

Koding med AI

Kristine Vanebo-Bøe

Innholdsfortegnelse

1 Koding med AI

1

1 Koding med AI

AI-verktøy som *ChatGPT* kan være svært nyttige hjelpemidler for studenter som lærer seg å programmere i R. Modellen kan forklare kode, gi forslag til løsninger, hjelpe deg å forstå feilmeldinger og vise hvordan ulike funksjoner fungerer i praksis. Men for å få best mulig utbytte, er det viktig å vite hvordan man bruker verktøyet riktig og kritisk.

Det aller viktigste er å være spesifikk når du stiller spørsmål. *ChatGPT* gir mye bedre svar når du tydelig forklarer hva du holder på med. For eksempel bør du skrive hvilke pakker du allerede har lastet inn (for eksempel `tidyverse` eller `bst290`), hvilket datasett du bruker, hvilke variabler som er relevante, og hva du allerede har prøvd. Da slipper du generelle svar, og får heller kode som passer til akkurat ditt prosjekt.

Et vanlig problem er at *ChatGPT* ofte legger inn kode for å laste inn pakker – som `library(tidyverse)` – hver gang du spør om noe. Det er ikke nødvendig å kopiere dette inn hver gang i R-skriptet ditt. Du trenger bare å laste inn pakkene én gang øverst i koden. Når man ser mange slike `library()`-linjer flere steder i en oppgave, kan det dessuten være et tydelig tegn på at *ChatGPT* har blitt brukt uten redigering. Det kan derfor være lurt å rydde litt i koden og fjerne unødvendige gjentakelser før du leverer.

En annen ting man må være oppmerksom på, er at *ChatGPT* noen ganger foreslår løsninger som bruker pakker du ikke har installert, eller som ikke brukes i kurset ditt. Verktøyet har tilgang til et bredt utvalg av R-pakker, og foreslår derfor ofte funksjoner fra verktøy som ikke nødvendigvis er relevante. Da er det viktig å bruke egen vurdering: sjekk alltid at pakken faktisk finnes, og at du forstår hva koden gjør. Hvis ikke, bør du heller spørre om en løsning som kun bruker de pakkene dere bruker i undervisningen.

Ved å simpelt spørre *ChatGPT* om den kan hjelpe deg med å lage et histogram i R vil den for eksempel gi deg forslag på om du vil bruke **Base R** eller **ggplot2**. I kurset bruker vi ikke **Base R**, og vi bruker også andre pakker enn **ggplot2** (**ggplot2** er inkludert **tidyverse**). I tillegg vil den ikke kunne gi deg en kode som er tilpasset dine behov, med datasettet du bruker samt variablene du ser på. Denne koden vil derfor ikke gå an å bruke i R.

kan du hjelpe meg lage et histogram i r?

Selvfølgelig! 😊

Jeg kan vise deg akkurat hvordan du lager et histogram i R — både i base R og med **ggplot2** (som gir finere grafer).

La meg bare spørre først:

Vil du lage histogrammet i **base R** eller med **ggplot2** (fra **tidyverse**)?

Og har du allerede et datasett du vil bruke — f.eks. en kolonne du vil lage histogram av?

For eksempel, hvis du har en variabel som heter **vekker** i et datasett som heter **data**, så kan du gjøre:

◆ Base R:

```
r  
  
hist(data$vekter,  
      main = "Histogram av vekter",  
      xlab = "Vekt (kg)",  
      col = "lightblue",  
      border = "white")
```

 Kopier kode

◆ Med ggplot2:

```
r  
  
library(ggplot2)  
  
ggplot(data, aes(x = vekter)) +  
  geom_histogram(binwidth = 2, fill = "skyblue", color = "white") +  
  labs(title = "Histogram av vekter", x = "Vekt (kg)", y = "Antall")
```

 Kopier kode

Ved å spesifisere at jeg har lastet inn pakkene og at jeg bruker `tidyverse`, gir R meg en bedre kode på et grunnleggende histogram, men det er fortsatt upresist og jeg vil ikke kunne bruke koden direkte inn i R.

Perfekt 🙌 Da bruker vi **ggplot2** (som følger med i tidyverse).

