

# Laboration - Covid-19

---

Covid-19 pandemin har påverkat hela världen med många dödsfall som följd. Ett kraftfullt verktyg för att förstå hur viruset påverkar världen är att analysera och visualisera data för att förstå trender.

---

## Data

Folkhälsomyndigheten har samlat in data kring nya fall, dödsfall med mera. Vi ska använda oss av följande dataset:

- [Covid-19 bekräftade fall](#) - refererad som **covid19** i texten nedan
- [Statistik för vaccination mot covid-19](#) - refererad som **vaccin** i texten nedan

Under rubriken "Ladda ner data" ska du ladda ned Excel-filerna.

---

## Uppgifter

Alla grafer ska exporteras till en undermapp som heter **Visualiseringar** i din mapp för labben. Exportera matplotlib/seaborngrafer i png-format och plotly-grafer som html-filer. Ge relevanta namn till dina filer så det blir lätt att referera till dem.

---

### Uppgift 1 - uppvärmning covid-19 data

Börja med att manuellt läsa Excel-filerna och de olika bladen i varje dokument. Använd därefter Pandas för att läsa in bladet: "Veckodata Riket" i covid19-filen.

a) Gör initial dataanalys för att snabbt få en överblick över datasetet. Använd metoder som ex. info(), describe(), value\_counts(), head(), columns, index för att snabbt få en överblick.

b) Slå ihop kolumnerna "år" och "veckonummer" till en kolumn med namn "Vecka" med följande format:

```
Vecka
2020v6
2020v7
...
2021v40
```

För deluppgifterna c-f, använd både Seaborn och Plotly express. För Seaborn, använd subplots så du får 2x2 grid med graferna.

c) Rita en linjediagram som visar **avlidna fall per vecka** från 2020v6 till nu.

d) Rita en linjediagram som visar **nya fall per vecka** från 2020v6 till nu.

e) Rita linjediagram av både **avlidna fall per vecka** och **nya fall per vecka** i samma fönster från 2020v6 till nu.

f) Rita linjediagram av **kumulativt antal fall** från 2020v6 till nu.

---

## Uppgift 2 - uppvärmning vaccindata

I uppgifterna nedan ska du använda dig av datasetet vaccin och bladet **Vaccinerade kommun och ålder**. Använd Plotly express för att rita diagrammen nedan för de uppgifter som kräver diagram. Använd Pandas för att svara på frågor om datasetet.

a) Hur många län finns representerade i datasetet?

b) Hur många kommuner finns representerade i datasetet?

c) Hur stor är befolkningen som är representerad i datasetet?

d) Beräkna hur många barn under 16 år det finns i Sverige. Du får leta upp statistik på hur stor totala befolkningen är i Sverige.

e) Rita stapeldiagram för **andel med minst 1 dos per län** och **andel färdigvaccinerade per län**

f) Rita ett stapeldiagram med län i x-axeln och staplar för **befolkning > 16år**, **antal minst 1 dos** och **antal färdigvaccinerade**.

---

## Uppgift 3 - KPlar & explorativ dataanalys

Definiera 3-6 KPlar utöver de som använts i ovanstående uppgifter och gör explorativ dataanalys på flera sheets på båda Excel-filerna. 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.

---

## Uppgift 4 (bonusuppgift)

Följ länkarna från [FHM](#) för att komma till ECDC, WHO och undersök Covid-19 globalt. Dokumentera vad du undersöker och vad du kommer fram till.

Notera att du får navigera och läsa till dig i deras hemsidor för att hitta relevant data.

---

## Bedömning

### Godkänt

- Löst uppgifterna 1-3 på korrekt sätt
- Variabelnamn är bra valda
- Koden är kommenterad med relevanta kommentarer
- Dokumenterat tydligt
- Committar ofta med bra commit-meddelanden

## Väl Godkänt

- Löst samtliga uppgifter
- Dokumenterat väl
- Återanvänt kod väl (exempelvis med funktioner)
- Lätt att följa både kod, dokumentation och visualiseringar