

Laboration - Skolstatistik

Vi hör mycket i media om att Sveriges PISA-resultat blir sämre och att svenska skolan tappar i kunskapsnivå. Låt oss undersöka lite närmre hur det ser ut med hjälp av nationella prov i årskurs 9.

Data

Skolverket har samlat mycket statistik kring nationella proven i årskurs 9 som vi kommer att analysera vidare.

Uppgifter

Alla grafer ska exporteras till en undermapp som heter **visualiseringar** i ditt repo för labben. Exportera matplotlib/seaborn-grafer i png-format och plotly-grafer som html-filer. Ge relevanta namn till dina filer så det blir lätt att referera till dem. Notera att du behöver ställa in bakgrundsfärg när du sparar matplotlib-grafer, annars blir de transparenta och svåra att läsa.

Tänk på data storytelling, graferna du skapar ska ha lämpliga rubriker, labels, annoteringar, clutter ska tas bort, jobba med attention mha färger mm.

Uppgift 0 - uppvärmning med riket

I den här uppgiften ska du visualisera med hjälp av matplotlib och/eller seaborn

a) Börja med att läsa in filen riket2023_åk9_np.xlsx och sheets för de olika ämnena. Ändra kolumnnamnen på respektive dataframe så du får något liknande som nedan.

	Plats	Huvudman	Totalt (A-F)	Flickor (A-F)	Pojkar (A-F)	Totalt (A-E)	Flickor (A-E)	Pojkar (A-E)	Totalt (poäng)	Flickor (poäng)	Pojkar (poäng)
0	Riket	Samtliga	93225	45800	47425	96.1	97.9	94.4	13.6	14.5	12.7
1	Riket	Kommunal	72771	35484	37287	95.7	97.6	93.8	13.3	14.3	12.4
2	Riket	Enskild	20414	10297	10117	97.6	98.7	96.6	14.4	15.2	13.5
3	Riket	Skolverket	40	19	21	100	100	100	15	15.9	14.2

b) Ta reda på antalet som fått F i matematik totalt, pojkar respektive flickor.

c) Rita ett stapeldiagram på totala poängen för de olika ämnena.

d) Rita fyra olika stapeldiagram i subplots på totala poängen för de olika ämnena för de olika huvudmännen.

e) Rita ett stapeldiagram i subplots på totala poängen för de olika ämnena för de olika huvudmännen.

Uppgift 1 - uppvärmning betyg och prov

I uppgifterna nedan ska du använda dig av datasetet betyg_o_prov_riksnivå.xlsx. Använd Plotly för att rita diagrammen nedan för de uppgifter som kräver diagram. Använd Pandas för att svara på frågor om datasetet.

a) Rita ett linjediagram för andel elever som saknar godkänt betyg i ett eller fler ämnen för läsårs 18-23. Ta med totalt, flickor och pojkar i samma graf.

b) Rita ett linjediagram för meritvärde för 16 ämnen för totalt, flickor och pojkar för läsår 18-23.

c) Undersök nu andra variabler, och hur de påverkar meritvärde och andel med minst ett saknat betyg. Data finns för läsår 22/23. Exempel kan man undersöka föräldrarnas utbildningsbakgrund. Rita lämpliga grafer.

Uppgift 2 - KPlar & EDA

Definiera 3-6 KPlar utöver de som använts i ovanstående uppgifter och gör explorativ dataanalys på en eller flera dataset som ni får söka själva. Ni får inte ta dataseten som ni redan jobbat med i uppgift 0 och 1. Excel och eller csv-filer kan ni hitta i [skolverkets statistik](#)

Se till att dokumentera vad du undersöker med markdown kombinerad med kod i Jupyter notebook eller om du använder Pythonskript, skriv en markdown vid sidan om. Dokumentera också vad du kommer fram till och spara alla visualiseringar.

Bedömning

Godkänt

- Löst uppgifterna 0-2 på korrekt sätt
- Variabelnamn är bra valda
- Koden är kommenterad med relevanta kommentarer
- Dokumenterat tydligt
- Committar ofta med bra commit-meddelanden
- Visualiseringarna är tydliga

Väl Godkänt

- Det finns bra data storytelling i visualiseringarna
- Dokumenterat väl och analyserat datan för att få fram viktiga insikter
- Koden är välstrukturerad och kod återanvänds (exempelvis med funktioner och/eller OOP) för att få till DRY
- Lätt att följa både kod, dokumentation och visualiseringar