

R grunnkurs

ASLAUG HURLEN FOSS, ANITA CECILIE DAL

OG SAMER ALTINAWI



Statistisk sentralbyrå
Statistics Norway

Presentasjon av kursholdere



Aslaug Hurlen Foss



Anita Cecilie Dal



Samer Altinawi



Statistisk sentralbyrå
Statistics Norway

Målet med kurset

- Bli kjent programmet R
- Bli litt kjent med Jupiter og hente kode fra Github
- Lese inn data
- Behandler data
- Lage tabeller og oppsummere
- Lage figurer



Agenda

- 10.00-10.20 Introduksjon til R
- 10.20-10.40 Øvelse: Jupiter og Github
- 10.40-11.00 Tall, tekst, liste og datasett
- 11.00-11.30 Øvelse med objekter
- 11.30-12.00 Lunsj i kantina
- 12.00-12.30 Bearbeiding av data
- 12.30-13.00 Øvelser
- 13.00-13.30 Koble og grafikk
- 13.30-14.00 Øvelse



Mitt møte med SAS - kultursjokk

Datasett

Log

Tekst, tekst, tekst

Objekter

Verdier

--

Vektor-lister

Matriser



Statistisk sentralbyrå
Statistics Norway

Hvorfor R

- Godt språk for nybegynnere og forskere
- Utviklet for statistikk
- God grafikk
- Basis + «pakker» med funksjoner
- Egen konferanse for r i statistikkproduksjon!

The R Project - The Use of R in Official Statistics - uRos2022



USE OF R IN OFFICIAL STATISTICS 2022
10TH INTERNATIONAL
VIRTUAL CONFERENCE

Hosted by Statistics Canada
December 6-8, 2022



Statistisk sentralbyrå
Statistics Norway

awesome-official-statistics-software



Italia
Estimering



Frankrike
Sesongjustering



Nederland
Dataeditering



Norge
Konfidensialitet



Internasjonalt: awesome-official-statistics-software

- Liste med programvare for offisiell statistikk
- Kategorisert etter prosessmodellen GSBPM
- Listen blir vedlikeholdt av Nederland
- Ukentlig sjekk av oppdatering
- Internasjonale retningslinjer
- Gjenbruk er bra!

Design frame and sample (GSBPM 2.1)

- CRAN 1.5-4 – a month ago license GPL (≥ 2)

R package [SamplingStrata](#). Optimal Stratification of Sampling Frames for Multipurpose Sampling Surveys.

- CRAN 1.0.4 – a year ago license EUPL

R package [R2BEAT](#). Multistage Sampling Allocation and PSU Selection.

Design variable descriptions (GSBPM 2.2)

- GitLab no releases found last commit december license MIT License

Excel [SDMX_Matrix_Generator](#). Excel-based visual SDMX artefact authoring tool which generates SDMX-ML for upload into an SDMX repository such as a registry. By OECD.

Sampling (GSBPM 4.1)

- CRAN 2.9 – 2 years ago license GPL (≥ 2)

R package [sampling](#). Several algorithms for drawing survey samples, including a variety of unequal probability sampling designs (high entropy, systematic, Rao-Sampford, etc.), and calibrating design weights.

- CRAN 4.0 – 3 years ago license GPL (≥ 2)

R package [surveyplanning](#). Tools for sample survey planning, including sample size calculation, estimation of expected precision for the estimates of totals, and calculation of optimal sample size allocation.

- CRAN 1.2.8 – 4 months ago license GPL (≥ 2)

R package [PracTools](#). Functions and datasets related to Valliant, Dever, and Kreuter (2018 2nd ed), [Practical Tools for Designing and Weighting Survey Samples](#).



Miljø for å kjøre R

- Jupiter – godt egnet for enkle ting
- Rstudio – den «vanlige» verktøyet for R.
- Muligheter i Rstudio versus Jupiter
 - Lett å se på data og objekter, pakker og figurer i Rstudio
 - Lett å bygge enkle dashboard og analyser i pdf
 - Bygge pakker som er viktig funksjon for metode og forskningsavdelingen



R i SSB

- R-produktteam
 - Testet workbench – ble for dyrt
 - Rstudio på ny R-server – jobbes med nå
 - Rstudio i Sky ?
 - Felles funksjoner i R

R - ressurser og kompetansebygging

På dette området finner du



Kom i gang med R



Ofte stilte spørsmål



Kurs i R

→ ↻ 🏠 github.com/statisticsnorway/fellesr

📄 fellesr-package.r
Oppdaterer dokumentasjon
last month

📄 fellesr.Rproj
Added using Roxygen
4 months ago

☰ README.md

Fellesfunksjoner i R for SSB (`fellesr`)

Her finner du en rekke funksjoner laget for å løse problemstillinger som er vanlige i SSB. Se under «Articles» [her](#) eller lenkene nedenfor for å velge emner.

Pakken `fellesr` inneholder følgende funksjoner:

Produksjonssonen

- [Uttrekk fra Dynarev i R](#) (`dynarev_uttrekk`)
- [Laste opp data til Statistikkbanken](#) (`statbank_lasting`)
- [R pakkehåndtering med renv](#) (`create_dependencies` , `install.packages` , `restore` , `snapshot` , `ssb_library`)

DAPLA

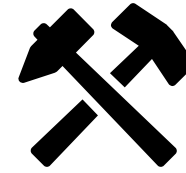
- [Lese inn og skrive filer på DAPLA](#) (`read_SSB` , `write_SSB` , `gcs_bucket` , `list.files`)
- [R pakkehåndtering med renv](#) (`create_dependencies` , `install.packages` , `restore` , `snapshot` , `ssb_library`)



Statistisk sentralbyrå
Statistics Norway

Metodebibliotek i SSB

- Metodebiblioteket er SSBs bibliotek for statistiske metoder
- Alle funksjoner er og testet i metodeseksjonen og er i dag i bruk i minst én produksjonsprosess.