Her er et generelt oppsett du kan bruke for å lage et histogram i tidyverse:

◆ Grunnleggende histogram

r

 Kopier kode

```
ggplot(data = ditt_datasett, aes(x = din_variabel)) +  
  geom_histogram(binwidth = 1, fill = "skyblue", color = "white") +  
  labs(title = "Histogram av variabelen",  
        x = "Navn på variabel",  
        y = "Antall observasjoner") +  
  theme_minimal()
```

◆ Forklaring:

- `data = ditt_datasett` → bytt ut med navnet på datasettet ditt
- `aes(x = din_variabel)` → bytt ut med kolonnen du vil lage histogram av
- `binwidth = 1` → bestemmer bredden på søylene (kan justeres for å få mer/mindre detaljer)
- `fill` og `color` styrer fargene
- `theme_minimal()` gir et rent design



datasettet er hentet fra ess runde 7, og jeg bruker variablene
essround, cntry idno, agea og rlgdgr. Landet jeg bruker er Norge

Hvis jeg så legger til hvilket datasett jeg bruker, variablene jeg vil se på og landet jeg vil forholde meg til gir den meg en mye mer utfyllende kode på et histogram som jeg kan bruke direkte inn i R.

```
r
library(tidyverse)

# Filtrer til Norge
ess_no <- ess %>%
  filter(cntry == "NO")

# Sjekk datastruktur
glimpse(ess_no)

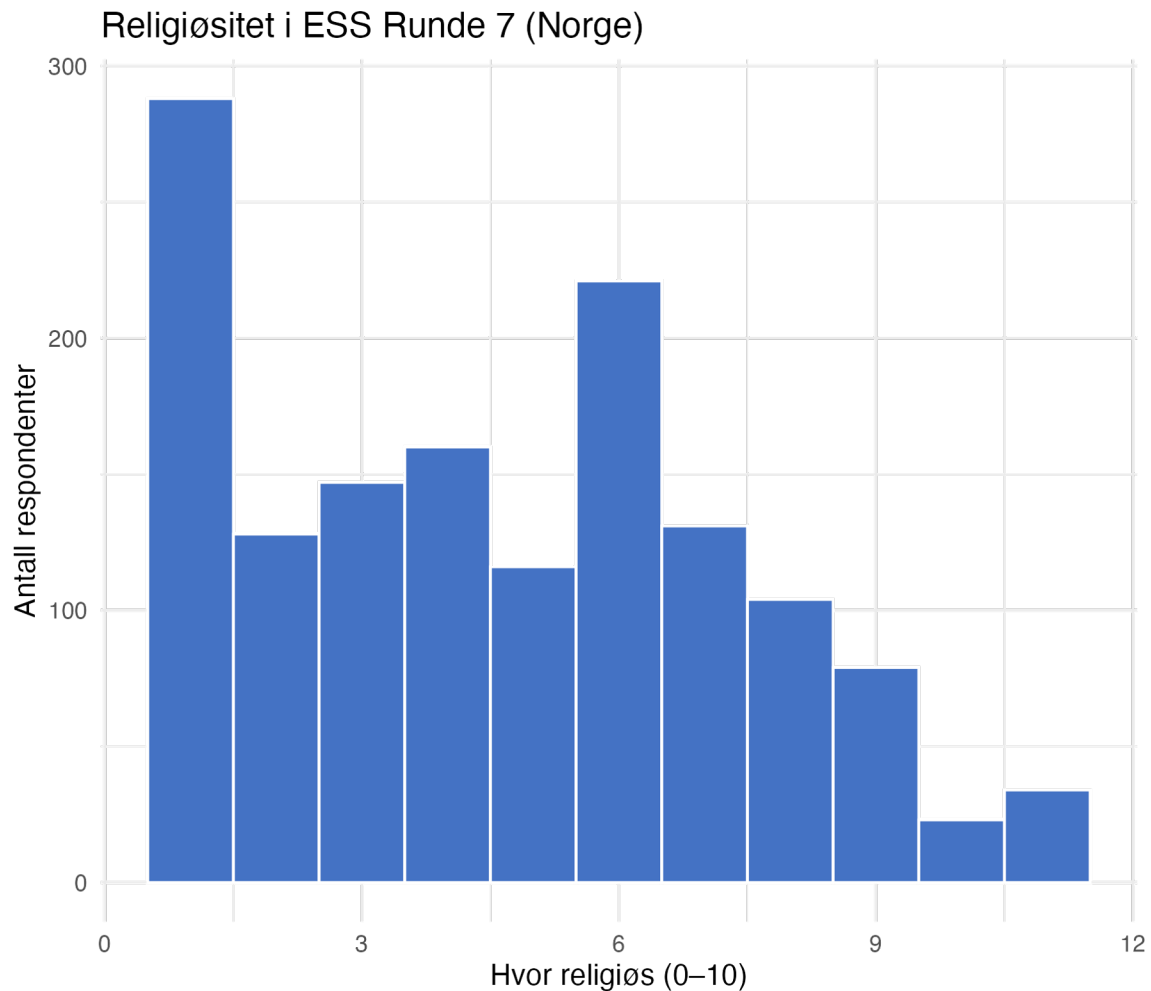
# Histogram av alder
ggplot(ess_no, aes(x = agea)) +
  geom_histogram(binwidth = 5, fill = "#69b3a2", color = "white") +
  labs(title = "Aldersfordeling i ESS Runde 7 (Norge)",
        x = "Alder",
        y = "Antall respondenter") +
  theme_minimal()

# Histogram av religiøsitet
ggplot(ess_no, aes(x = rlgdgr)) +
  geom_histogram(binwidth = 1, fill = "#4472C4", color = "white") +
  labs(title = "Religiøsitet i ESS Runde 7 (Norge)",
        x = "Hvor religiøs (0-10)",
        y = "Antall respondenter") +
  theme_minimal()
```

