TALL | SOM FORTELLER

Introduction course

DAY 5 – 26TH JANUARY 2021

SUSIE JENTOFT & ASLAUG HURLEN FOSS

Statistisk sentralbyrå
Statistics Norway

Agenda

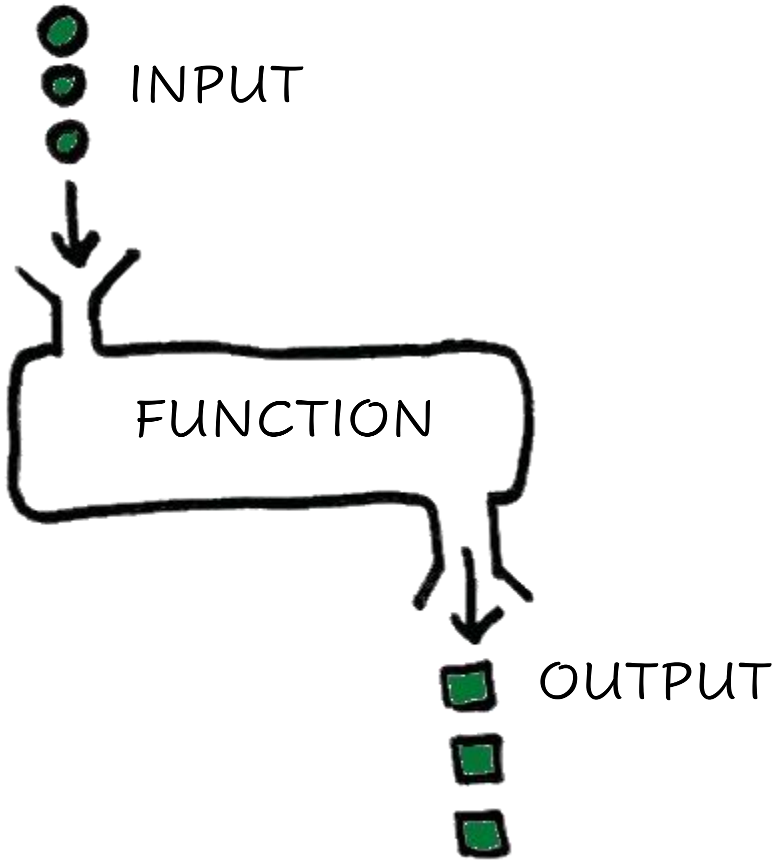
	Monday 25th January	Tuesday 26th January	Wednesday 27th January
12:00	• General: indexing, vectors and lists	• Summary • Functions	• Building packages
12:45	Exercise 7	Exercise 9	Exercise 11
13:15	Break	Break	Break
13:30	• Sorting • Control with if and else • Loops	• RMarkdown • Dashboards	• Other useful packages and resources
14:15 – 15:00	Exercise 8	Exercise 10	Discusion and summary

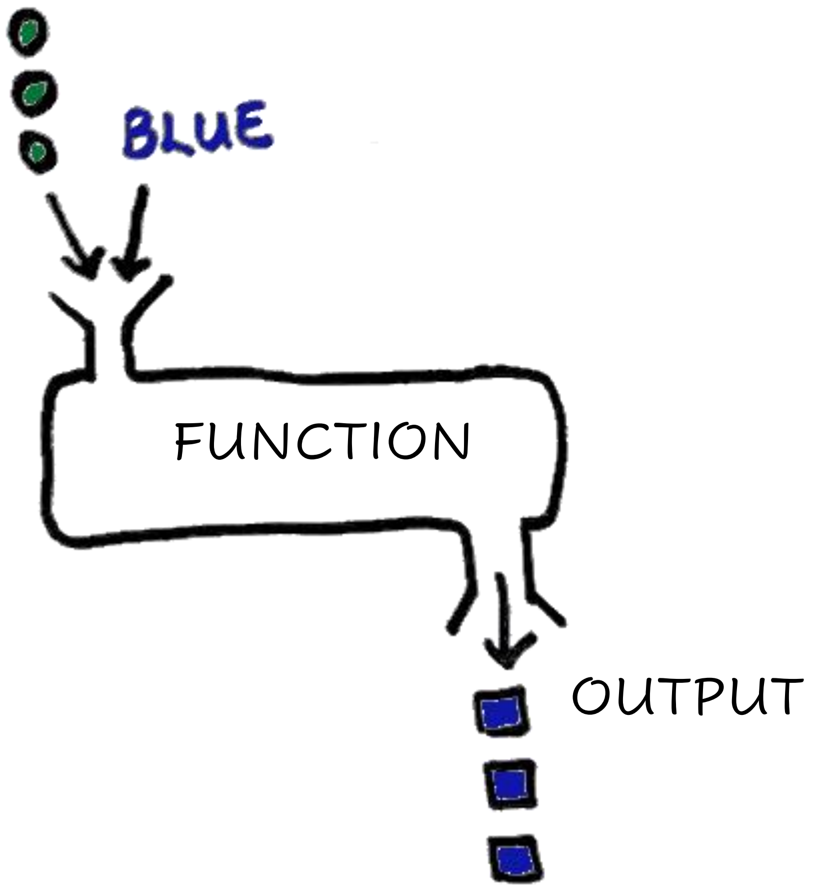
Statistisk sentralbyrå
Statistics Norway

Review of exercise 8

Statistisk sentralbyrå
Statistics Norway

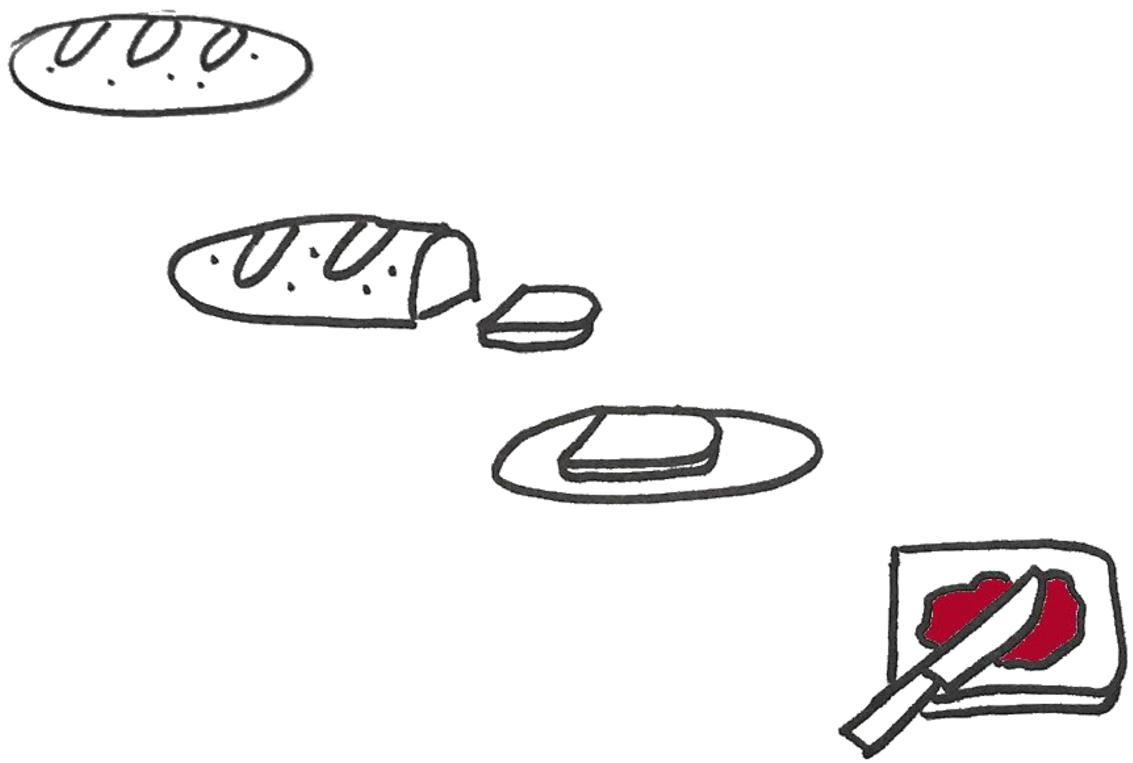
What is a function?



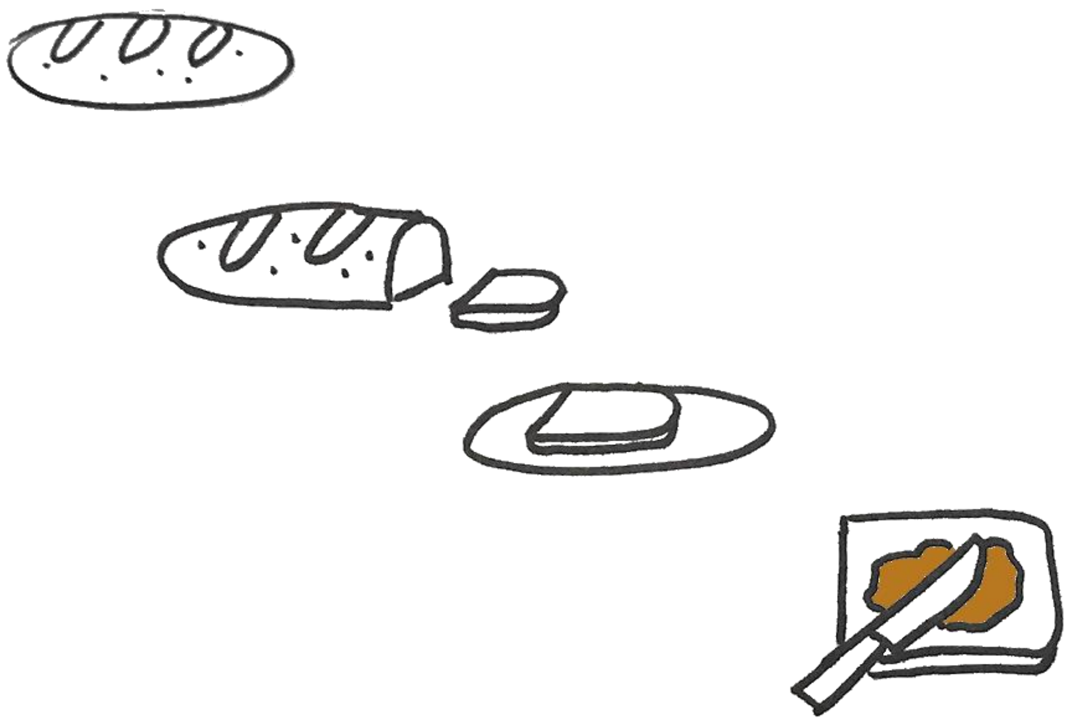


Why do we use functions?