Under bygging

← → ↺ 🏠 statisticsnorway.github.io/metodebibliotek/catalog.html

Metodebiblioteket Browse Catalogue ▾

Metodebiblioteket functions

Feb. 28, 2022
sklearn.svm

LinearSVC

PYTHON IMPUTATION CLASSIFICATION

Linear Support Vector Classification. Similar to SVC with parameter kernel='linear', but implemented in terms of liblinear rather than libsvm, so it has more flexibility in the choice of penalties and loss functions and should scale better to large numbers of samples.

CATEGORIES

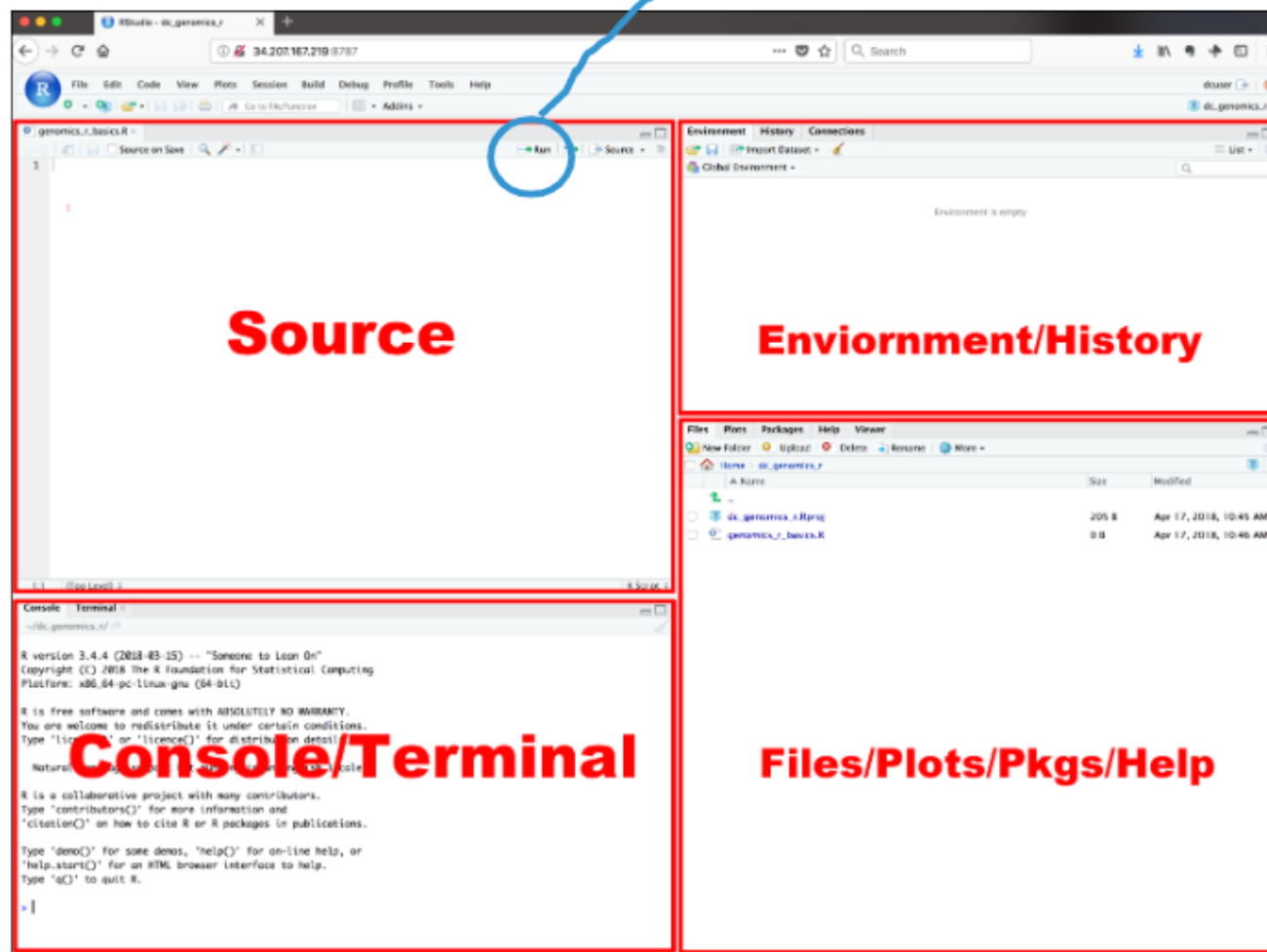
- Articles (7)
- classification (2)
- confidentiality (4)
- dataediting (2)
- imputation (2)
- python (1)
- R (6)



Statistisk sentralbyrå
Statistics Norway



Kjøre-knappen
(eller ctrl + enter)