Kopier kode

Den gir meg da to koder til histogram av både `rlgdgr` og `agea`, som man kan kopiere rett inn i R. Til tross for at jeg har fortalt *ChatGPT* at jeg allerede har lastet inn `tidyverse`-pakken inkluderes den før koden til selve grafen. I tillegg får jeg feilmelding i R når jeg prøver å bruke koden til histogram av religiøsitet, fordi den ikke har lagt inn at variabelen `rlgdgr` er nødt til å bli gjort om til numerisk for å utføre koden.

Resultatet av histogrammet av religiøsitet ser slik ut og er fullt mulig å bruke i oppgaver for å vise fordelingen av religiøsitet:



For å gjøre det enklere å bruke *ChatGPT* effektivt kan du lage en fast struktur når du spør. Du kan for eksempel skrive: “Jeg jobber i R med datasettet ESS runde 7, og har lastet inn `tidyverse` og `bst290`. Variablene jeg bruker er `essround`, `cntry`, `idno`, `agea` og `rlgdgr`. Jeg skal se på data fra Norge. Jeg har trimmet datasettet til å kun inkludere disse variablene fra Norge og navngitt det `ess7`, samt fjernet NAs. Jeg vil lage et univariat histogram som viser fordelingen av `rlgdgr`. Kan du lage en kode for dette, med 10 bins i tema svart/hvitt, og at teksten på y-aksen er «Number of observations» og på x-aksen «How religious are you (0-10)»?»

Klart! Her er en enkel og ren ggplot-kode som lager et univariat histogram for `rlgdgr` med 10 bins, svart/hvitt-tema, og de aksenavnene du ønsker:

```
r Kopier kode

library(tidyverse)

ggplot(data = ess7, aes(x = rlgdgr)) +
  geom_histogram(color = "white", bins = 10) +
  labs(
    x = "How religious are you (0-10)",
    y = "Number of observations",
    title = "Distribution of Religiosity in Norway (ESS Round 7)"
  ) +
  theme_bw()
```

Til slutt er det viktig å huske at *ChatGPT* er et støtteverktøy, ikke en erstatning for egen forståelse. Det beste er å bruke svarene som forklaringer eller inspirasjon, og ikke bare kopiere koden direkte. Ved å kombinere egne ferdigheter med AI-hjelp lærer du både mer og jobber mer effektivt – uten å miste kontrollen over det faglige innholdet.