

# Inlämningsuppgift 1

## Instruktioner

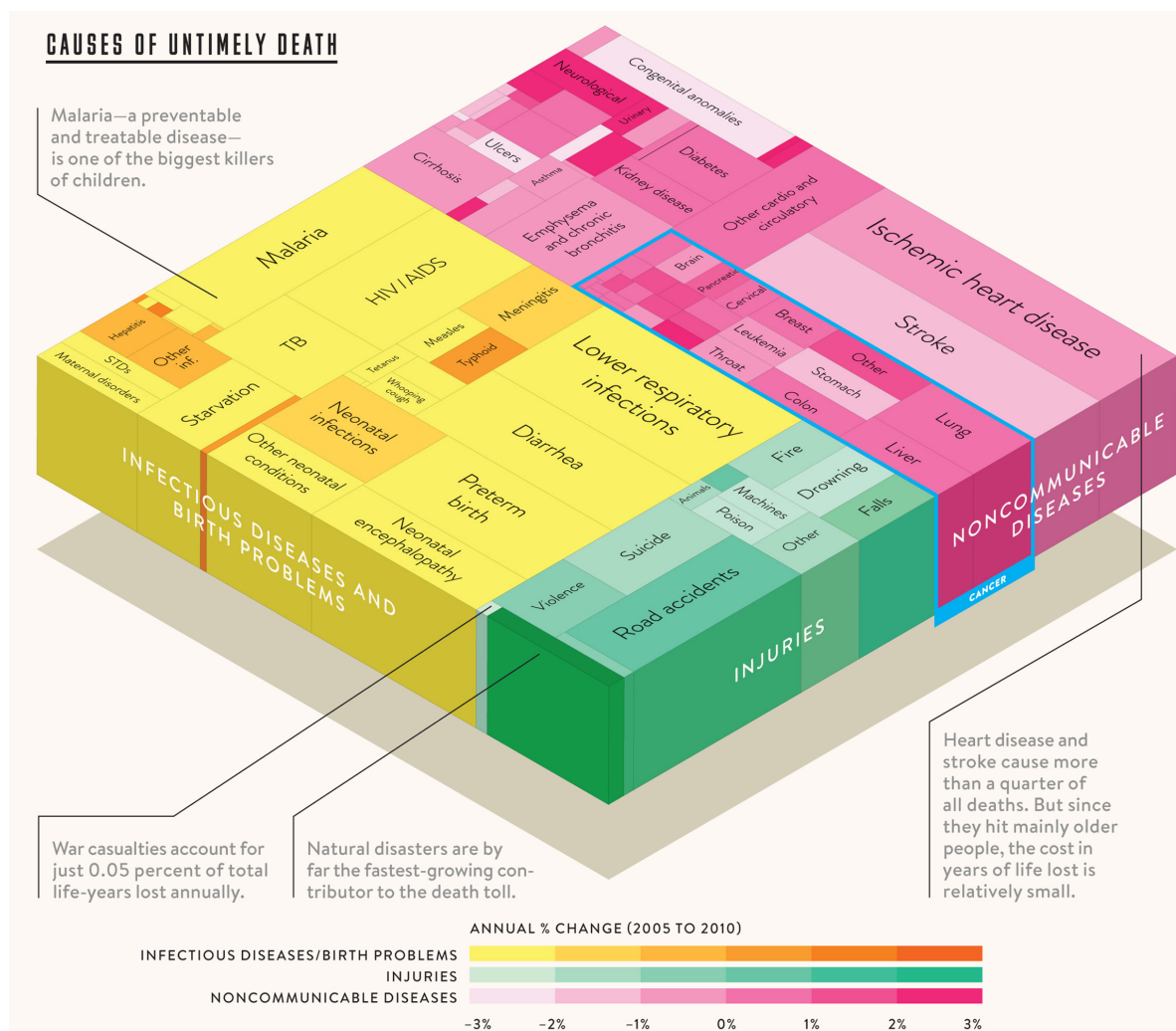
Lös uppgifterna på följande sidor *i er grupp* och lämna in följande i *Canvas > Inlämningsuppgifter*:

- Ett Quarto-dokument med filnamn (ex. för Grupp\_1): Grupp\_1.qmd där ni svarar på frågorna och skapar kod för visualiseringarna. **Kom ihåg att ange namn på dem som deltagit i arbetet.**
- Ett pdf-dokument som är kompilerat från qmd-filen (eller sparat som pdf från en html- eller word-version av Quartodokumentet). **Visa er R-kod i rapporten.**

## Uppgift 1

I Washington Post (2013-12-27) finns en artikel där Microsofts grundare Bill Gates presenterar sitt favoritdiagram från året (Figur 1).

Bill Gates säger följande: “I love this graph because it shows that while the number of people dying from communicable diseases is still far too high, those numbers continue to come down. In fact, fewer kids are dying, more kids are going to school and more diseases are on their way to being eliminated. But there remains much to do to cut down the deaths in that yellow block even more dramatically. We have the solutions. But we need to keep the up support where they’re being deployed, and pressure to get them into places where they’re desperately needed.”



Figur 1: Treemap – Förtida död

Se t.ex. <https://en.wikipedia.org/wiki/Treemapping> för en beskrivning av vad en treemap är.

