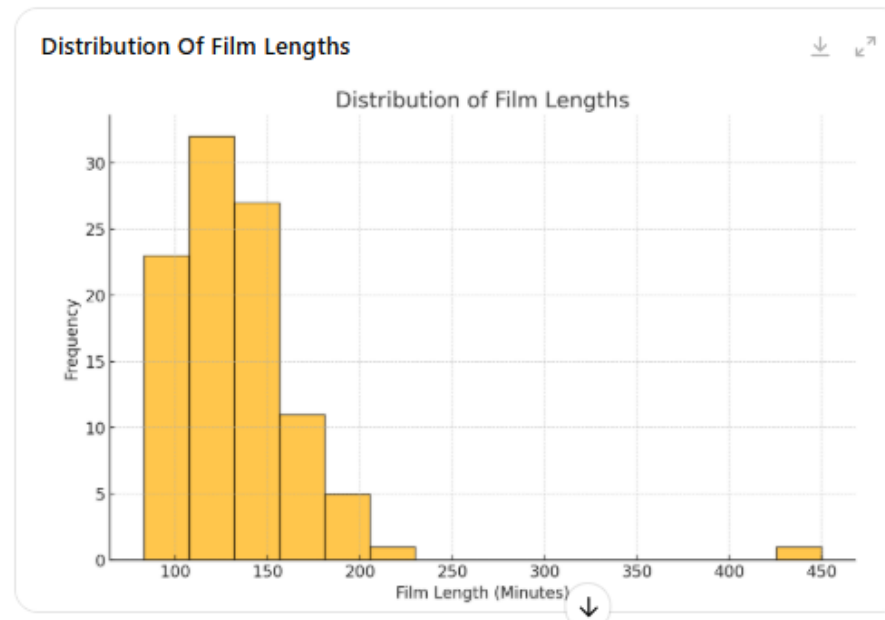
→ A measure of data asymmetry:

- **Positive skew (>0):** More films have shorter lengths, with a few very long ones.
- **Negative skew (<0):** More films have longer lengths, with a few very short ones.
- **Near zero:** The data is symmetric.

📌 **Interpretation:**

Since the skewness is **positive (4.16)**, it means there are **more shorter films**, with a few very long ones pulling the average up. This indicates a right-skewed distribution.

Now, I'll generate a **bar chart** 📊 showing the distribution of film lengths. 🎬 🔄



Kétségek

- Valós adattáblák vs. Generált adattáblák
- Kreativitás vs. Önálló gondolkodás





Hallgatói projekt

- Tapasztalatok
- Csoportmunka
- Prezentációk
- Alkalmazott rendszerek
- Válaszok validálása, ellenőrzése





Tanulás

- Hogyan tudják a hallgatók tanulásra használni?
- Kockázatok





Oktatás módszertani javaslatok

- <https://claude.ai/chat/b0b75689-88>

Teaching Time Series Analysis to 600 Economics

Undergraduates

Here's a detailed methodology for conducting a 90-minute lecture on time series analysis basics for 600 undergraduate economics students, with access to smartphones, internet, and a projector:

1. PREPARATION (before lecture):

- Make lecture materials/notes available online
- Generate QR codes for materials and online quizzes
- Prepare Kahoot or Mentimeter quizzes
- Prepare Excel example files for demonstrations

2. LECTURE STRUCTURE (90 minutes):

INTRODUCTION (10 minutes):

- Demonstrate practical significance with 2-3 engaging examples:
 - Stock price movements
 - COVID case trends
 - Inflation patterns
- Display QR code for accessing materials

THEORETICAL BLOCK 1 (15 minutes):



	At the begining of the carrier	8 years experience	15 years experience	25 years experience	40 years experience
Developing presentation and teaching skills	✓	✗	✗	✗	✗
Gamification	✓	✗	✓	✓	✓
Data visualization	✓	✗	✓	✓	✗
Real data	✓	✓	✓	✓	✓
Generated data	✗	✗	✗	✗	✓
Flipped Classroom	✗	✓	✗	✓	✗
Differential teaching	✗ Don't fear feedbacks	✓ Active teaching methods	✗ Stories Competition facilitator role	✓ GenAI, Chatbots Oral exams	✗ Individual works peer review



Kutatás

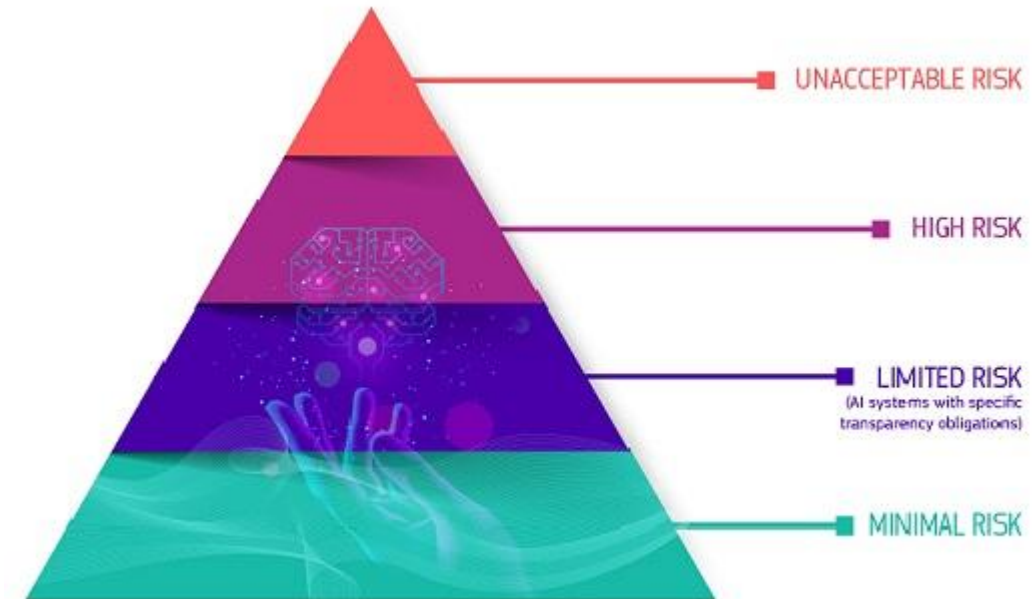
- Összegzés
- Elemzés





Kihívások

- Melyik rendszert használjuk?
- Órai aktivitás: licenszek?
- AI-támogatott értékelés
- AI Act [Regulation (EU) 2024/1689]
 - Kockázat alapú megközelítés,
 - Define AI literacy
- Kockázatok (pld. adatvédelem)



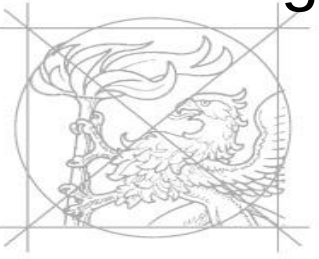
Forrás: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

Oktatási intézményekben használt mesterséges intelligencia-megoldások, amelyek meghatározhatják valakinek az oktatáshoz való hozzáférését és szakmai életének menetét (pl. vizsgák pontozása).



Összegzés

- Nincs legjobb rendszer
- Jó kisegítő eszköz, de nem mindenkinek
- Nem helyettesíti a munkánkat
- Újfajta szerepkörök oktatóknak, kutatóknak
- AI jártasság fejlesztése
- Nagyon sok kockázata van





Köszönöm a figyelmet

kovacs.peter@eco.u-szeged.hu