



grunnkurs

www.sli.do
#461161

SUSIE JENTOFT & ASLAUG FOSS



Statistisk sentralbyrå
Statistics Norway

Mål

- Kjennskap til programmet R og RStudio
- Åpne RStudio og kjøre enkle beregninger
- Lese inn data
- Behandler data
- Lage tabeller og oppsummere
- Lage figurer



Agenda

Dag 1: 20. september 2021

12:00 – Introduksjon til kurs og R

12:45 – Øvelse 1

13:30 - Les inn forskjellige data

14:00 – Øvelse 2

Dag 2: 22. september 2021

12:00 – Oppsummering

12:15 – Databehandling og tabeller

12:45 – Øvelse 3

13:30 – Kobling og figurer

14:00 – Øvelse 4

14:20 – Løsninger og oppsummering



Kjøreregler

- **Bruk chat** til å stille spørsmål – Aslaug/Tatsiana sjekke
- **Rekk opp hånd** i Teams om det er noe som er uklart
- **Øvelser: 3 chatrom** i Teams. Snakk gjerne med hverandre eller skrive i chaten.
- **Demper lyd** når dere ikke snakker men gjerne har på video for å kunne ser hverandre.



Deres bakgrunn

<https://admin.sli.do/event/60a9ozqz/polls>



Hva er



?

- Programmeringsspråk og miljø for statistikk
- Utviklet Ross Ihaka & Robert Gentleman (1993)
- Base + tilleggspakke
- RStudio:
 - IDE/Utviklingsmiljø
 - Åpen source + kommersielle



©EnchantedLearning.com



Statistisk sentralbyrå
Statistics Norway

Hvorfor ?

- Åpen source/gratis
- Stor samfunn rundt verden
- God grafikk
- Utviklet for statistikk
- Brukt internt og i andre statistikk byrå:
 - <https://github.com/SNStatComp/awesome-official-statistics-software>
- Alle sone i SSB **OG PÅ DAPLA**





Kjøre-knappen
(eller ctrl + enter)

The screenshot shows the R Studio interface with several annotations. A blue circle highlights the 'Run' button in the top toolbar, with a line pointing to the text 'Kjøre-knappen (eller ctrl + enter)'. Red boxes outline the four main panes: 'Source' (top-left), 'Environment/History' (top-right), 'Console/Terminal' (bottom-left), and 'Files/Plots/Pkgs/Help' (bottom-right). The 'Source' pane shows a script named 'genomics_r_basics.R' with a line number '1'. The 'Environment/History' pane shows 'Global Environment' with 'Environment is empty'. The 'Console/Terminal' pane shows the R version 3.4.4 (2018-03-15) and copyright information. The 'Files/Plots/Pkgs/Help' pane shows a file explorer view of the 'dc_genomics_r' directory, listing files 'dc_genomics_r.Rproj' (205 B) and 'genomics_r_basics.R' (0 B).

Source

Environment/History

Console/Terminal

Files/Plots/Pkgs/Help

Kode-fil

Åpne en ny kode fil

- File > New File > R script

Eller



Åpen en eksisterende fil

- File > Open File ...

Eller



Lagre fil

- File > Save

Eller



R og Git



- På Github: https://github.com/statisticsnorway/R_grunnkurs
 - Klikk på «Clone»
 - Kopiere adressen (HTTPS)
- I RStudio:
 - Velg File > New project > Version kontrol > Git
 - Og så lim inn url under «Repository URL»
 - Klikk på Create Project
 - Lagre filer du endre (for eks. øvelser) med et nytt navn

Eksempel kode for i dag

Eksempel kode for i dag heter **R_grunnkurs_dag1.R**

Det ligger i Github repository og i Teams under filer



Kjøring av kode

- R kan brukes som en kalkulator:

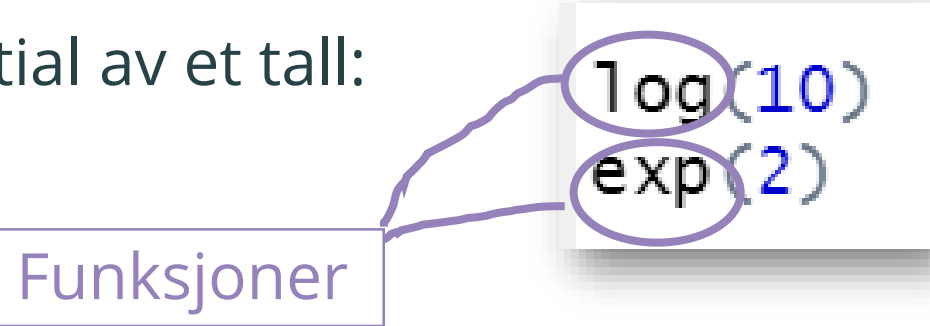
```
2 + 2  
5 - 2  
25 ÷ 4  
5 ^ 2
```

- Kjøre kode med trykke på  eller ctrl + enter
 - Dette kjøres kodelinje hvor en står
- # Brukes til kommentar og ikke kjøres