Make_sandwich_with: jam



Make_sandwich_with: peanutbutter



Functions i R

`c()`

`data.frame()`

`head()`

`log()`

`list()`

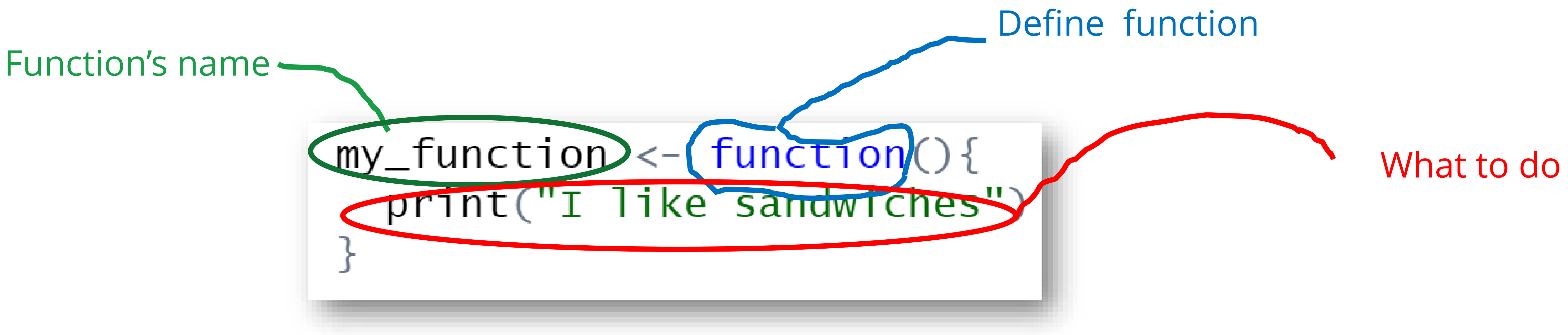
`exp()`

`read_csv()`

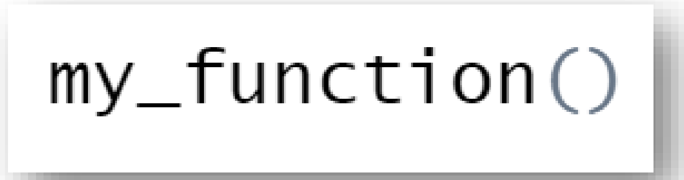
`seq()`

`glimpse()`

Make a function in R



```r
my_function <- function(){
  print("I like sandwiches")
}
```

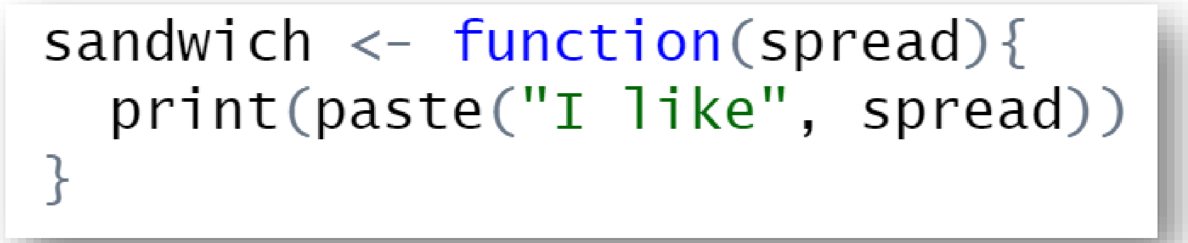


```r
my_function()
```

Statistisk sentralbyrå
Statistics Norway

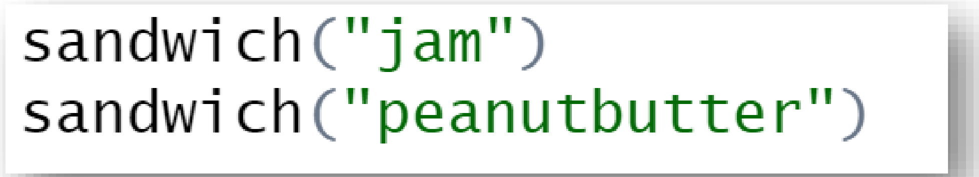
Functions with a parameter

- Parameter is written in brackets:



```r
sandwich <- function(spread){
  print(paste("I like", spread))
}
```

- Write parameter value when calling function

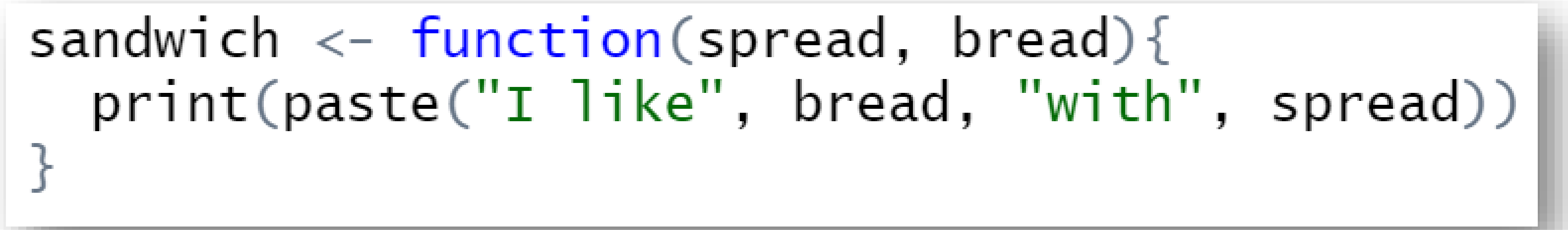


```r
sandwich("jam")
sandwich("peanutbutter")
```

```
"I like jam"
"I like peanutbutter"
```

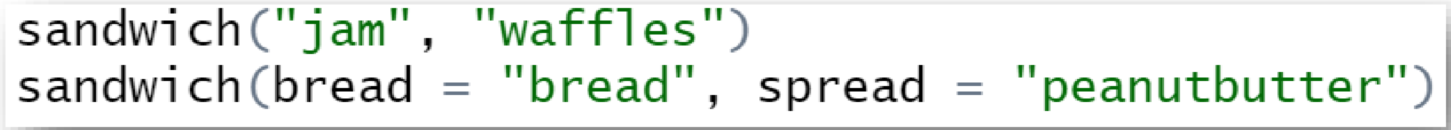
Functions with several paramters

- Include several parameters



```r
sandwich <- function(spread, bread){
  print(paste("I like", bread, "with", spread))
}
```

- Use same order or specify



```r
sandwich("jam", "waffles")
sandwich(bread = "bread", spread = "peanutbutter")
```

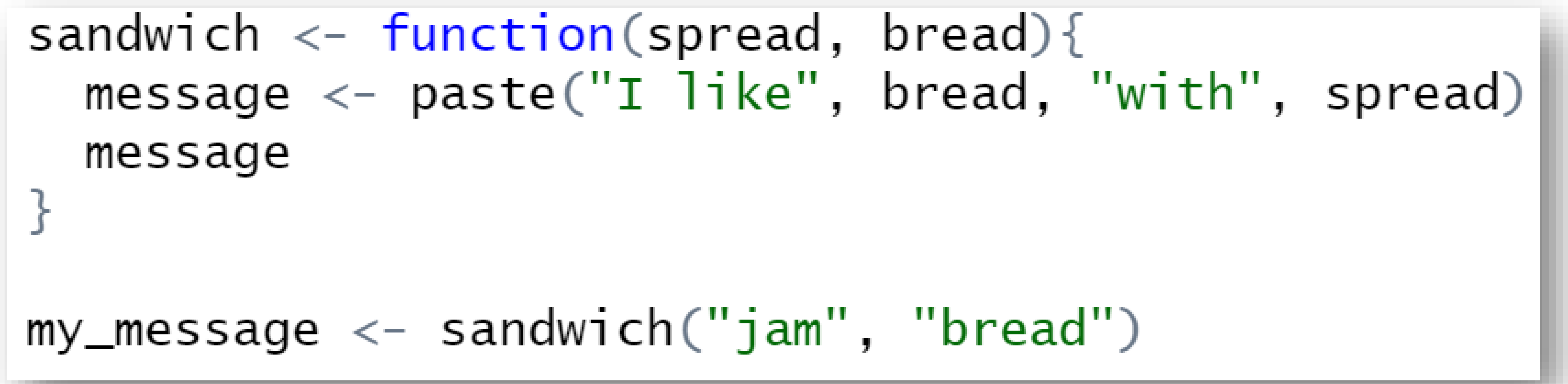
```
"I like waffles with jam"
"I like bread with peanutbutter"
```

- Parameter type: number, string, logical/boolean, vector, list, dataset

Statistisk sentralbyrå
Statistics Norway

Function output

- Default: last line
- Alternative: return()
- Can be a string, number, vector, list, dataset (but only one object)

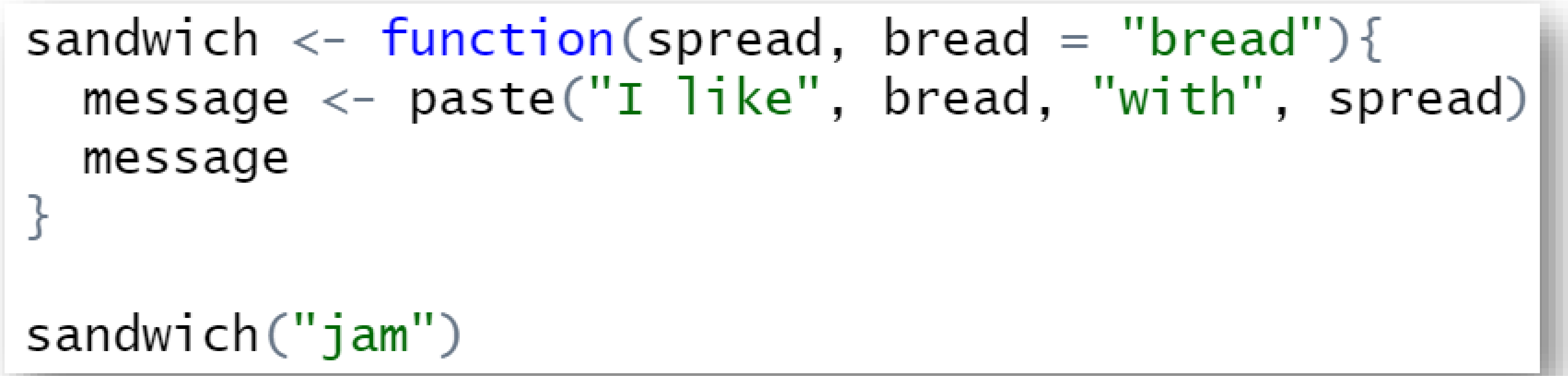


```r
sandwich <- function(spread, bread){
  message <- paste("I like", bread, "with", spread)
  message
}

my_message <- sandwich("jam", "bread")
```

Statistisk sentralbyrå
Statistics Norway

Default parameter

- Set a default parameter value so the user doesn’t have to specify
- Default parameter comes after the other parameter



```r
sandwich <- function(spread, bread = "bread"){
  message <- paste("I like", bread, "with", spread)
  message
}

sandwich("jam")
```

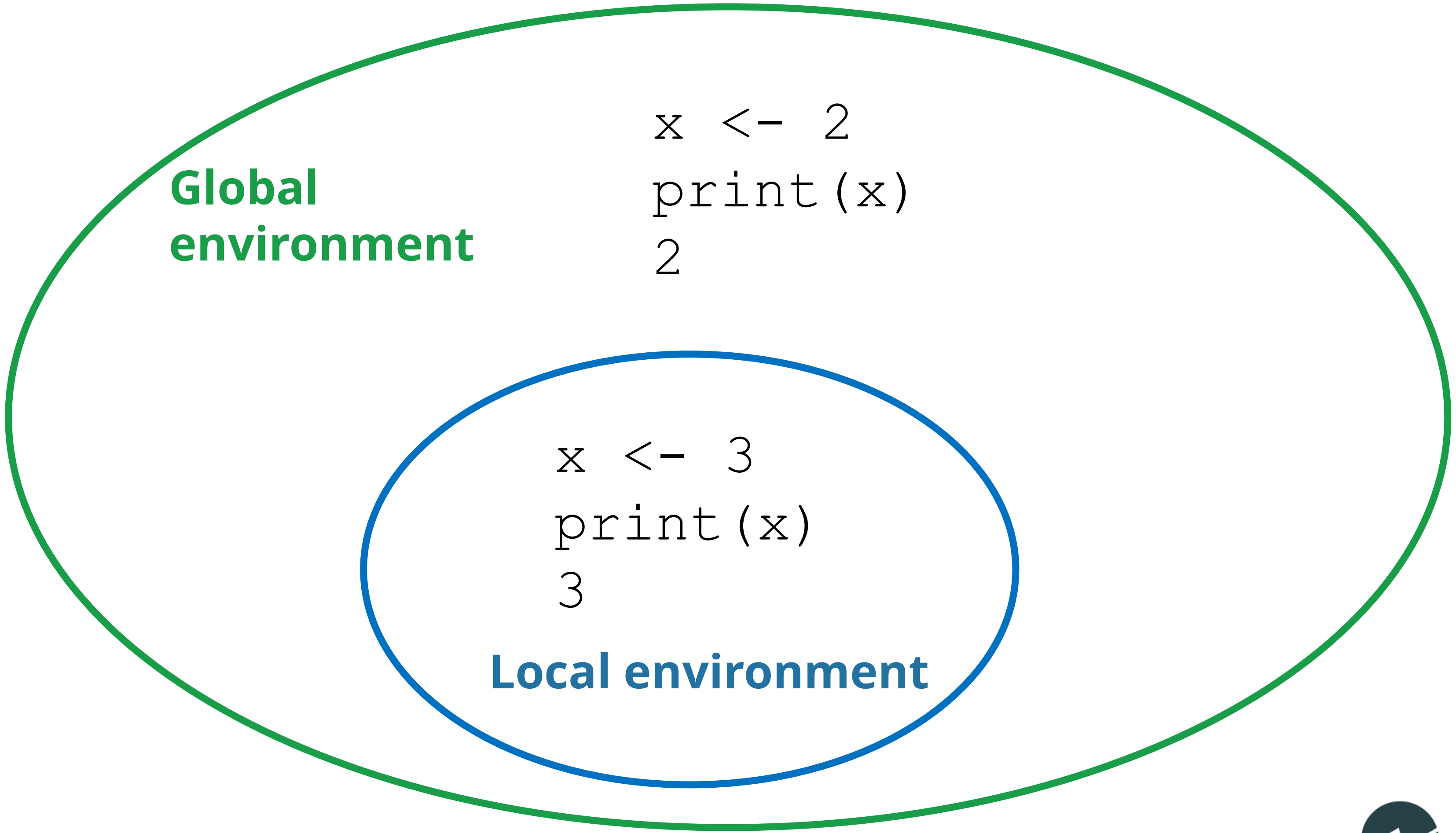
- Called “named parameter” “keyword arguments”