**Frågor att besvara:**

- a) Vilka information ger visualiseringen?
- b) Vilka variabler beskrivs med hjälp av
  - storlek på areor
  - färger
  - färgtoner (mörka och ljusa toner)?
- c) Hur påverkar 3D-effekten möjligheten att jämföra areornas storlekar?
- d) Beskriv ytterligare styrkor/svagheter med visualiseringen utifrån olika principer om visualisering.
- e) Kan du föreslå några förbättringar? Hur skulle detta kunna ha gjorts bättre?

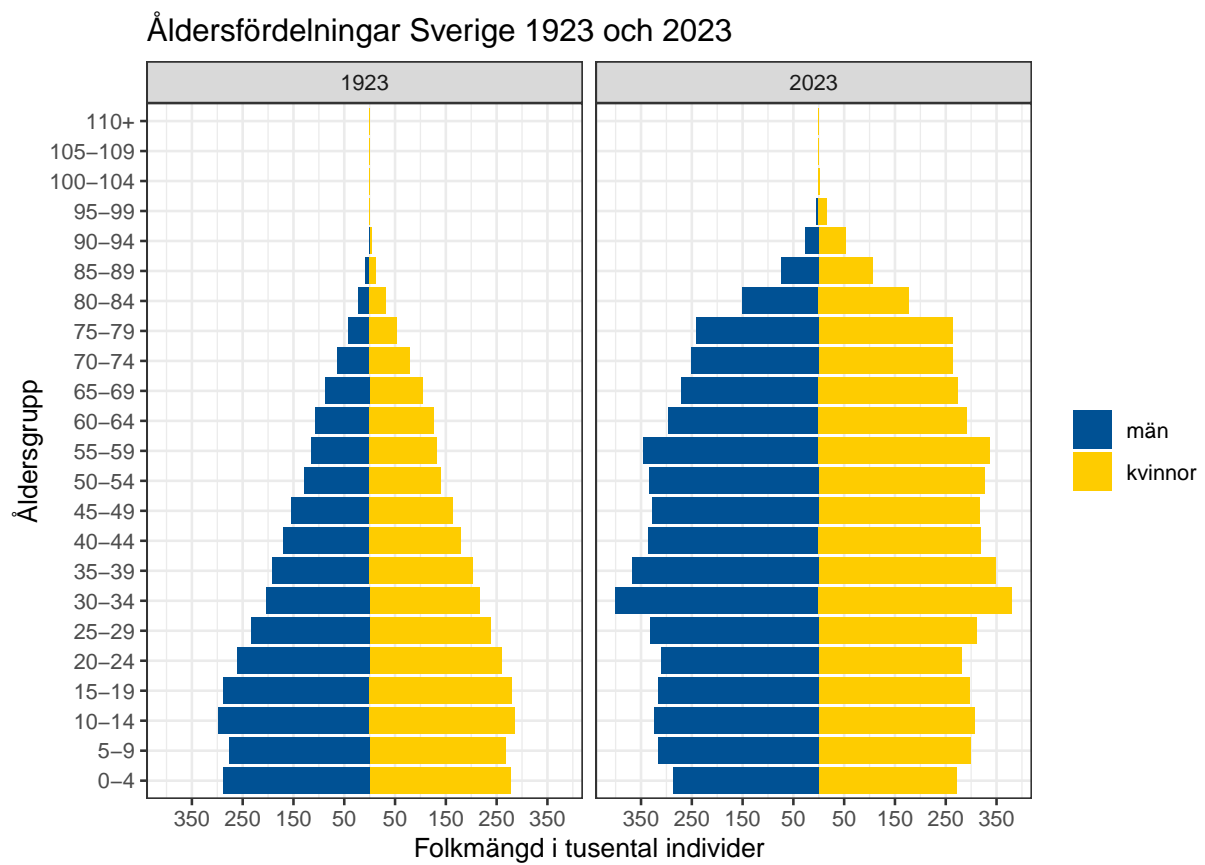
## Uppgift 2

Datamaterialet `population_1923_2023` är nedladdat från [Statistikdatabasen](#) hos Statistiska centralbyrån (SCB) och därefter något modifierat för att underlätta visualisering. Materialet innehåller Sveriges folkmängd i åldersgrupper (0-4, 5-9,..., 105-109, 110+) för åren 1923 och 2023.

Skapa befolkningspyramider för Sverige för åren 1923 och 2023 med R och `ggplot2`. Resultatet bör likna Figur 2. Utgå gärna från koden i “chunken” nedan.

```
pyramid <- ggplot(data = population_1923_2023,
                  aes(x = count,
                      y = reorder(age, order),
                      fill = sex)) +
  geom_col(data = subset(population_1923_2023, sex=="men"),
           aes(x = -count)) +
  geom_col(data = subset(population_1923_2023, sex=="women"),
           aes(x = count)) +
  scale_x_continuous(breaks = ...,
                     labels = c(...)) +
  scale_fill_manual(values = c(...),
                    labels = c(...)) +
  ... +
  ... +
  facet_wrap(~year) +
  theme...

pyramid
```



Figur 2: Sveriges befolkning 1923 respektive 2023.

### Uppgift 3

Visualisera förändringen i åldersfördelning i den svenska befolkningen under perioden 1860-2023 med `ggplot` och datamaterialet `population_1860_2023`. Ni väljer själva hur ni på ett bra sätt ska visualisera utvecklingen.

---

#### **i** Notera

Datamaterialen i Uppgift 2-3 har vissa brister, särskilt för de tidigare åren, se kommentarer [här](#).