

SOK-1004 H24 Forelesning 2

RStudio

Derek J. Clark

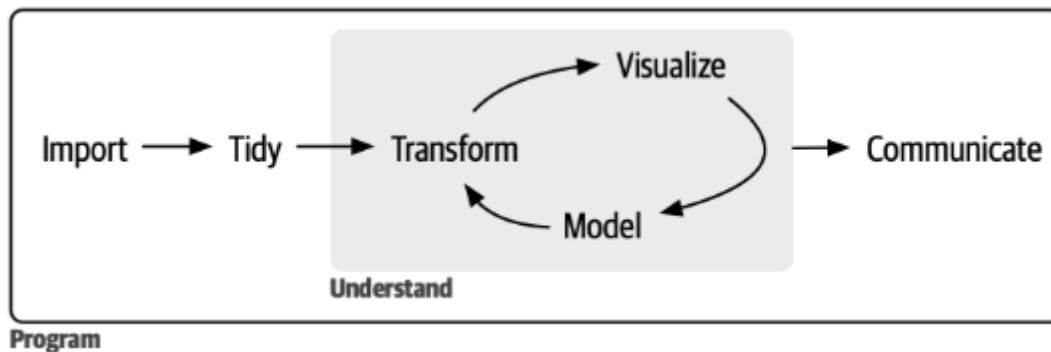
1. Innledning

Vi ønsker å forvandle ustrukturert informasjon om økonomiske forhold til sikker, troverdig kunnskap. Vi bruker **R** som er et programmeringsspråk med åpen kilde som er ganske lett å bruke og som er i kontinuerlig utvikling. **RStudio** er verktøyet som vi bruker for å kjøre kommandoer i R (samt skrive dokumenter, presentasjoner osv).

Først finner vi et spørsmål vi har lyst å svare på, for eksempel:

- hvor mange husholdninger lever under fattigdomsgrensen i Tromsø kommune?
- er inntektene til personer jevnere fordelt på Kvaløya eller på fastlandet?

Så leter vi etter data for å besvare vårt spørsmål:



[Kilde](#)

2. Oversikt i RStudio

Det er fire vinduer:

1. Konsoll - for å kjøre kode direkte
2. Tekstredigerer - skrive og kjøre større mengder med kode og tekst osv
3. Øvre faner - minne, tidligere kjørt kode
4. Nedre faner - filer, grafikk, pakker med mer

3. Funksjoner og pakker

- Funksjon: En regel $f()$ som tar et argument X og tilordner en unik verdi $Y : f(X) = Y$
- Eksempel: La $X = \{1, 1, 2, 3\}$ en liste med tall. Funksjonen `sum()` summerer tallene: `sum(X)=7`
- R-kjernen (Base-R): Funksjoner som ligger til grunn i R
- R-pakker: Samlinger av funksjoner laget av utviklermiljøet til R, må lastes opp

Vi skal ofte bruke `tidyverse` som er et bibliotek av pakker som inneholder blant annet

- `dplyr` som hjelper med transformasjon og bearbeiding av data
- `ggplot2` som gir visualisering i form av figurer

4. Komme i gang med koding

Den beste måten å lære koding på er å kode selv! ChatGPT/ChatUiT er et veldig nyttig hjelpemiddel når du har lært noe om koding.

Pensumboka [R for Data Science \(2e\)](#) er en viktig ressurs med mange gode forklaringer og eksempler.

5. Installere pakker

En pakke må installeres før du kan bruke den. Du kan gjøre det i konsollen ettersom pakken behøves installasjon kun én gang. Etter at pakken er installert kan du kaller opp dens bibliotek for å bruke den.

I dag skal vi bruke `tidyverse` for å manipulere og visualisere data. Selve datasettet skal vi hente fra en pakke som heter `Ecdat` . [Ecdat](#)

I konsollen skriv: `install.packages("tidyverse")` og trykk Enter, skriv deretter `install.packages("Ecdat")` og trykk Enter. NB du kan installere pakker fra “Packages” i nedre fane, eller du kan installere flere samtidig ved å bruke `c()` (som betyr “concatenate”, sette sammen på norsk): `install.packages(c("tidyverse", "Ecdat"))`

6. Last inn `SOK-1004_Forelesning_2_h24.R` og kom i gang!