

Labb 2 - Pandas och data processing

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.

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 grafer i png-format och plotly-grafer som html-filer. Ge relevanta namn till dina filer så det blir lätt att referera till dem.

Tänk på data storytelling, graferna du skapar ska ha lämpliga rubriker, labels, annoteringar mm.

Uppgift 1

I den här uppgiften ska du visualisera med hjälp av matplotlib

a) Börja med att läsa in filen riket2023_åk9_np.xlsx och sheets för de olika ämnena (engelska, matematik, svenska, svenska som andraspråk). 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) För de olika ämnena, rita en stapeldiagram på totala poängen för de olika huvudmännen. Placera alla diagram i en plott med subplottar

Uppgift 2

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 läsåret 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 läsåret 18-23.

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

Grading

If you have taken ideas of codes from someone or found them online, it is **important** that you state the source and understand how the codes work. Write a comment next to these codes.

Criteria for G:

- solved all non-bonus tasks correctly
- added relevant comments on the codes

- the codes are well-structured

Criteria for VG:

- solved all tasks including bonus tasks correctly

Submission

- create a public github repo when you start working with this lab
- the repo should show that you have made commits into it
- when you are done, send the repo URL to Learnpoint by the deadline