Statistisk sentralbyrå
Statistics Norway

Control in functions

- `warning()` : gives a warning message to user (but continues)
- `stop()` : gives a error message and stops
- `print()` : prints message to console. Useful for bug-fixing

Local vs global objects



```
x <- 2
print(x)
2
```

```
x <- 3
print(x)
3
```

tidyverse functions

- Data/vector is the first paremeter
- Can be used with mutate() and summarise() (for example)
- “if” does not work with vector parameters in functions (use ifelse)

Exercise 9

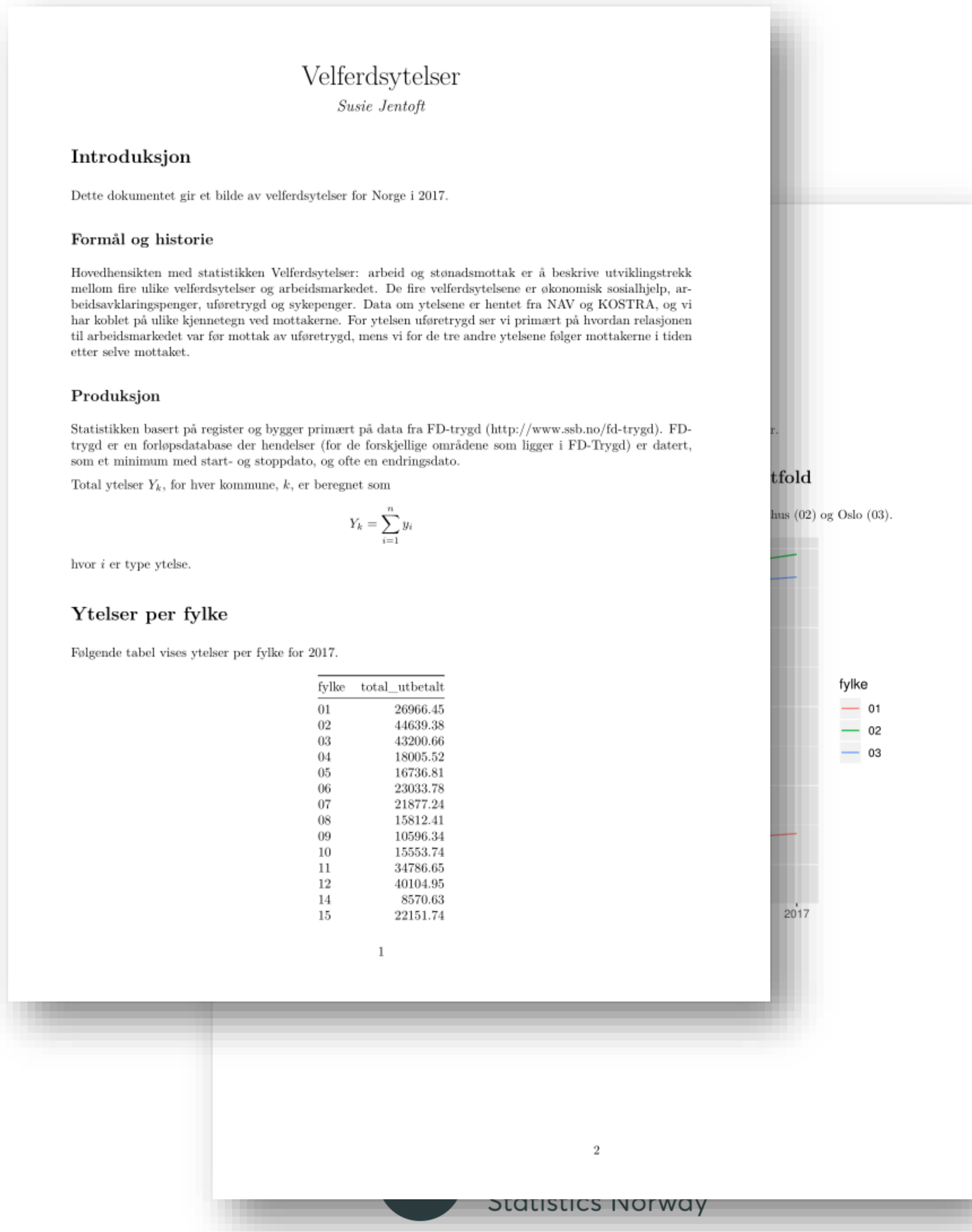
- Complete exercise 9 in the file Exercises_day5.R

RMarkdown

Analyze. Share. Reproduce.

Your data tells a story. Tell it with R Markdown.

Turn your analyses into high quality documents, reports, presentations and dashboards.



Velferdsytelser

Susie Jentoft

Introduksjon

Dette dokumentet gir et bilde av velferdsytelser for Norge i 2017.

Formål og historie

Hovedhensikten med statistikken Velferdsytelser: arbeid og stønadsmottak er å beskrive utviklingstrekk mellom fire ulike velferdsytelser og arbeidsmarkedet. De fire velferdsytelsene er økonomisk sosialhjelp, arbeidsavklaringspenger, uføretrygd og sykepenger. Data om ytelsene er hentet fra NAV og KOSTRA, og vi har koblet på ulike kjennetegn ved mottakerne. For ytelsen uføretrygd ser vi primært på hvordan relasjonen til arbeidsmarkedet var før mottak av uføretrygd, mens vi for de tre andre ytelsene følger mottakerne i tiden etter selve mottaket.

Produksjon

Statistikken basert på register og bygger primært på data fra FD-trygd (http://www.ssb.no/fd-trygd). FD-trygd er en forløpsdatabase der hendelser (for de forskjellige områdene som ligger i FD-Trygd) er datert, som et minimum med start- og stoppdato, og ofte en endringsdato.

Total ytelser Y_k, for hver kommune, k, er beregnet som

$$Y_k = \sum_{i=1}^{n} y_i$$

hvor i er type ytelse.

