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. 分析和可视化数据以了解趋势。

Data

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

瑞典公共卫生局收集了有关新病例、死亡人数等的数据。 我们将使用以下数据集：

* 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.

所有图形都应导出到实验室文件夹中名为 Visualizations 的子文件夹中。 以 png 格式导出 matplotlib/seaborne 图形，并将绘图图形导出为 html 文件。 为您的文件指定相关名称，以便于引用它们。

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.

首先手动阅读 Excel 文件和每个文档中的不同工作表。 然后使用 Pandas

加载工作表：covid19 文件中的“每周数据王国”。

1. 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.

执行初始数据分析以快速获得数据集的概览。 使用方法如。 信息（），

describe()、value\_counts()、head()、columns、index 快速获得概览。

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

b) 按以下格式将“年”和“周数”列合并到名为“周”的列中：

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 - KPIer & explorativ dataanalys

Definiera 3-6 KPIer 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.