Statistisk sentralbyrå
Statistics Norway

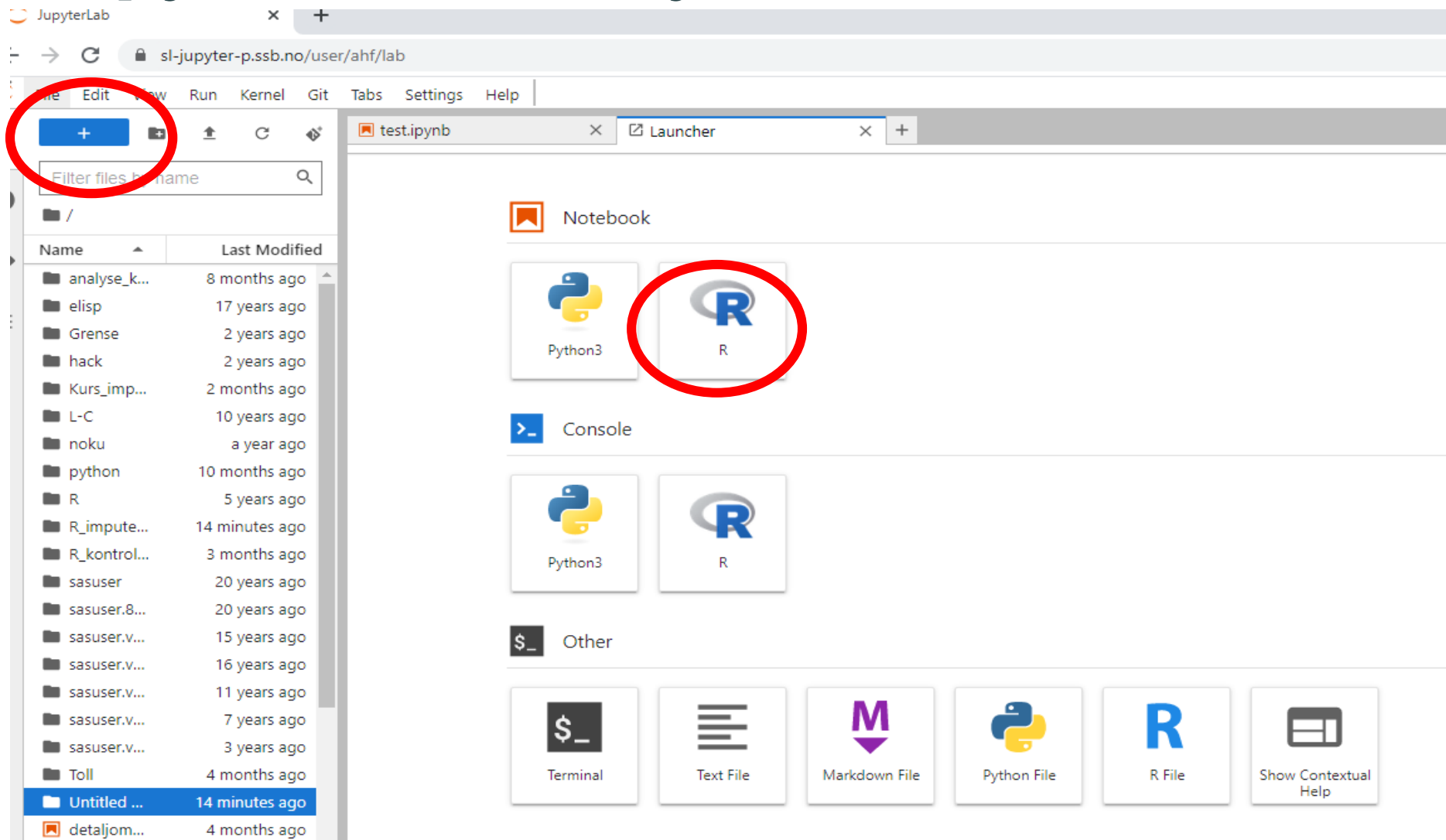
Jupyter – Julia, Python, R

The image shows a screenshot of the JupyterLab web interface. The interface includes a file browser on the left, a central code editor, and a right-hand sidebar. Several green callout bubbles are overlaid on the interface to explain its components:

- Fil-utforsker**: Points to the file browser on the left side of the interface.
- Hvilket nr cellen ble kjørt 1,2,3,**: Points to the cell execution order indicator in the code editor.
- Kjøre koden i cellen**: Points to the 'Run' button (a green play icon) in the code editor.
- Kode eller tekst**: Points to the dropdown menu in the code editor that allows switching between 'Code', 'Markdown', and 'Raw' modes.
- Celler**: Points to the code cell in the editor.

The interface also shows a terminal at the bottom with the text 'Fully initialized R | Idle Mem: 1.15 / 10.00 GB' and a status bar at the bottom right indicating 'Mode: Command' and 'Ln 1, Col 1'.

Jupyter – Julia, Python, R



The screenshot displays the JupyterLab web interface. The top navigation bar includes tabs for 'File', 'Edit', 'View', 'Run', 'Kernel', 'Git', 'Tabs', 'Settings', and 'Help'. Below this is a toolbar with a red circle highlighting the '+' button for creating new files. The left sidebar shows a file browser with a search bar and a list of files and folders, including 'analyse_k...', 'elisp', 'Grense', 'hack', 'Kurs_imp...', 'L-C', 'noku', 'python', 'R', 'R_impute...', 'R_kontrol...', 'sasuser', 'sasuser.8...', 'sasuser.v...', 'sasuser.v...', 'sasuser.v...', 'sasuser.v...', 'sasuser.v...', 'Toll', 'Untitled ...', and 'detaljom...'. The main workspace area is divided into three sections: 'Notebook', 'Console', and 'Other'. The 'Notebook' section shows two options: 'Python3' and 'R', with the 'R' option highlighted by a red circle. The 'Console' section shows 'Python3' and 'R' options. The 'Other' section includes 'Terminal', 'Text File', 'Markdown File', 'Python File', 'R File', and 'Show Contextual Help'.

Name	Last Modified
analyse_k...	8 months ago
elisp	17 years ago
Grense	2 years ago
hack	2 years ago
Kurs_imp...	2 months ago
L-C	10 years ago
noku	a year ago
python	10 months ago
R	5 years ago
R_impute...	14 minutes ago
R_kontrol...	3 months ago
sasuser	20 years ago
sasuser.8...	20 years ago
sasuser.v...	15 years ago
sasuser.v...	16 years ago
sasuser.v...	11 years ago
sasuser.v...	7 years ago
sasuser.v...	3 years ago
Toll	4 months ago
Untitled ...	14 minutes ago
detaljom...	4 months ago