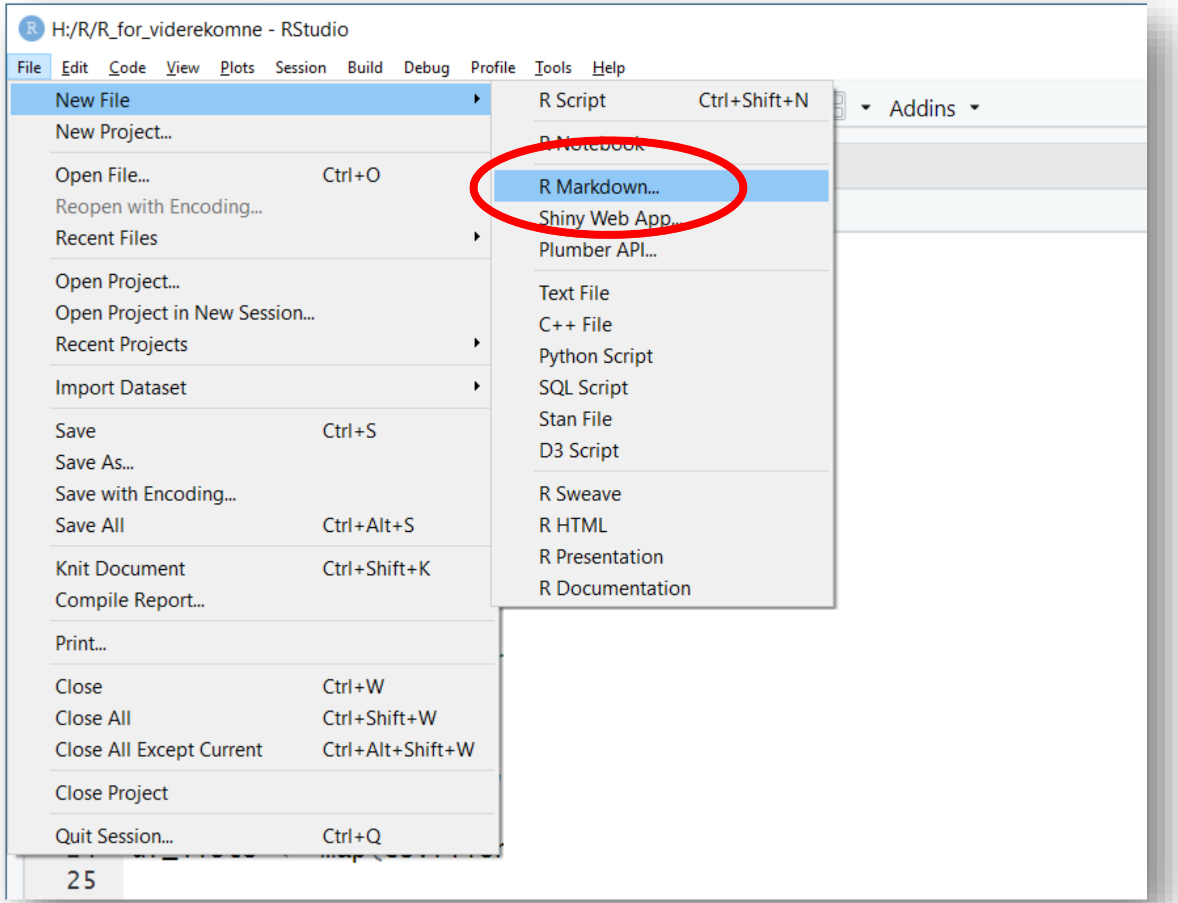
Ytelser per fylke

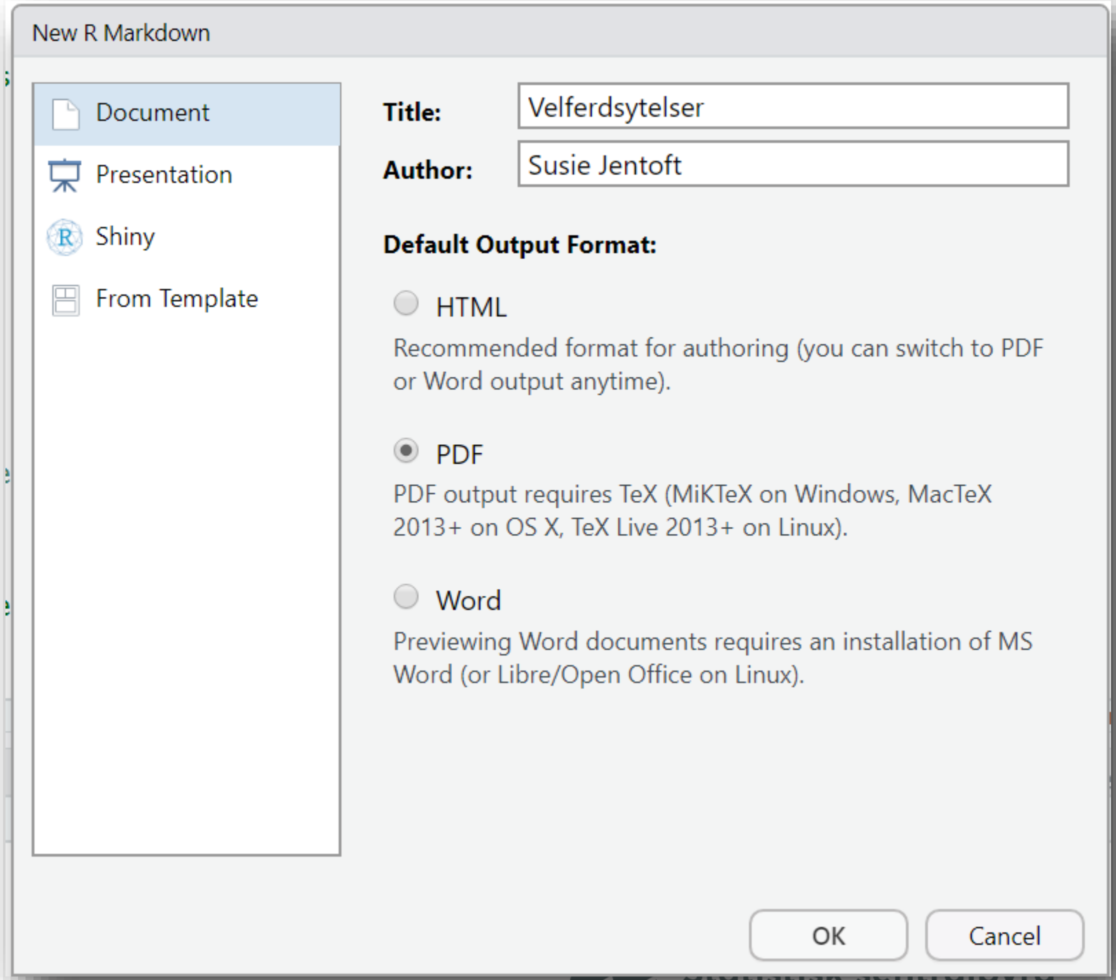
Følgende tabel vises ytelser per fylke for 2017.

fylke	total_utbetalt
01	26966.45
02	44639.38
03	43200.66
04	18005.52
05	16736.81
06	23033.78
07	21877.24
08	15812.41
09	10596.34
10	15553.74
11	34786.65
12	40104.95
14	8570.63
15	22151.74

1

RMarkdown




New R Markdown

Title: Velferdsytelser

Author: Susie Jentoft

Default Output Format:

HTML

Recommended format for authoring (you can switch to PDF or Word output anytime).

[✓ PDF]

PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).

Word

Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux).

RMarkdown: General

- Need Rtools
  - Including MikTex for compilation

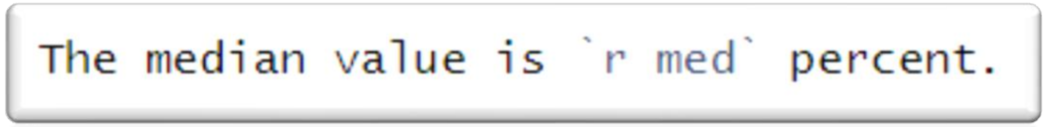
- Files saved as .Rmd

- Export as pdf /html click on the button

Statistisk sentralbyrå
Statistics Norway

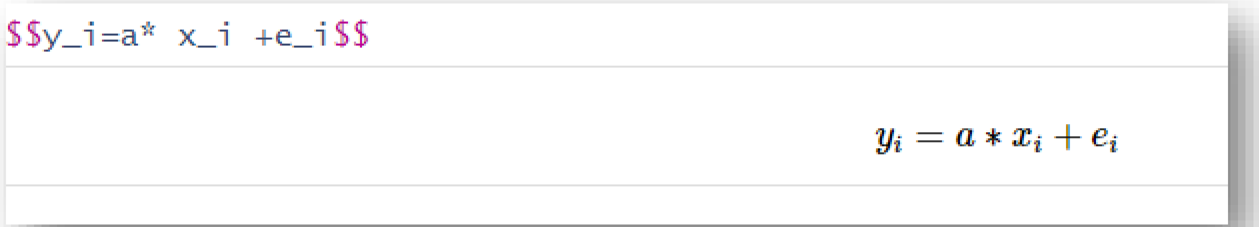
Code - Write text

- Write sentence as normal
- Use # for heading in text (## and ###): # My heading
- Use ** around words you want to have in bold letters : **bold**
- Write r code in text with: `r  «code»`



```
The median value is `r med` percent.
```

- To write formulas use $ around formulas (inline) or $$ (new line)



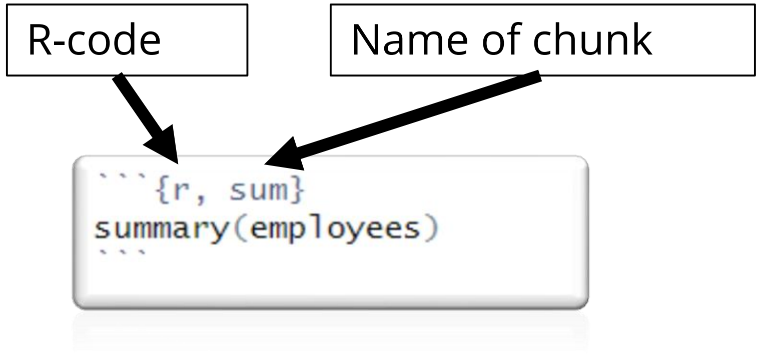
```
$$y_i=a* x_i +e_i$$
```

$$y_i = a * x_i + e_i$$

Statistisk sentralbyrå
Statistics Norway

RMarkdown: Code in r

- You can embed an R code chunk like this:



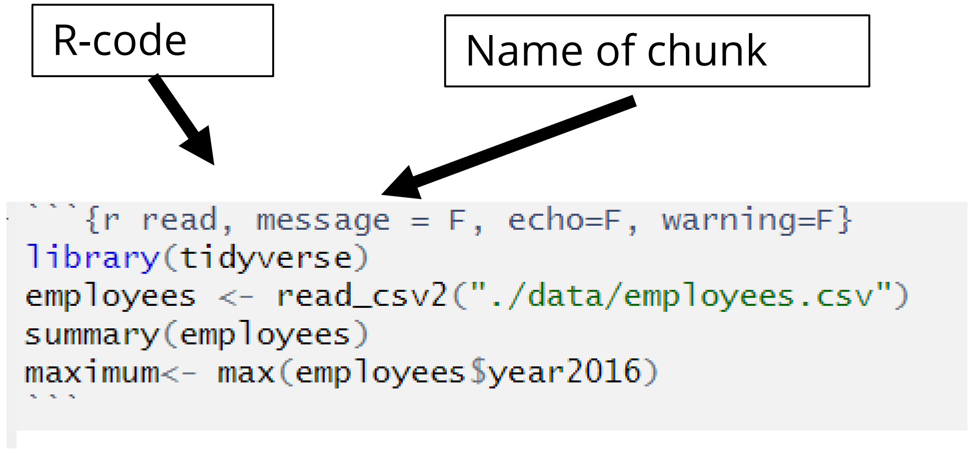
~~~
```{r, sum}
summary(employees)
```
~~~

- echo = F   To not show code
- message = F   To not show messages
- warning=F To not show warnings

Statistisk sentralbyrå
Statistics Norway

Code in r

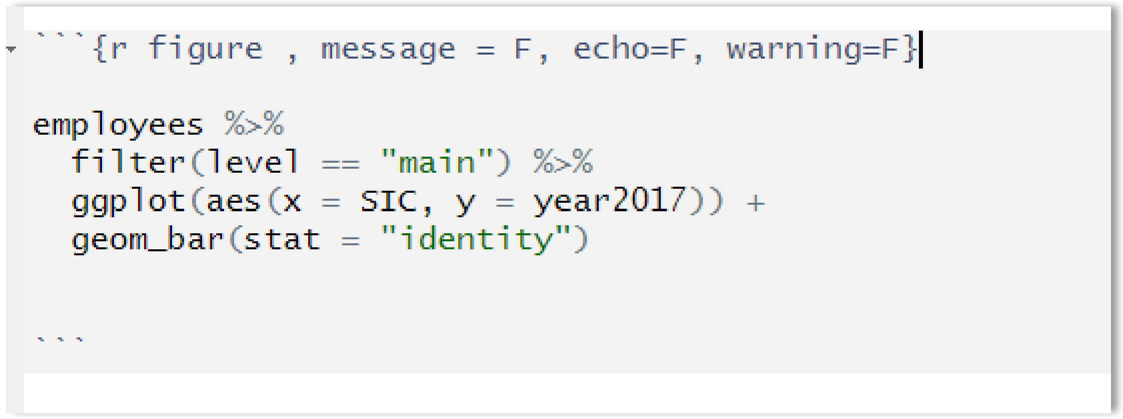
- You can embed an R code chunk like this:



~~~
```{r read, message = F, echo=F, warning=F}
library(tidyverse)
employees <- read_csv2("./data/employees.csv")
summary(employees)
maximum<- max(employees$year2016)
```
~~~

- echo = F   To not show code
- message = F   To not show messages
- Waring=F To not show warnings

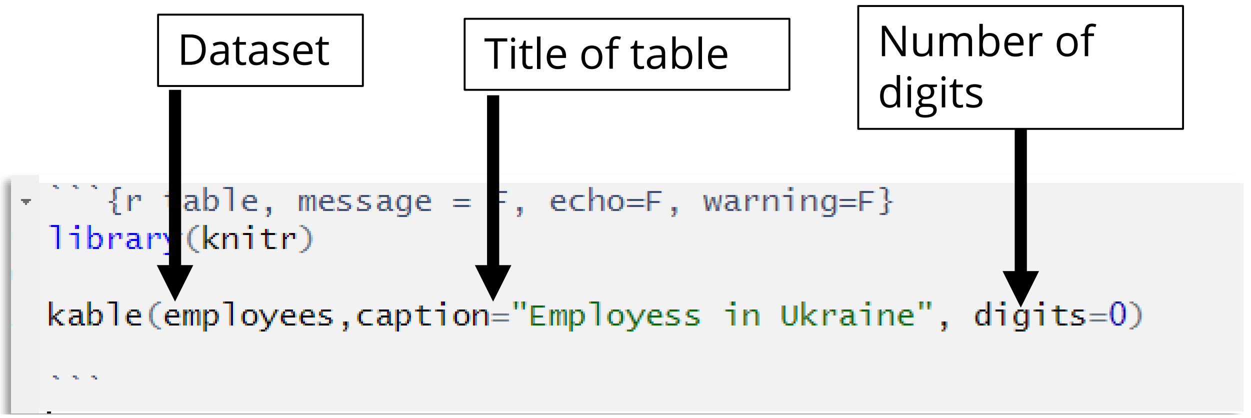
Figures



~~~
```{r figure , message = F, echo=F, warning=F}

employees %>%
  filter(level == "main") %>%
  ggplot(aes(x = SIC, y = year2017)) +
  geom_bar(stat = "identity")


```
~~~