Grunnfunksjoner

Log og exponential av et tall:

Funksjoner



```
log(10)  
exp(2)
```

Alle funksjoner i R har hjelpfiler:

```
?log
```



String

```
"Hello, world!"
```

- Både enkel (' ') og dobbel snutter (" ") kan brukes – konsistent!
- CASE SENSITIVE
- Lim sammen med **paste()**
- Plukk ut deler med

```
substr(string, start, stopp)
```



Objekter

- Å lage et objekt er å gi det et navn
- Vi bruker `<-` for å allokere noe til et objekt

```
x <- 2
```

- Skriv ut hva innhold er ved å skrive objektnavn

Objekter

- CASE SENSITIVE
- MÅ starte med en bokstav
- Ikke mellomrom (bruk _)
- Forskjellige innhold: for eks. tall, eller string
- Kan overskrives

Kommune ≠ kommune

```
x <- 'Oslo'
```



Objekttype

- Finne ut mer med

```
str( )
```

- Sette navn av objektet inn i ()
- Vanlig objekttype:

Objekttype	Beskrivelsen
chr	Character/string
num	Nummer
Date	Dato
Factor	Kategorisk (bestemt nivå)
data.frame	Datasett



Vektorer

- Objekter som inneholder flere verdier.
- Bruk `c()` som står for concatenate og komma mellom
- Må være samme objekttype

```
c(1, 7, 10)
```

```
c("Kongsvinger", "Oslo")
```



Vektor som et objekt

- Gir vektorer et navn ved
- Finne lengde av vektoren ved funksjonen
- For å finne ut objekttypen bruk

```
<-
```

```
length()
```

```
str( )
```



Statistisk sentralbyrå
Statistics Norway

Øvelse 1:

- Øvelsene til oppgavesett 1 er på fil: **øvelser_1.R**



Logiske setninger

- Sammenlign objekter ved ==

- Skrives ut: TRUE/FALSE

- Kan brukes på vektorer

- Bind sammen med & for og, | for eller

Kode i R	Betydning
==	Sammenlign om de er like
!=	Om de er ulike
>	Større enn
>=	Større enn eller lik
<	Mindre enn
<=	Mindre enn eller lik
%in%	er inn i



Hvis eller...

- For å sammenlign og gjøre noen kan vi bruke funksjon

`ifelse()`

`ifelse(sted == "Oslo", "0301", "3401")`

condition

Om condition er riktig

Om condition
er feil

Lage et datasett

- Objekt kan være et datasett (rader og koloner)
- Kombinere forskjellige type vektorer

```
data.frame()
```

```
data.frame(objekt1, objekt2)
```

```
data.frame(kolonnenavn1=c(1,2), kolonnenavn2=c("banan", "eple"))
```

Se på datasett

- Skriv datasettnavn
- Klikk på datasett navn under Environment
- Eller



View()

R-bibliotek/tilleggpakker

- Samlinger av kode
- CRAN
- Installerer en gang (mange allerede installert)

```
install.packages("pakkenavn")
```



- Hentes inn hver gang

```
library(pakkenavn)
```



Les inn datasett: .csv fil

```
library(tidyverse)  
read_csv()
```

Stig til datasett

Skilletegn = , Desimaltegn = .

```
library(tidyverse)  
read_csv2()
```

Stig til datasett

Skilletegn = ; Desimaltegn = ,



Les inn datasett: SAS fil

- Vi kan lese inn sas datasett (.sas7bdat) direkte til R med haven pakke.

```
library(haven)  
read_sas()
```

Stig til datasett

```
?read_sas
```

Les inn datasett: Statbanken

- Hent inn Statbanken data direkte til R ved API.
- Trenger tabell nr.
- Spesifisere: ContentsCode, Tid osv.
- E-læringskurs:

<https://ssbviz.shinyapps.io/ApiData/>

```
library(PxWebApiData)  
ApiData()
```



Ser nærmere på data

```
glimpse(datanavn)
```

Variabler og type

```
summary(datanavn)
```

Gjennomsnitt, osv

```
names(datanavn)
```

Variabelnavn



Øvelse 2

- Øvelsene til oppgavesett 1 er på fil: **øvelser_2.R**