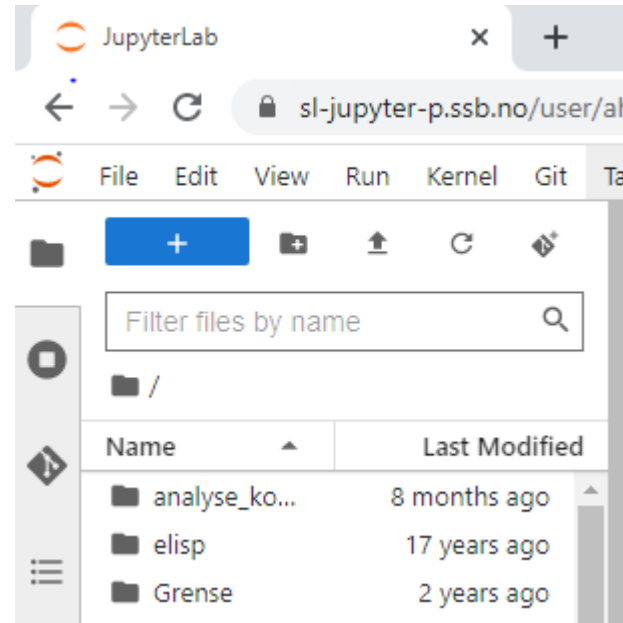
Jupyter – Julia, Python, R

Filutforsker

Terminal og
kjerne

Git

Innhold



Statistisk sentralbyrå
Statistics Norway

Hva er markdown?

- Markdown er en enkel syntaks som formaterer tekst som overskrifter, lister, fet skrift og så videre.
- Mest brukte kommandoer:
 - # - tittel (## - overskrift)
 - **tekst** - tekst i fet skrift
 - *tekst* - tekst i kursiv skrift
 - \$ x=by +e \$ - matematiske formler



Eksempel

```
# Tittel
```

```
## Overskrift
```

```
Dette er min tekst med formler $ y=bx+e $ og bruk av kursiv
```

Tittel

Overskrift

Dette er **min** tekst med formler $y = bx + e$ og bruk av *kursiv*

Analyse av statistiske kontroller i Tvinn

Aslaug Hurlen Foss

29 3 2022

Innledning

Som input kontroll har Tvinn registrert to statistiske kontrollfunksjoner. Relativ priskontroll og generell priskontroll. Parametrene til disse funksjonene blir levert av SSB.

Data

Grunnlaget for analysen er alle varer i mengde og verdi fortollet i 2019. Datasettet inneholder omtrent 22 millioner observasjoner.

Variabler i kontrollene: * verdi * mengde

Grupperingsvariabler for analysen:

- Eksport/import
- Mengde og annen mengde – hvilken variabel?
- Varegrupper - det kan bli brukt de to første sifre

Definisjon av rettet og treffsikkerhet

Antall rettet av de som er slått ut i denne kontrollen:

- Rettet 1: Behandlet i klargjøringsprosessen i SSB med kodene: Retting, Omberegning og Slettet
- Rettet 2: Endret verdi i mengde, verdi eller varenummer. Eller at varen har blitt slettet.

Det kan altså være at det blir rettet andre variabler som landkode og ikke verdi, mengde eller varenummer. Kontrollene som skal analyseres her er forholdet mellom verdi og mengde i de forskjellige varegruppene.

Treffsikkerhet er antall rettede verdier delt på antall vareposter som er slått ut i en kontroll, oppgitt i prosent.



Statistisk sentralbyrå
Statistics Norway

Kursmaterialet

<https://github.com/statisticsnorway/kurs-r-grunnkurs-jupyter>

← → ↻ 🏠 [github.com/statisticsnorway/R_kontrollfunksjoner](#)

Search or jump to... / Pull requests Issues Marketplace Explore

[statisticsnorway](#) / [R_kontrollfunksjoner](#) Public Edit Pins

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

[main](#) 2 branches 0 tags [Go to file](#) [Add file](#) [Code](#)

Your main branch isn't protected
Protect this branch from force pushing, deletion, or require status checks before merging

aslaugfoss Add files via upload

Eksempler.R	Add files via upload
Losninger.R	Add files via upload
Oppgaver.R	Add files via upload
Presentasjon.pdf	Add files via upload

Clone ⓘ

HTTPS SSH GitHub CLI

https://github.com/statisticsnorway/R_kontrollfunksjoner

Use Git or checkout with SVN using the web URL.

Open with GitHub Desktop

Download ZIP

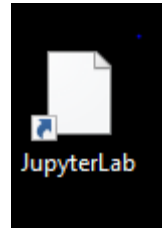
1 / months ago



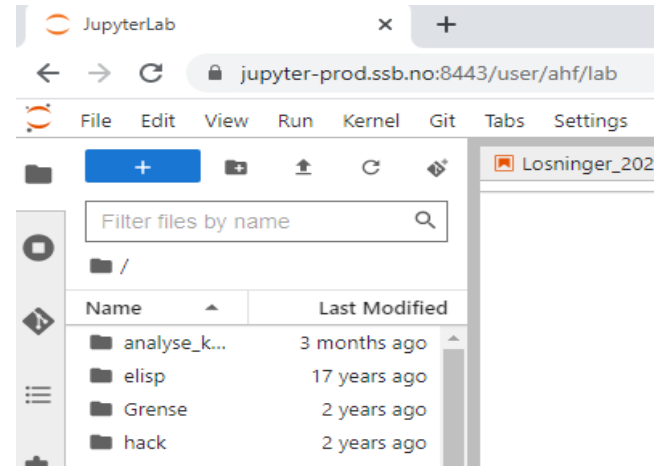
Statistisk sentralbyrå
Statistics Norway

Starte opp Jupyter i produksjonssonen

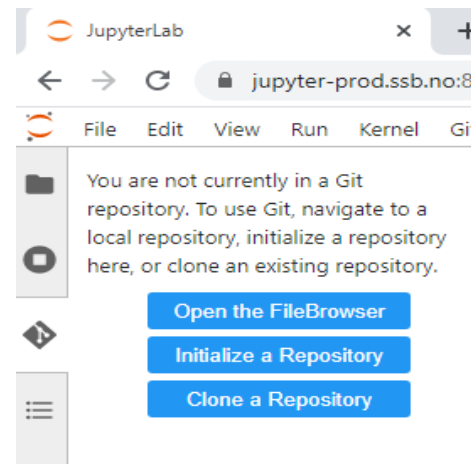
- Trykk på ikonet:



- Stå i «filutforsker»



- Trykk på Github-ikonet:



Clone a repo

Enter the Clone URI of the repository

Cancel

CLONE

Øvelse

- Gå til kursmaterialet på Github og kopier adressen
(<https://github.com/statisticsnorway/kurs-r-grunnkurs-jupyter>)
- Starte Jupyter i produksjonssonen
- Klon kurset over til deg selv
- Lage en ny Notebook i R
- Lage en tekst i Markdown med overskrift og fet skrift. Lag navn på notebooken.



Tall, tekst, liste og datasett

R som kalkulator

- Regnestykkene kan tastes rett inn:

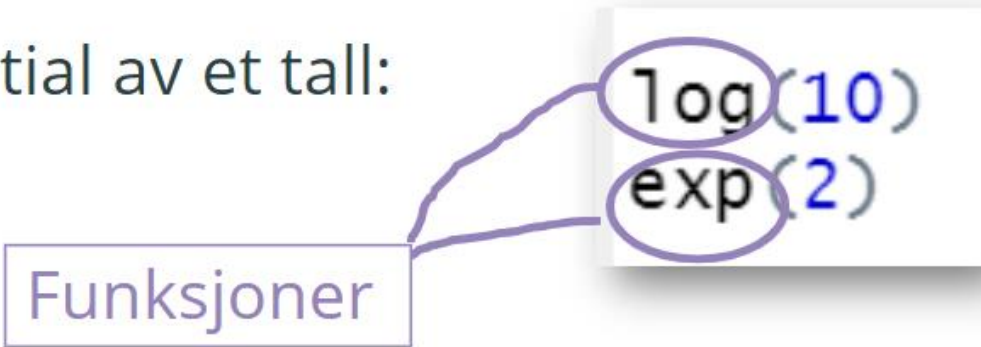
$2 + 2$
 $5 - 2$
 $25 * 4$
 $5 \wedge 2$



Grunnfunksjoner

Log og exponential av et tall:

Funksjoner



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

Alle funksjoner i R har hjelpfiler:

```
?log
```



Tekststrenger - streng

'Nittedal'

- Både enkel (' ') og dobbel snutter (" ") kan brukes –konsistent!

- CASE SENSITIVE

Kommune ≠ kommune

- Lim sammen med paste()

paste()

- Plukk ut deler med substring()

substring(steng, start, stop)



Objekter

- Et objekt inneholder data
- Å lage et objekt er å gi det et navn

```
Kommune <- 'Nittedal'
```

- Vi bruker <- for å allokere noe til et objekt
- Skriv ut hva innhold er ved å skrive objektnavn

```
Kommune  
> 'Nittedal'
```



Objekter

- CASE SENSITIVE

Kommune ≠ kommune

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

Kommune <- 'Kongsvinger'



Objekttyper

- char – karakter, tekststreng
- num – numerisk verdi
- date – dato
- factor – kategorisk bestemt nivå
- data.fram - datasett



Liste - vektorer

- Objekter inneholder flere verdier
- Bruk `c()` for å kombinere verdier og komma mellom
- Må være samme objekttype

```
c( 'Kongsvinger', 'Oslo')  
c(1,3,6)
```



Vektorer som objekt

- Gir vektorer et navn ved

`<-`

- Finne lengde av vektoren ved funksjonen

`length()`

- For å finne ut objekttypen bruk

`str()`



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(Kommune== 'Kongsvinger', '3401', '0301')
```

Betingelse

Sann

Usann



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 navnet på datasettet – datasettet blir printet ut
- Se på de 5 første linjene av datasettet: `head()`



Pakker og biblioteker

- Pakker er en «modul» – samling av funksjoner
- Cran – bibliotek med samling av pakker
- Hvordan bruke pakker:
 - Installering av pakker, en gang: `install.packages('pakkenavn')`
 - For hver gang koden kjører: `library(pakkenavn)`



Les inn datasett: .csv fil

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

Sti til datasett

Skilletegn = , Desimaltegn = .

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

Sti til datasett

Skilletegn = ; Desimaltegn = ,



Les inn sas fil

- VI kan lese inn sas datasett (.sas7bdat) direkte til R med haven pakken

```
library(haven)
```

```
read_sas('sti til datasettet')
```

```
?read_sas
```



Lese inn data fra statistikkbanken

- Bruk pakken PxWebApiData
- Trenger tabell nummer

```
library(PxWebApiData)
helse_bank <- ApiData(12293, KOKkommuneregion0000 = TRUE , ContentsCode =TRUE,Tid = TRUE)
dat<-helse_bank[[2]]
meta<-helse_bank[[1]]
```

- E-læringskurs: https://statisticsnorway.shinyapps.io/ApiData_tutorial/



Lese inn data fra Klass

– database for klassifiseringer og kodelister

- Bruk pakken klassR
- Trenger id nummer til klassifiseringen – adressen i URL

```
#Henter gyldig fylkesliste fra Klass
library(klassR)
fy_klass <- GetKlass(klass = 214, date = "2022-01-01")
```

- E-læringskurs: https://statisticsnorway.shinyapps.io/KlassR_tutorial/



Ser nærmere på data

