

Laboration VIS2

732G48 - Visualisering

Linköpings universitet
Institutionen för datavetenskap, IDA
Avdelningen för Statistik och maskininlärning, STIMA

Introduktion

Se till att ni sitter i de grupper som är angivna i Excelfilen i *LISAM*. Under *Inlämningar* kan ni även finna deadline för laborationen och det tillhörande datamaterialet, och det är även här ni ska lämna in er slutgiltiga rapport. Detta är ett examinerande moment i kursen och alla laborationer måste vara godkända innan betyget **G** eller **VG** ges på hela kursen.

Er uppgift med dessa instruktioner är att skapa snygga, korrekta och tydliga diagram, **men glöm inte bort att tolka dessa!** Utan tolkning av diagrammen kommer rapporten inte bli godkänd.

Datamaterial

Under denna laboration kommer först lönestatistiken från NotReal AB användas igen (*VIS1 Lön.xlsx*). För mer information om dessa siffror kan ni läsa i laboration 1.

I filen *VIS2 COVID-19 Oct2021.xlsx* redovisas det siffror över antalet bekräftade fall av COVID-19 i Region Östergötland, Stockholm samt totalt i hela Sverige mellan 2020-02-04 och 2021-10-07. Datamaterialet är hämtat från [Folkhälsomyndigheten](#) den 12 Oktober 2021.

I filen *VIS2 Husdjur Sep2023.xlsx* redovisas det totala antalet kor, får och svin i Sverige per år från 1961-2007. Datamaterial är hämtat från SCBs statistikdatabas den 17 September 2023.

Uppgifter

Uppgift 1 - Samband mellan ålder och lön vid NotReal AB

Datamaterialet från NotReal AB ska analyseras, och nu ska *sambandet* mellan ålder och lön på arbetsplatsen undersökas.

- Skapa ett spridningsdiagram i **SPSS**, med ålder på x-axeln och lön på y-axeln. Vad kan ni säga om sambandet mellan ålder och lön?

Ledning: Se *föreläsning 1* om spridningsdiagram och vad vi kan utreda samt dokumentet *Tolkningstips*.

- Skapa ett spridningsdiagram i **RStudio**. Vad kan ni säga om sambandet utifrån detta diagram?

Ni bör se en avvikande observation (extremvärde), nämligen att en ung person tjänar väldigt mycket pengar. Vad kan vi säga om personen utifrån diagrammet?

Uppgift 2 - Antal bekräftade fall av COVID-19 i Sverige

Vi ska nu behandla det COVID-19 relaterade datamaterialet från Folkhälsomyndigheten.

- a) Skapa ett linjediagram i **RStudio** för att visualisera utvecklingen av det totala antalet bekräftade fall av COVID-19 i *Sverige*. Finns det några intressanta mönster eller vändpunkter? Tydliga toppar?

Ledning: `scale_x_datetime(date_breaks = "4 month", date_labels = "%y %b %d")`

OBS! Skriv in kod från ledning på egen hand, kopiera inte!

- b) Skapa ett linjediagram i **Excel** för att visualisera utvecklingen av antalet bekräftade fall av COVID-19 i *Östergötland*. Finns det några intressanta mönster eller vändpunkter? Tydliga toppar?

Uppgift 3 - Samband mellan olika husdjur

Vänd er nu till det datamaterial från SCBs Statistikdatabas angående husdjur.

- a) Skapa en punktdiagramsmatris i **RStudio** och undersök de parvisa sambanden mellan de tre husdjuren. Tolka sambanden.

Ledning: Kom ihåg att ta bort variabeln "År" från datamaterialet.

Att lämna in

Alla diagram och tillhörande tolkningar ska sammanställas till en tydligt strukturerad rapport, där vardera visualisering följer den checklista som anges i *LISAM*. Använd den rapportmall som hittas i *Kursdokument* på *LISAM*. Rapporten ska lämnas in via *Inlämningar* innan deadline på inlämningstillfället.