Tables

- Package knitr



~~~
```{r table, message = F, echo=F, warning=F}
library(knitr)

kable(employees,caption="Employess in Ukraine", digits=0)

```
~~~

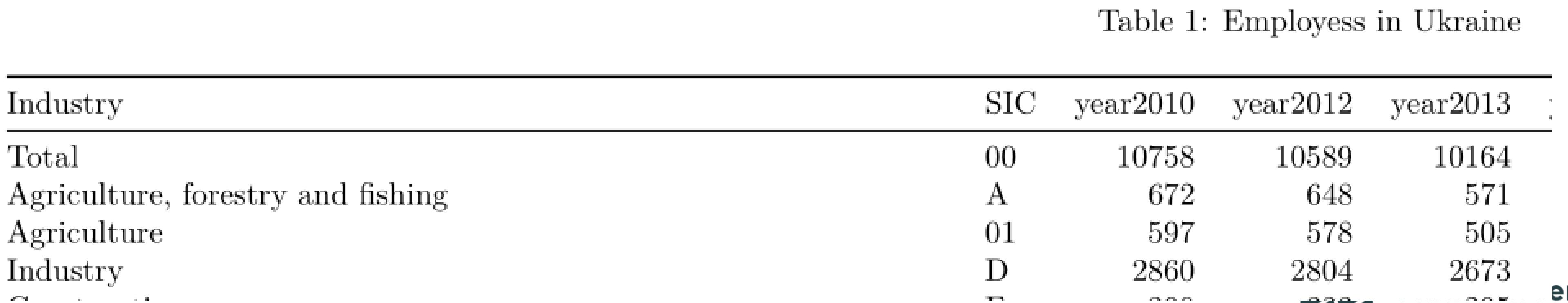
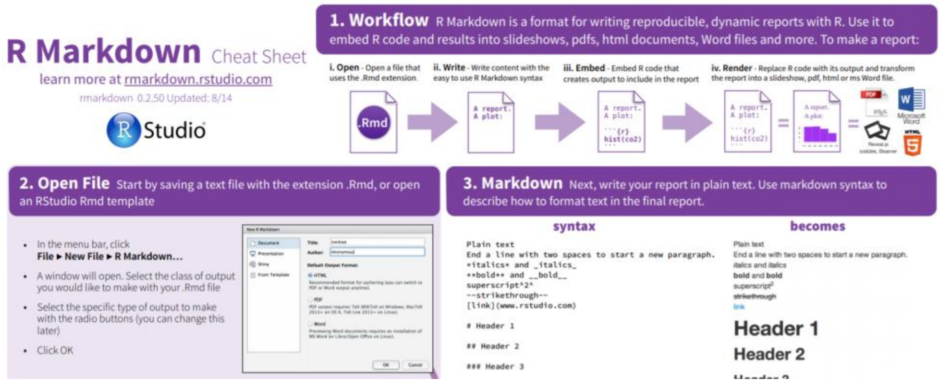


Table 1: Employess in Ukraine

Industry	SIC	year2010	year2012	year2013
Total	00	10758	10589	10164
Agriculture, forestry and fishing	A	672	648	571
Agriculture	01	597	578	505
Industry	D	2860	2804	2673

Resources

Cheatsheet: https://rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf



R Markdown Cheat Sheet
learn more at rmarkdown.rstudio.com
rmarkdown 0.2.50 Updated: 8/14

1. Workflow R Markdown is a format for writing reproducible, dynamic reports with R. Use it to embed R code and results into slideshows, pdfs, html documents, Word files and more. To make a report:

i. Open - Open a file that uses the .Rmd extension.
ii. Write - Write content with the easy to use R Markdown syntax
iii. Embed - Embed R code that creates output to include in the report
iv. Render - Replace R code with its output and transform the report into a slideshow, pdf, html or ms Word file.

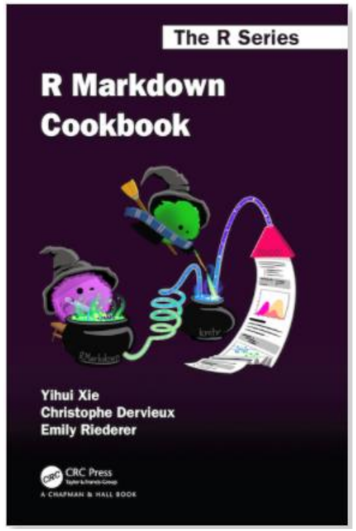
2. Open File Start by saving a text file with the extension .Rmd, or open an RStudio Rmd template

- In the menu bar, click **File ▸ New File ▸ R Markdown…**
- A window will open. Select the class of output you would like to make with your .Rmd file
- Select the specific type of output to make with the radio buttons (you can change this later)
- Click OK

3. Markdown Next, write your report in plain text. Use markdown syntax to describe how to format text in the final report.

syntax	becomes
`Plain text`	Plain text
`End a line with two spaces to start a new paragraph.`	End a line with two spaces to start a new paragraph.
`*italics* and _italics_`	*italics* and *italics*
`**bold** and __bold__`	**bold** and **bold**
`superscript^2^`	superscript^2
`--strikethrough--`	~~strikethrough~~
`[link](www.rstudio.com)`	link
`# Header 1`	Header 1
`## Header 2`	Header 2
`### Header 3`	Header 3

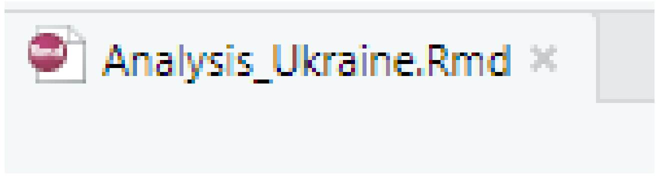
Cookbook: https://bookdown.org/yihui/rmarkdown-cookbook/



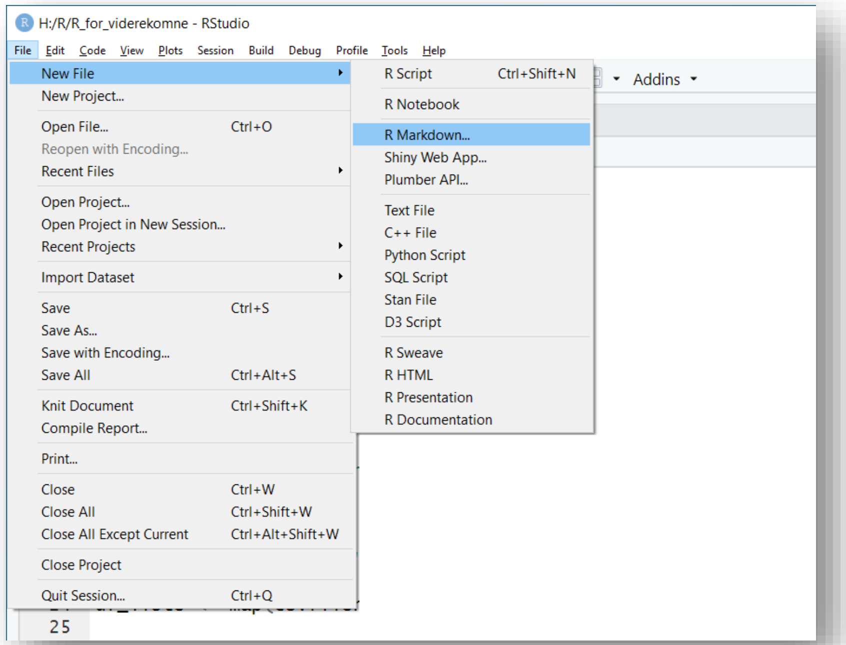
Statistisk sentralbyrå
Statistics Norway

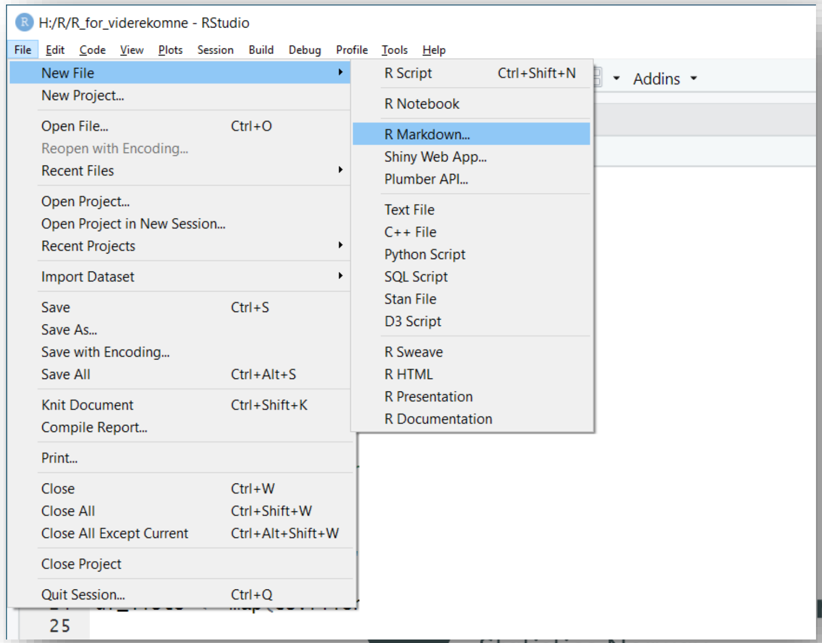
Example and exercise

- Example:

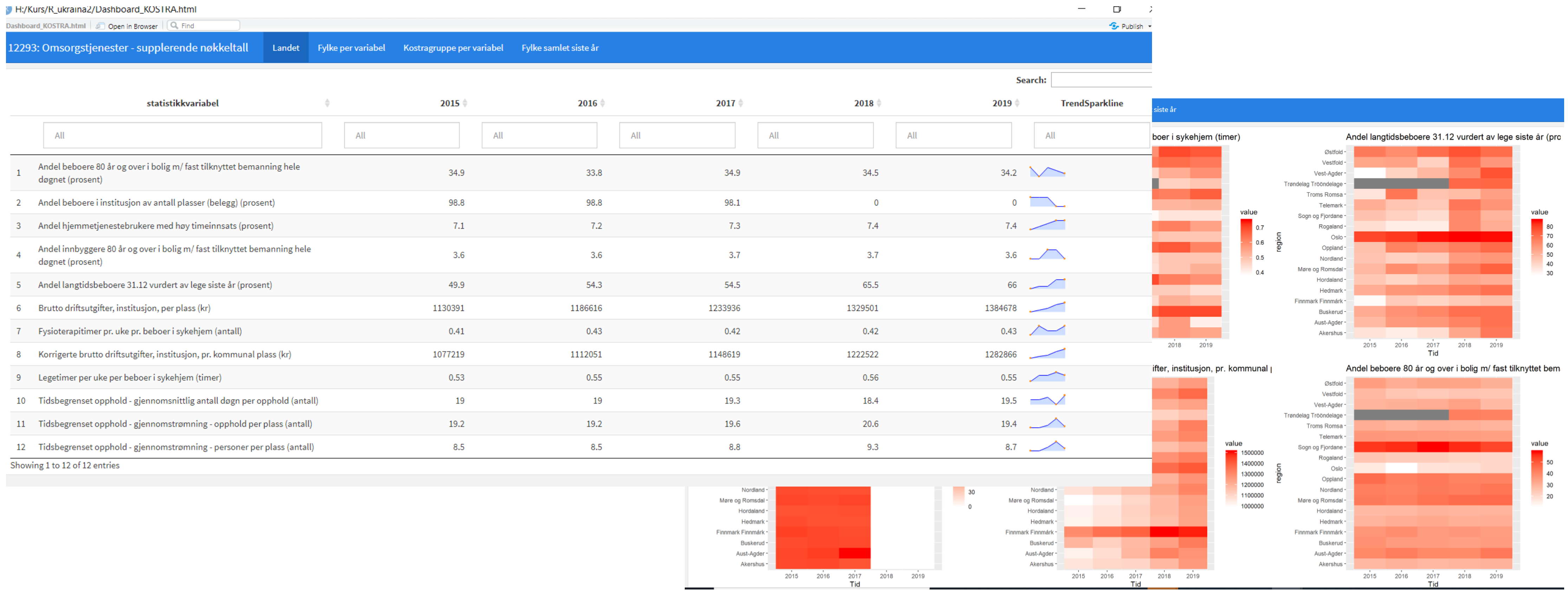

Analysis_Ukraine.Rmd

- Exercise: Try markdown!
  - Open a R markdown document with pdf and run the document with knitr
  - Change the text, add a heading, add a formula
  - Make a plot and a table