```
glimpse(datanavn)
```

Variabler og type

```
summary(datanavn)
```

Gjennomsnitt, osv

```
names(datanavn)
```

Variabelnavn



Eksempler kjøring i R



Øvelse

- Oppgaver 1- 12 på filen: Oppgaver.ipynb
- Løsninger på oppgavene fins på filen: Losninger.ipynb
- Hvis det er for mange oppgaver, plukk ut det du synes er mest interessant for deg!



Bearbeiding av datasett

Bearbeiding med pakken «tidyverse»

- Gjør koden ryddigere – effektiv og lesbar
- Pipelines «rørledning» - med `%>%` som pipe operator

`%>%` - Denne operatøren vil videresende en verdi, eller resultatet av et uttrykk, til neste uttrykk.

Base R:

```
leave_house(get_dressed(get_out_of_bed(wake_up(me))))
```

tidyverse:

```
me %>%  
  wake_up() %>%  
  get_out_of_bed() %>%  
  get_dressed() %>%  
  leave_house()
```

Lage nye variabler: mutate()

- Kan brukes som en del av en pipeline

```
datanavn %>%  
  mutate(nyvariabel = 1000)
```

Gir variabel et navn

```
datanavn %>%  
  mutate(nyvariabel = oldvariabel * 1000)
```

Gir variabel et navn

Eksisterende variabel

+ hva skal gjøres



Statistisk sentralbyrå
Statistics Norway

Lage nye variabler: mutate()

- Kombinere med `ifelse()`
- For å lagre husk `<-`
- Kan overskrives
- Flere variablene kan lagres samtidig (, for å skille)
- Endre variabeltype (`as.character()`, `as.numeric()`)

```
datanavn %>%  
  mutate(variabelnavn = as.character(variabelnavn))
```



Velg noen rader: filter()

- For å velge ut noen rader bruker vi filter()
- Skriv logiske setning inn i parentes.
- Flere logiske setninger kan brukes sammen (skille med ,)

```
datanavn %>%  
  filter(condition)
```

- Igjen: Ingenting lagres uten <-



Velg ut noen variabler: select()

- Brukes med pipelines
- Skriv variabelnavn i parentes
- En eller flere variabler (skille med ,)
- Brukes sammen med andre funksjoner (for eks. filter())

```
datanavn %>%  
  select(variabelnavn)
```

```
datanavn %>%  
  filter(condition) %>%  
  select(variabelnavn)
```

Oppsummering/aggregering: summarise()

- Ta oppsummering (summen, gjennomsnitt, median, antall) av en variabel med summarise()

```
datanavn %>%  
  summarise(oppsummeringsnavn = mean(variabelnavn))
```

Gir oppsummering et navn

Hva skal gjøres:

- mean()
- median()
- sum()
- n()

Eksisterende variabel



Statistisk sentralbyrå
Statistics Norway



Gruppering: group_by()

- Gjøre alle prosesser etterpå innen hver gruppe

```
datanavn %>%  
  group_by(grupperingsvariabel) %>%  
  summarise(oppsummeringsnavn = mean(variabelnavn))
```



Gruppering: group_by() og spread()

- Kombinere flere variabler med ,
- For en 2 x 2 frekvenstabell:

```
datanavn %>%  
  group_by(grupperingsvariabel1, grupperingsvaraibel2) %>%  
  summarise(oppsummeringsnavn = n()) %>%  
  spread(grupperingsvariabel1, oppsummeringsnavn)
```



Endre variabelnavn: rename()

```
datanavn %>%  
  rename(nyttnavn = gammeltnavn)
```

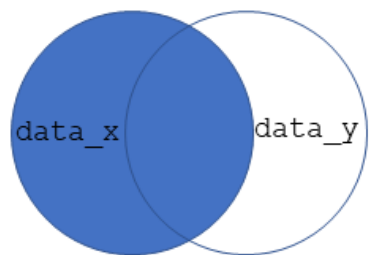
Øvelser 3

- Oppgave 13 til 18

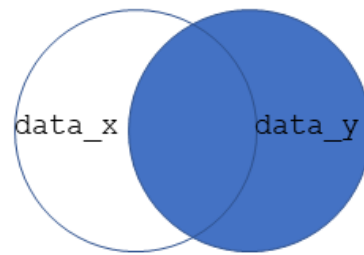


Koble og figurer

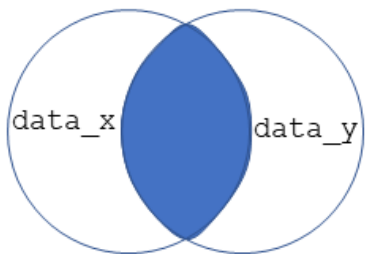
Koble to datasett



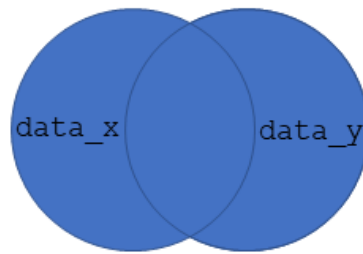
`left_join(data_x, data_y)`



`right_join(data_x, data_y)`



`inner_join(data_x, data_y)`



`full_join(data_x, data_y)`



Koble to datasett

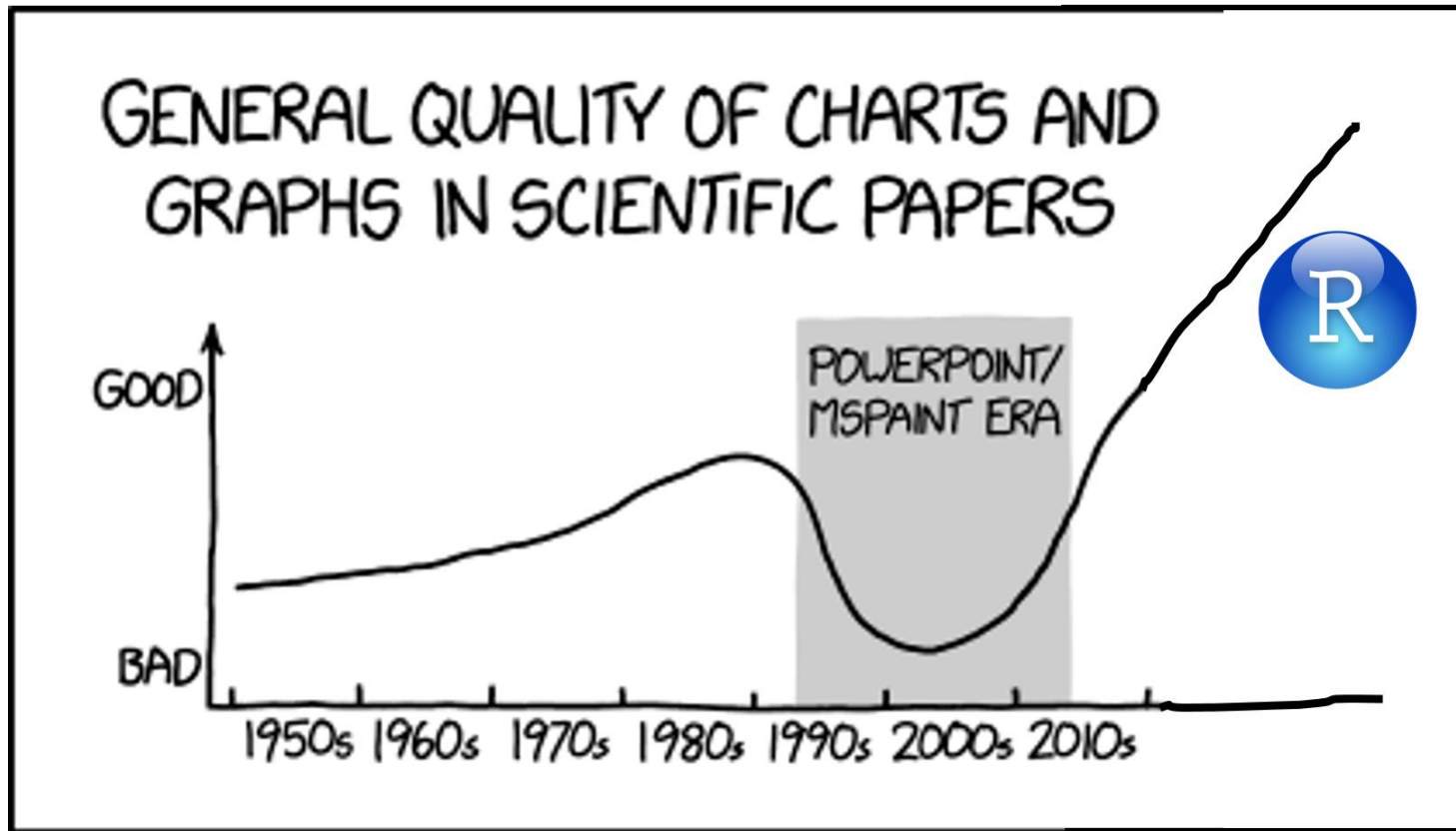
- Bruk **by** = for å spesifisere nøkkel variabel til å koble på

```
by = c("aar" = "year")
```

```
kobletdata <- left_join(datanavn1, datanavn2, by = variabelnavn)
```

- Flere variabler kan brukes for å koble på (som en vektor)

Plotting



Plotting med ggplot()

- **aes** : aesthetics, hvilke variabler
- **geom_** : hva slags figur
- **stat** : hva slags statistisk aggregat å presentere

Table 18-1 A Selection of Geoms and Associated Default Stats

<i>Geom</i>	<i>Description</i>	<i>Default Stat</i>
<code>geom_bar()</code>	Bar chart	<code>stat_bin()</code>
<code>geom_point()</code>	Scatterplot	<code>stat_identity()</code>
<code>geom_line()</code>	Line diagram, connecting observations in order by x-value	<code>stat_identity()</code>
<code>geom_boxplot</code>	Box-and-whisker plot	<code>stat_boxplot()</code>
<code>geom_path</code>	Line diagram, connecting observations in original order	<code>stat_identity()</code>
<code>geom_smooth</code>	Add a smoothed conditioned mean	<code>stat_smooth()</code>
<code>geom_histogram</code>	An alias for <code>geom_bar()</code> and <code>stat_bin()</code>	<code>stat_bin()</code>



Søylediagram

```
ggplot(aes(variabelnavn)) +  
  geom_bar()
```

Bruke + for å legge til figurtype

Spesifisere variabelen

Spesifisere søylediagram