Dashboard in R


H:/Kurs/R_ukraina2/Dashboard_KOSTRA.html
Dashboard_KOSTRA.html
Open in Browser
Find
Publish
12293: Omsorgstjenester - supplerende nøkkeltall
Landet
Fylke per variabel
Kostragruppe per variabel
Fylke samlet siste år
Search:
siste år
boer i sykehjem (timer)
Andel langtidsbeboere 31.12 vurdert av lege siste år (pro
ifter, institusjon, pr. kommunal
Andel beboere 80 år og over i bolig m/ fast tilknyttet bem
value
region
Tid

	statistikkvariabel	2015	2016	2017	2018	2019	TrendSparkline
	All	All	All	All	All	All	All
1	Andel beboere 80 år og over i bolig m/ fast tilknyttet bemanning hele døgnet (prosent)	34.9	33.8	34.9	34.5	34.2	
2	Andel beboere i institusjon av antall plasser (belegg) (prosent)	98.8	98.8	98.1	0	0	
3	Andel hjemmetjenestebrukere med høy timeinnsats (prosent)	7.1	7.2	7.3	7.4	7.4	
4	Andel innbyggere 80 år og over i bolig m/ fast tilknyttet bemanning hele døgnet (prosent)	3.6	3.6	3.7	3.7	3.6	
5	Andel langtidsbeboere 31.12 vurdert av lege siste år (prosent)	49.9	54.3	54.5	65.5	66	
6	Brutto driftsutgifter, institusjon, per plass (kr)	1130391	1186616	1233936	1329501	1384678	
7	Fysioterapitimer pr. uke pr. beboer i sykehjem (antall)	0.41	0.43	0.42	0.42	0.43	
8	Korrigerte brutto driftsutgifter, institusjon, pr. kommunal plass (kr)	1077219	1112051	1148619	1222522	1282866	
9	Legetimer per uke per beboer i sykehjem (timer)	0.53	0.55	0.55	0.56	0.55	
10	Tidsbegrenset opphold - gjennomsnittlig antall døgn per opphold (antall)	19	19	19.3	18.4	19.5	
11	Tidsbegrenset opphold - gjennomstrømning - opphold per plass (antall)	19.2	19.2	19.6	20.6	19.4	
12	Tidsbegrenset opphold - gjennomstrømning - personer per plass (antall)	8.5	8.5	8.8	9.3	8.7	

Showing 1 to 12 of 12 entries

Statistisk sentralbyrå
Statistics Norway

Flexdashboard: Easy interactive dashboards for R

- Use R Markdown to publish a group of related data visualizations as a dashboard.
- Flexible and easy to specify row and column-based layouts. Components are intelligently re-sized to fill the browser
- Optionally use Shiny to drive visualizations dynamically.

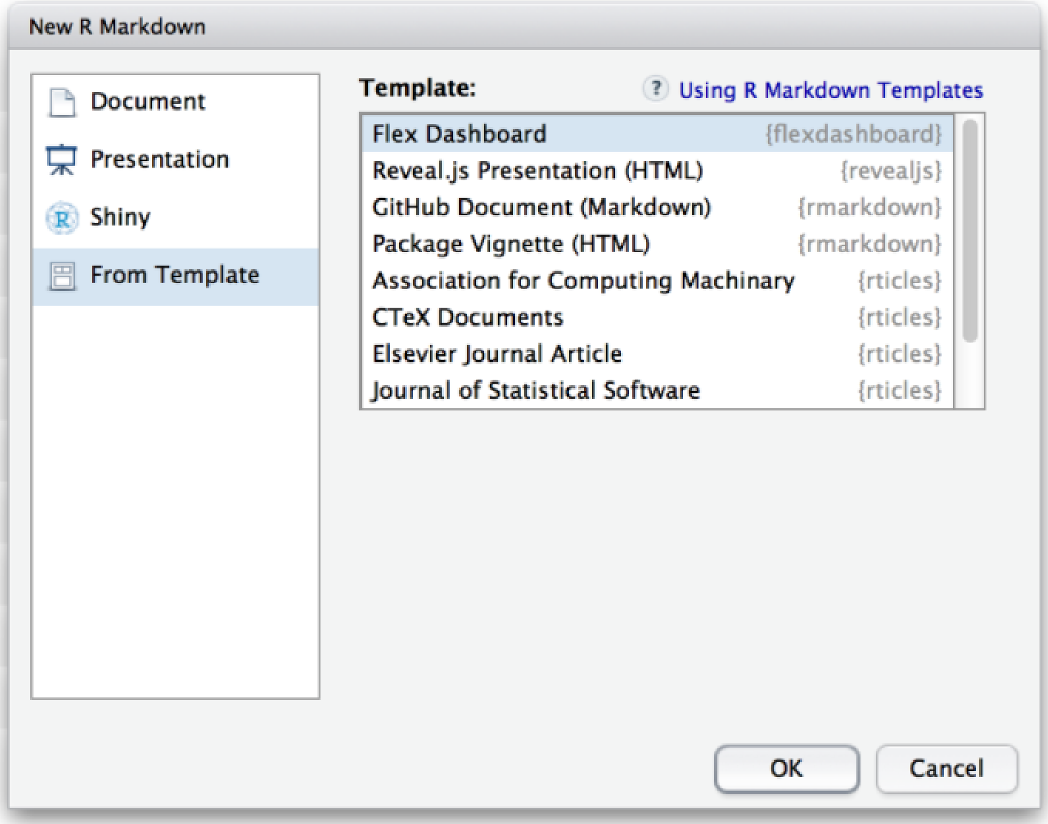
- https://rmarkdown.rstudio.com/flexdashboard/

Getting Started

Install the **flexdashboard** package from CRAN as follows:

```
install.packages("flexdashboard")
```

To author a flexdashboard you create an R Markdown document with the `flexdashboard::flex_dashboard` output format. You can do this from within RStudio using the **New R Markdown** dialog:


New R Markdown

Document
Presentation
Shiny
From Template

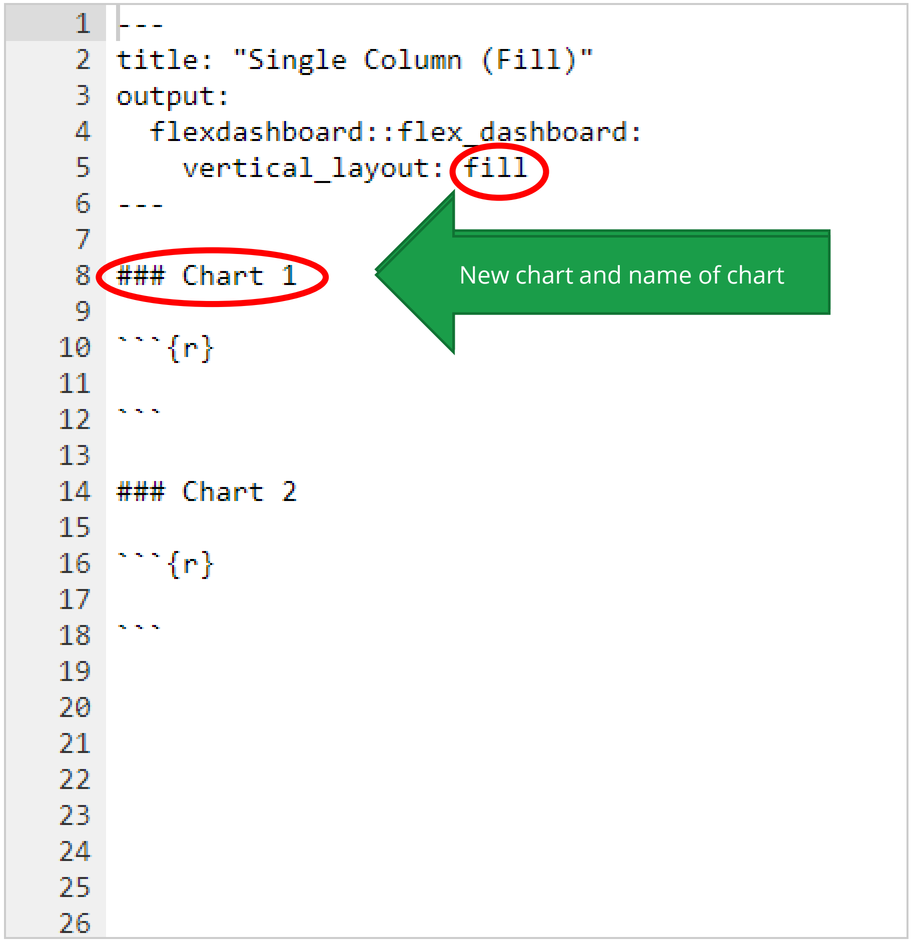
Template: Using R Markdown Templates

Flex Dashboard	{flexdashboard}
Reveal.js Presentation (HTML)	{revealjs}
GitHub Document (Markdown)	{rmarkdown}
Package Vignette (HTML)	{rmarkdown}
Association for Computing Machinary	{rticles}
CTeX Documents	{rticles}
Elsevier Journal Article	{rticles}
Journal of Statistical Software	{rticles}

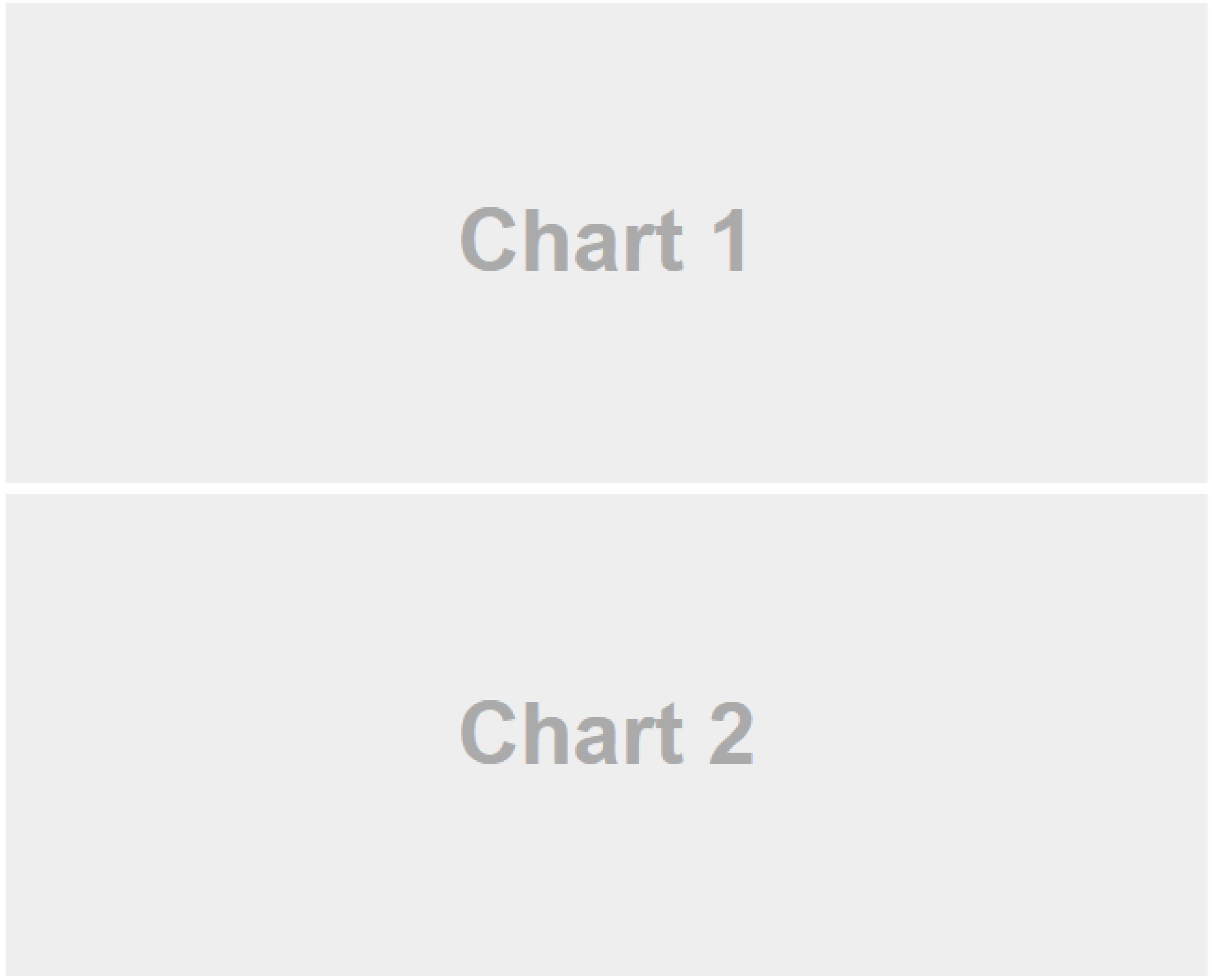
Statistisk sentralbyrå
Statistics Norway