```
ggplot(aes=c(x = variabelnavn1, y = variabelnavn2)) +  
  geom_bar(stat="identity")
```

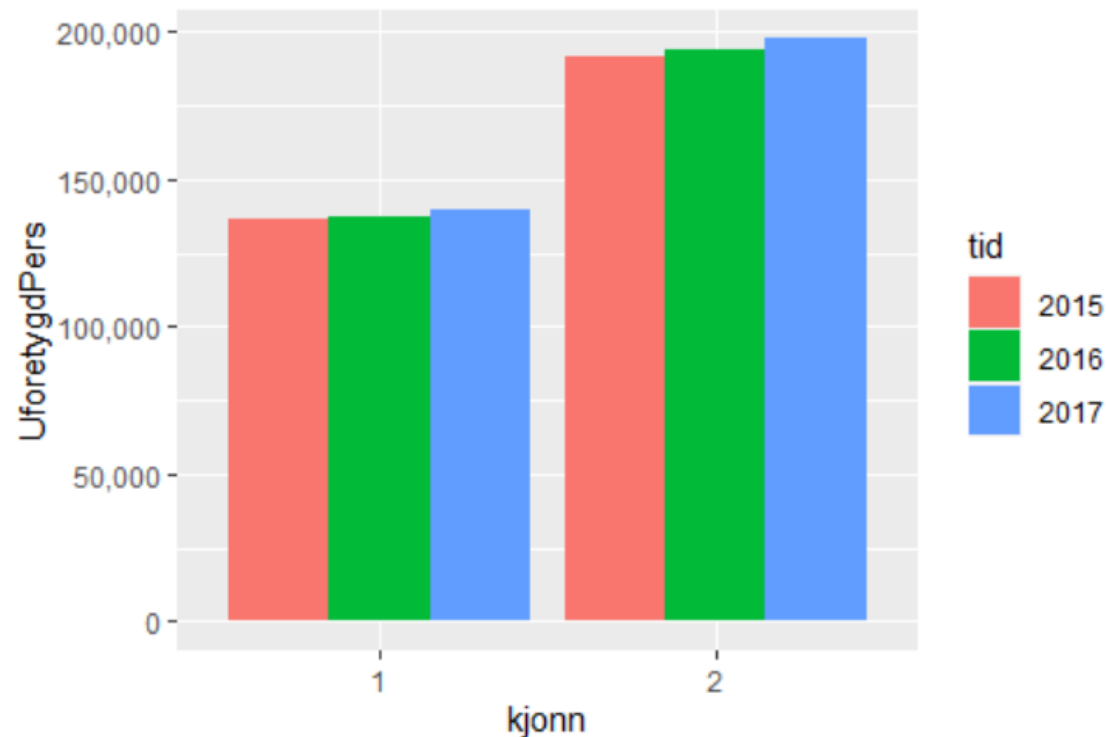
Spesifisere x og y variablene

Spesifisere å bruke verdi



Søylediagram

- Bruk fill() i aes for å spesifisere en variabel for farge
- Kombinere med filter først



Punktdiagram

- Sammenlign to numeriske variabler

```
ggplot(aes(x = variabelnavn1, y = variabelnavn2)) +  
geom_point()
```

- Legg til regresjonslinje med

```
geom_smooth(method = "lm")
```

- Farge punktene etter gruppe

```
geom_point(aes(color = variabelnavn))
```

Drag-and-drop ggplot med pakken esquisse



- Fungerer ikke på Jupyter!

Plotly – interaktive figurer

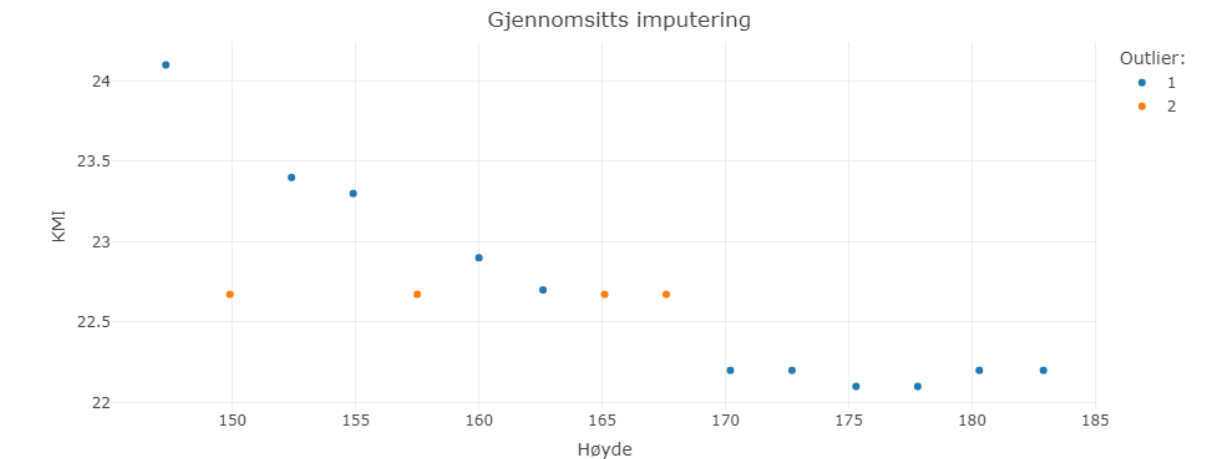
- Funksjonalitet:
 - hover – få informasjon når markøren holdes over et punkt i grafen
 - Zoome
 - Selektare
- Kan omgjøre plot i ggplot til plotly



Interaktive figurer - plotly

- Funksjonalitet:
 - hover – få informasjon når markøren holdes over et punkt i grafen
 - Zoome
 - Selekttere
- Kan omgjøre plot i ggplot til plotly
 - library(plotly)
 - fig1<- ggplot(aes(x=variabel1, y=variabel2)) + geom_point()
 - ggplotly(fig1)

```
fig <- women_1 %>%  
  plot_ly(x = ~hoyde, y = ~kmi,  
    type = "scatter", mode = "markers", split = ~imp,.....  
    hovertemplate = paste(.....  
      "%{axis.title.text}: %{y:}<br>",  
      "%{axis.title.text}: %{x:}<br>",  
      "<extra></extra>"  
    )..) %>%  
  layout(title = "Gjennomsitts imputering", xaxis = list(title = "Høyde"),  
    yaxis = list(title = "KMI").., legend=list(title=list(text='Outlier:')))
```



Øvelse 4

- Oppgaver 19- 27

R- videregående

- Kommer i mars med Susie
 - Flipped classroom
- **Målsetning:** Å gjøre medarbeiderne i stand til å jobbe med mer kompleks oppgaver i R og strukturere kode ved bruk av funksjoner.
 - **Innhold:**
 - Datasett, variabeltype og indeksering
 - Loops
 - Funksjoner
 - Kjøring av R og python sammen



Oppsummering

- Husk library()
- Les inn filer: read_csv2() read_sas()
- Ny variabel: mutate()
- Velg noen linje: filter()
- Aggregere/oppsummere: summarise()
- Figur: ggplot(), aes(), geom_...()

- <https://ssbno.sharepoint.com/sites/Kompetanseogutvikling/SitePages/Ressursside-om-R.aspx>
- Yammer: R i SSB
- Google



Takk!