Single Column (Fill)

Dashboards are divided into columns and rows, with output components delineated using level 3 markdown headers (`###`). By default, dashboards are laid out within a single column, with charts stacked vertically within a column and sized to fill available browser height. For example, this layout defines a single column with two charts that fills available browser space:

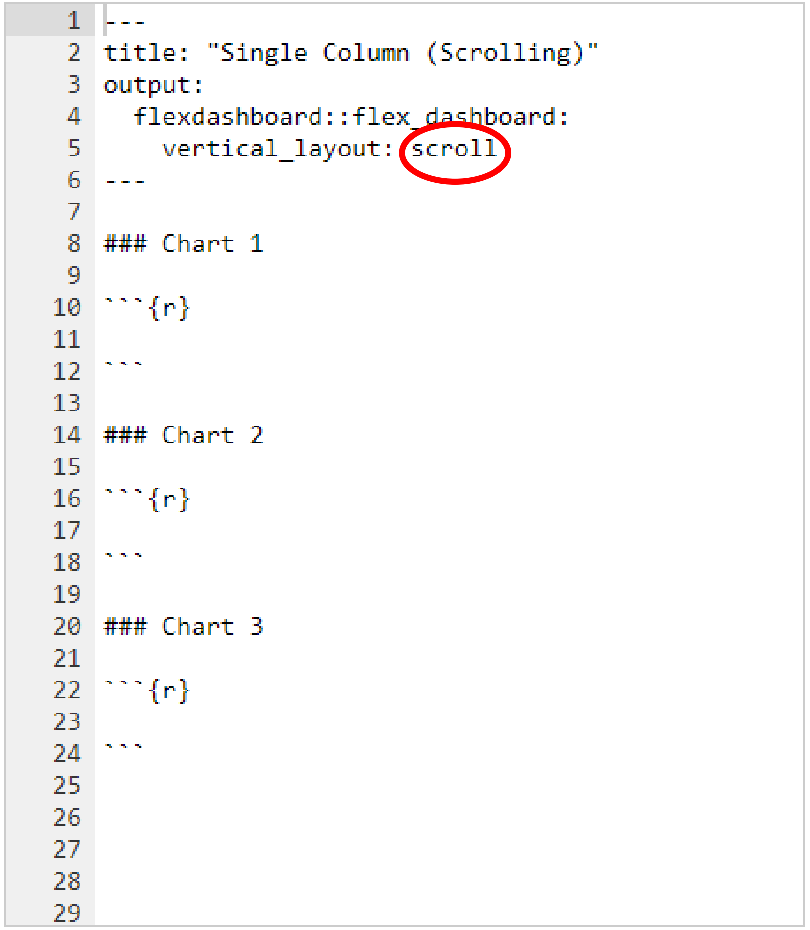


~~~
---
title: "Single Column (Fill)"
output:
  flexdashboard::flex_dashboard:
    vertical_layout: fill
---

### Chart 1

```{r}

```

### Chart 2

```{r}

```
~~~



Single Column (Scroll)

Depending on the nature of your dashboard (number of components, ideal height of components, etc.) you may prefer a scrolling layout where components occupy their natural height and the browser scrolls when additional vertical space is needed. You can specify this behavior via the `vertical_layout: scroll` option. For example, here is the definition of a single column scrolling layout with three charts:

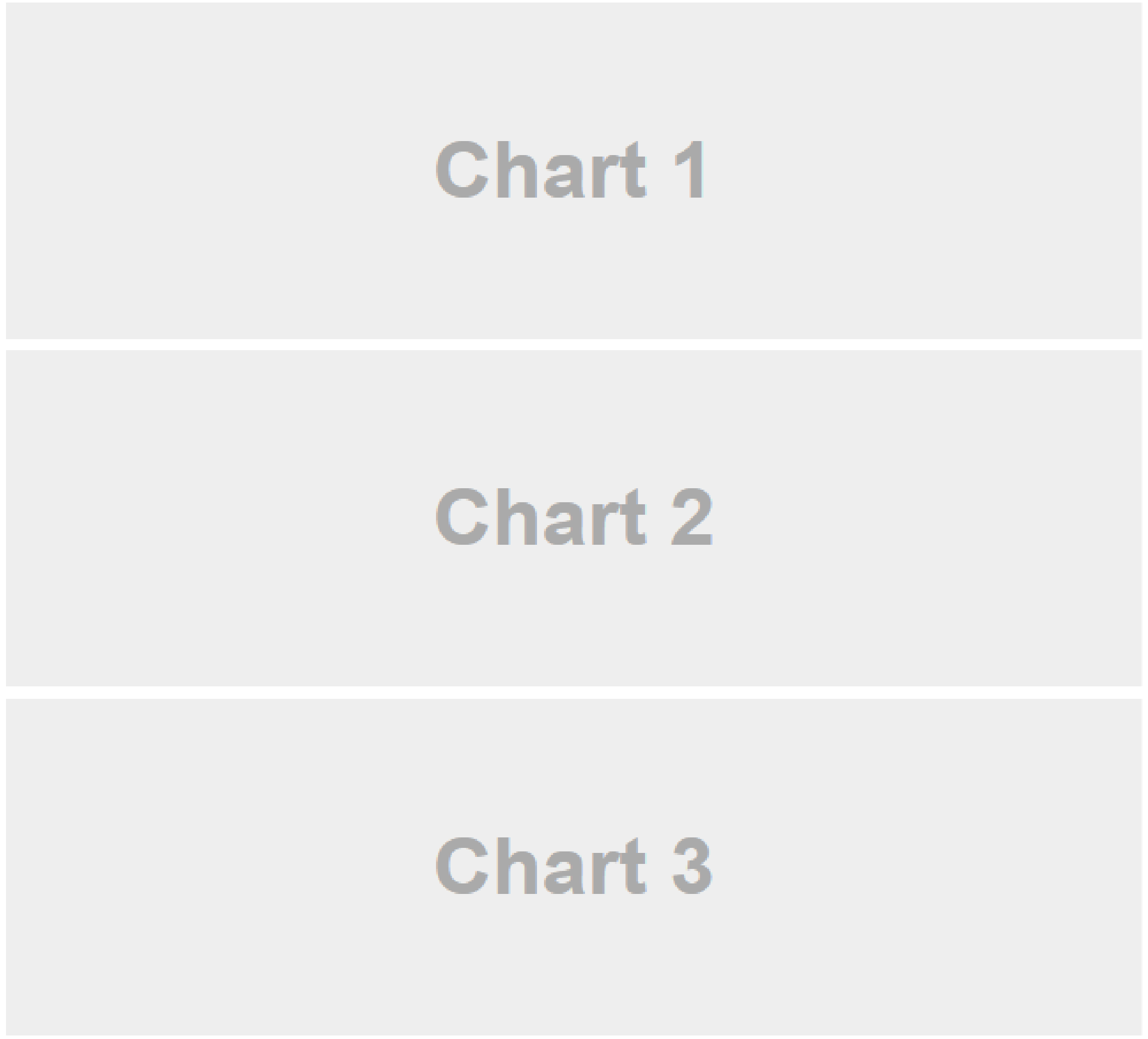


~~~
---
title: "Single Column (Scrolling)"
output:
  flexdashboard::flex_dashboard:
    vertical_layout: scroll
---

### Chart 1

```{r}

~~~

Chart 2

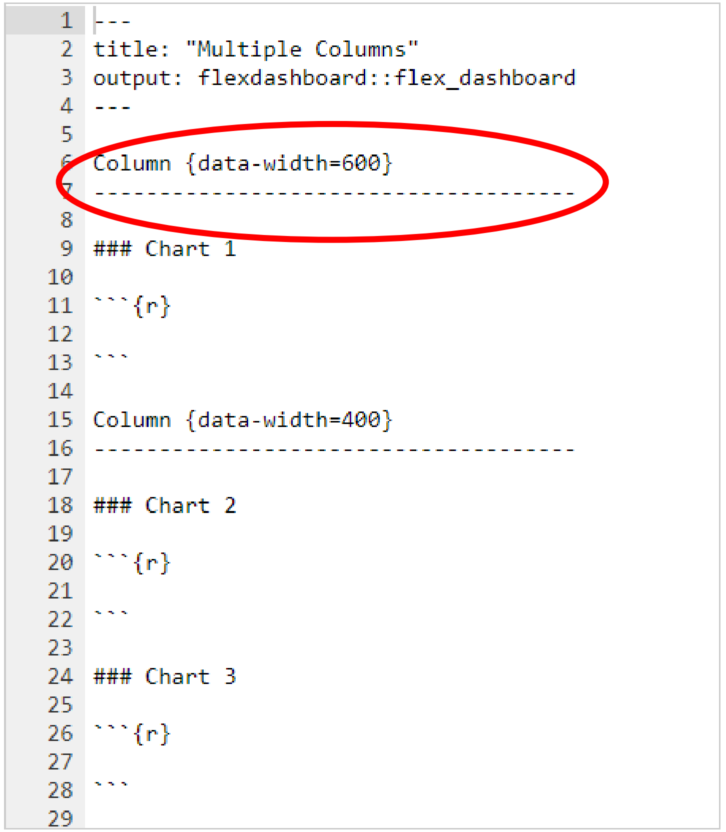
```{r}

```

Chart 3

```{r}

```
```



entralbyrå
way
```

## Multiple Columns

To lay out charts using multiple columns you introduce a level 2 markdown header (`--------------`) for each column. For example, this dashboard displays 3 charts split across two columns:



~~~
---
title: "Multiple Columns"
output: flexdashboard::flex_dashboard
---

Column {data-width=600}
-------------------------------------

### Chart 1

```{r}

~~~

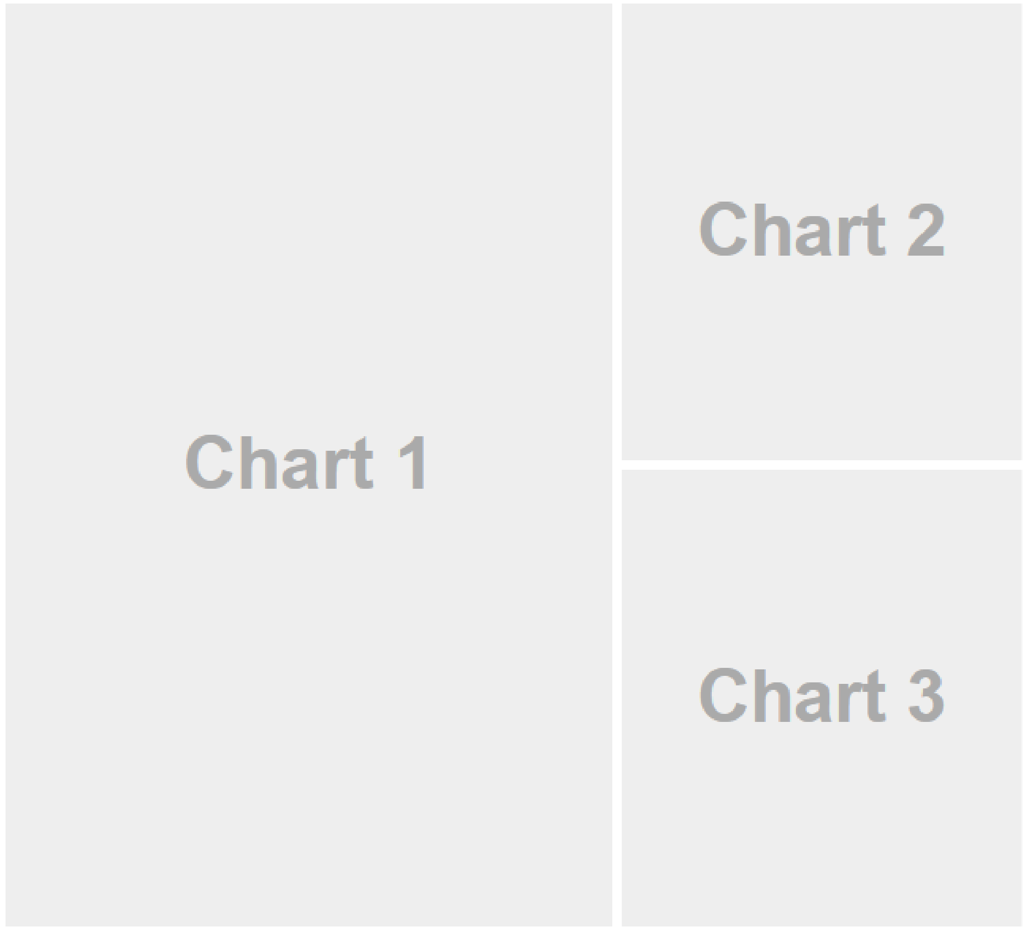
Column {data-width=400}
-------------------------------------

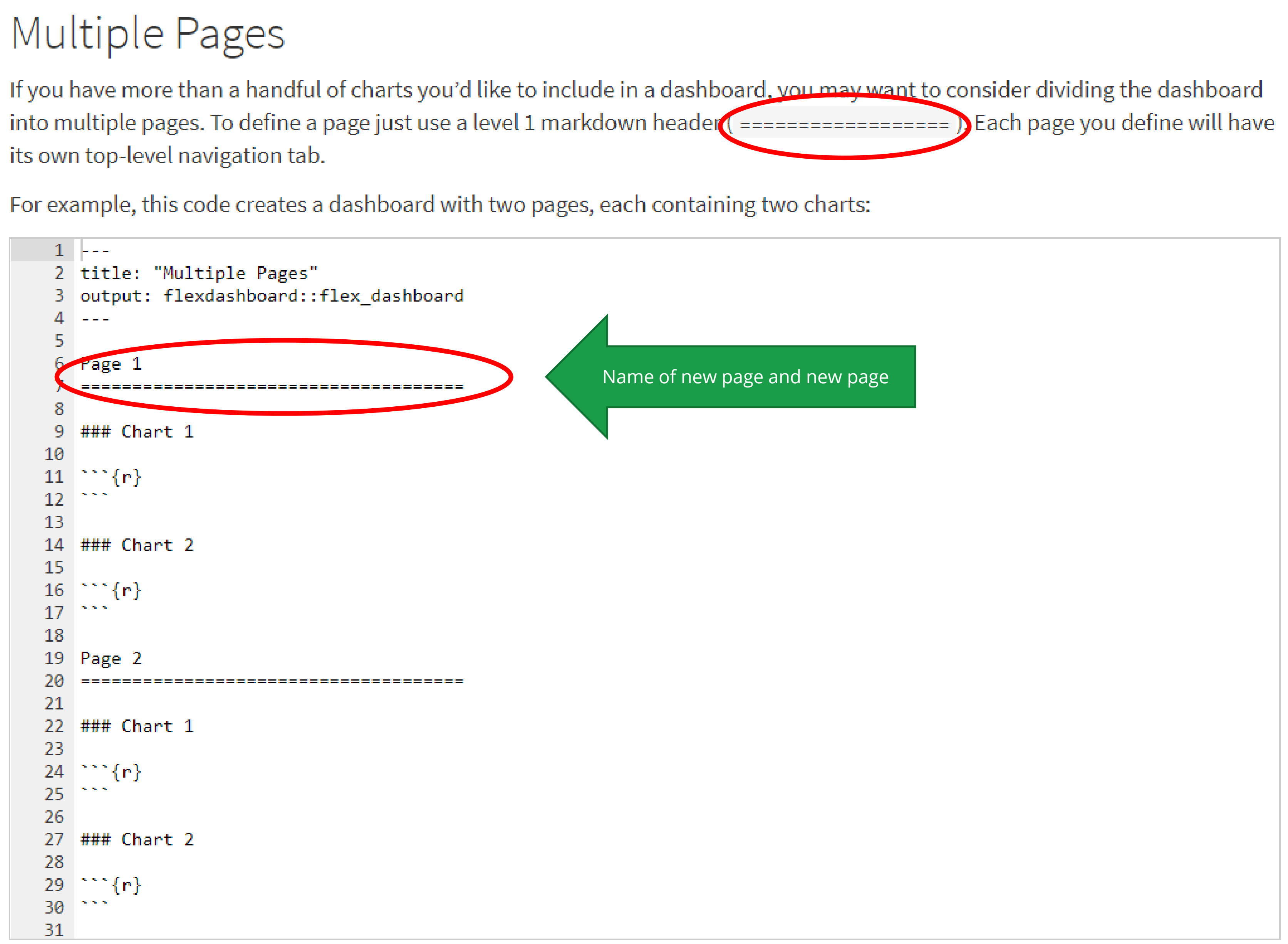
### Chart 2

```{r}

```

Chart 3

```{r}

```

```



In this example we've moved Chart 1 into its own column which it will fill entirely. We've also given the column a larger size via the `data-width` attribute to provide additional emphasis to Chart 1.
```



Multiple Pages

If you have more than a handful of charts you'd like to include in a dashboard, you may want to consider dividing the dashboard into multiple pages. To define a page just use a level 1 markdown header (`==================`). Each page you define will have its own top-level navigation tab.

For example, this code creates a dashboard with two pages, each containing two charts:

~~~
---
title: "Multiple Pages"
output: flexdashboard::flex_dashboard
---

Page 1
=====================================

### Chart 1

```{r}
~~~

Chart 2

```{r}
```

Page 2
=====================================

Chart 1

```{r}
```

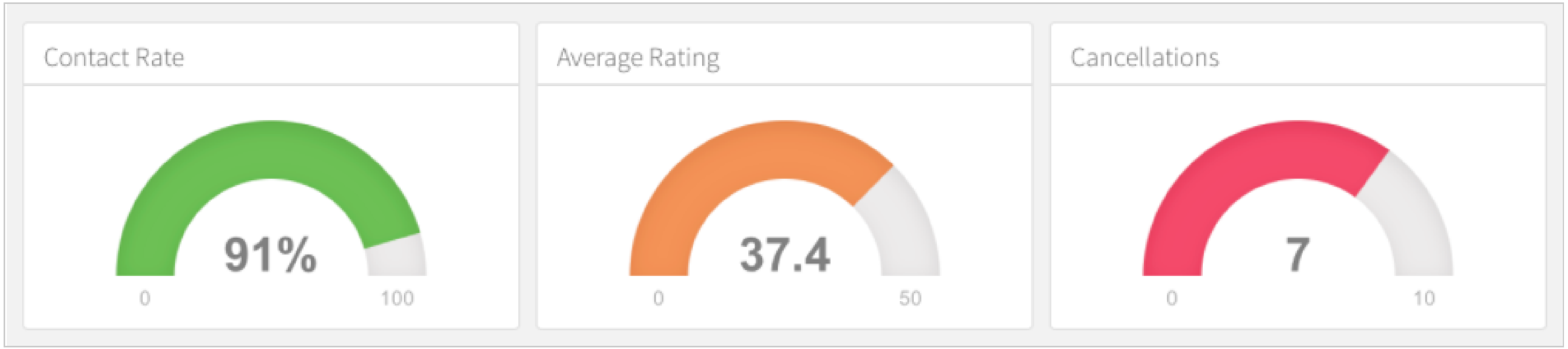
Chart 2

```{r}
```

```
```

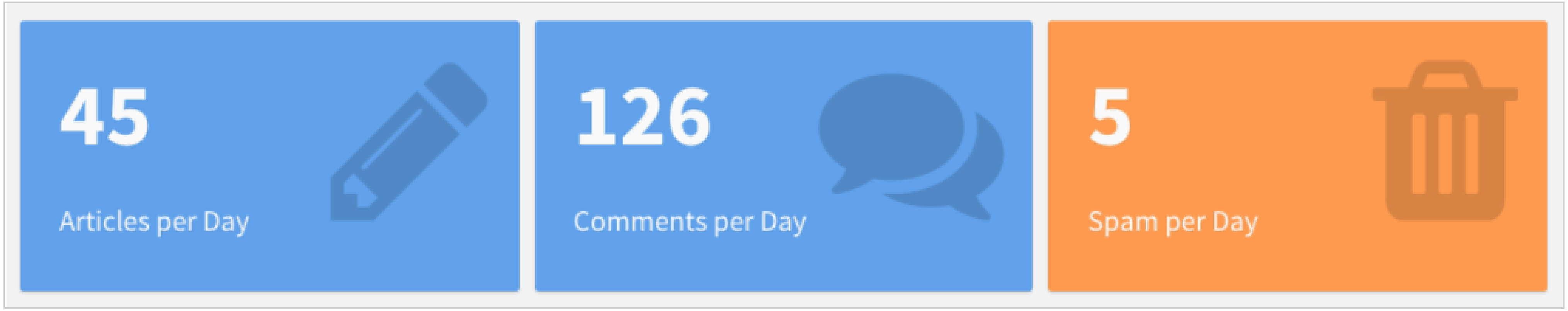
Gauges

Gauges display values on a meter within a specified range. For example, here is a set of 3 gauges:



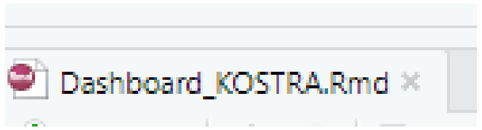
Value Boxes

Sometimes you want to include one or more simple values within a dashboard. You can use the `valueBox` function to display single values along with a title and optional icon. For example, here are three side-by-side sections each displaying a single value:



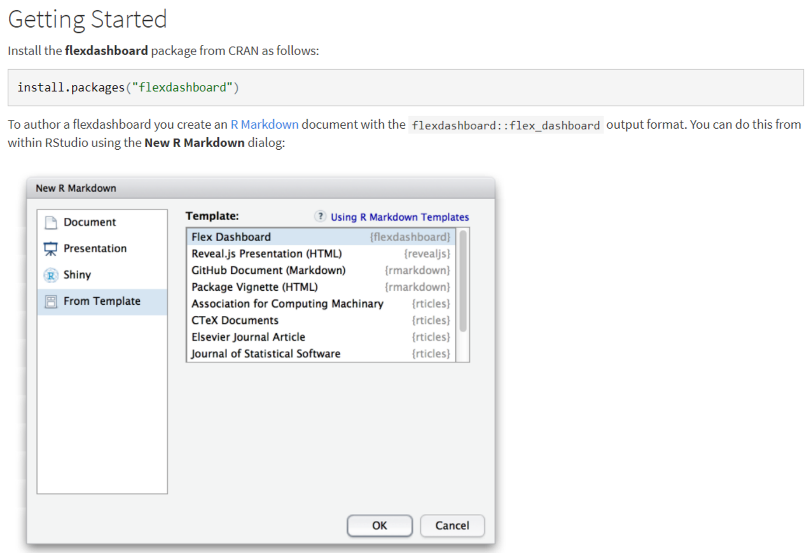
lbyrå

Example and exercise

- Example: 

Dashboard_KOSTRA.Rmd

- Exercise: Try dashboard!
  - Open a R markdown document with a dashboard template and
    - Make plots
    - Run the document with knitr



Getting Started

Install the **flexdashboard** package from CRAN as follows:

```
install.packages("flexdashboard")
```

To author a flexdashboard you create an R Markdown document with the `flexdashboard::flex_dashboard` output format. You can do this from within RStudio using the **New R Markdown** dialog:

Flex Dashboard	{flexdashboard}
Reveal.js Presentation (HTML)	{revealjs}
GitHub Document (Markdown)	{rmarkdown}
Package Vignette (HTML)	{rmarkdown}
Association for Computing Machinary	{rticles}
CTeX Documents	{rticles}
Elsevier Journal Article	{rticles}
Journal of Statistical Software	{rticles}

Statistisk sentralbyrå
Statistics Norway

Exercise 10

Statistisk sentralbyrå
Statistics Norway