



Managing dependencies



Managing and executing analysis workflow



Versioning and collaborating on code (and some other files)

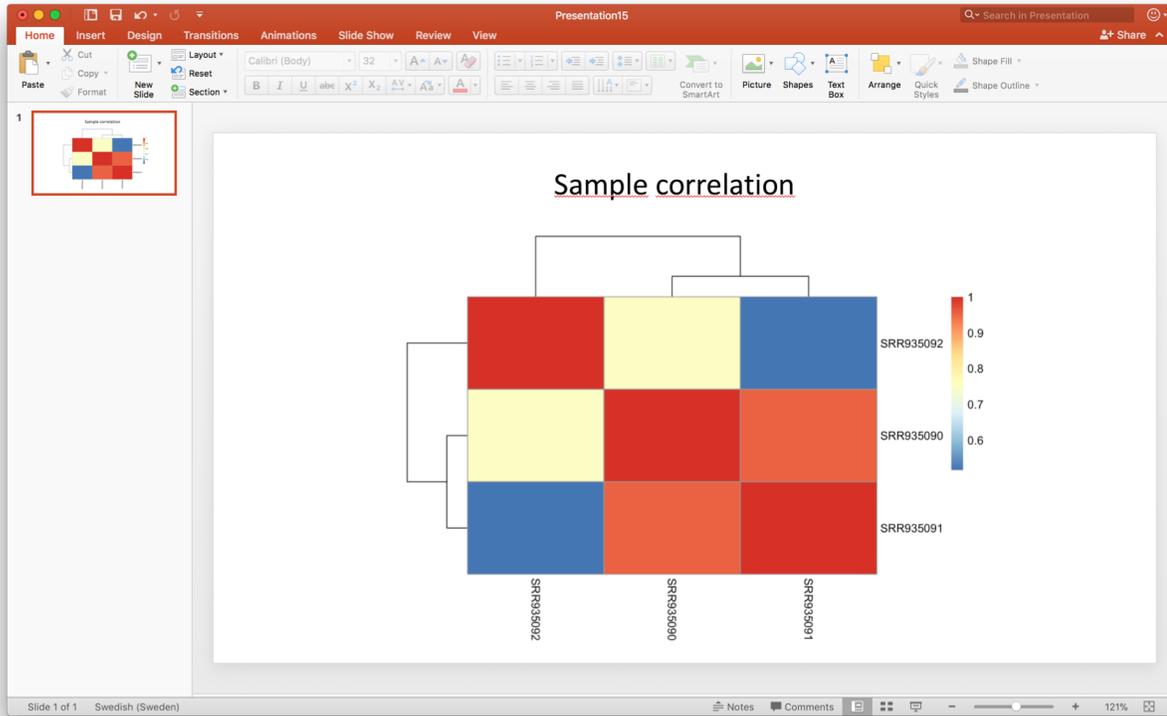


Connecting code and reporting

and...



Isolating and exporting environment



- R Markdown makes your analysis more reproducible by connecting your code, figures and descriptive text.
- You can use it to make reproducible reports, rather than e.g. copy-pasting figures into a Word document.
- You can also use it as a notebook, in the same way as lab notebooks are used in a wet lab setting.

# Supplementary material

John Doe, Joan Dough, Jan Doh, Dyon Do

18 March, 2018

## Read in the data

We have *count data* for three samples:

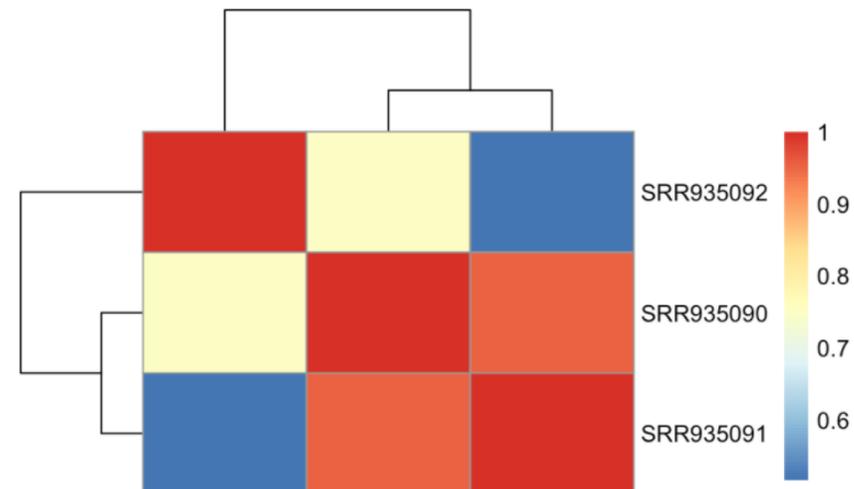
- SRR935090
- SRR935091
- SRR935092

```
# Read counts:
counts_file <- "results/tables/counts.tsv"
counts <- read.delim(counts_file, skip=1, header=F, row.names=1)
sample_names <- t(read.delim(counts_file, nrow=1, header=F))
colnames(counts) <- c("description", gsub(".*(SRR[0-9]*)\\.\\.\\.\\.", "\\1", sample_names))
```

## Plot sample correlation

Next, we will plot the sample correlation of the *count data*:

```
heatmap(cor((counts[,-1])), show_colnames=F)
```



```
---
title: "Supplementary material"
author: John Doe, Joan Dough, Jan Doh, Dyon Do
date: "`r format(Sys.time(), '%d %B, %Y')`"
output: html_document
---
```

```
```{r, include=FALSE}
library("pheatmap")
```
```

```
# Read in the data
```

We have *count data* for three samples:

```
- SRR935090
- SRR935091
- SRR935092
```

```
```{r}
# Read counts:
counts_file <- "results/tables/counts.tsv"
counts <- read.delim(counts_file, skip=1, header=F, row.names=1)
sample_names <- t(read.delim(counts_file, nrow=1, header=F))
colnames(counts) <- c("description",
                     gsub(".*(SRR[0-9]*)\\.\\.\\.","\\1", sample_names))
```
```

```
# Plot sample correlation
```

Next, we will plot the sample correlation of the *count data*:

```
```{r, fig.height=3, fig.width=5}
pheatmap(cor((counts[,-1])), show_colnames=F)
```
```

# Supplementary material

John Doe, Joan Dough, Jan Doh, Dyon Do

18 March, 2018

## Read in the data

We have *count data* for three samples:

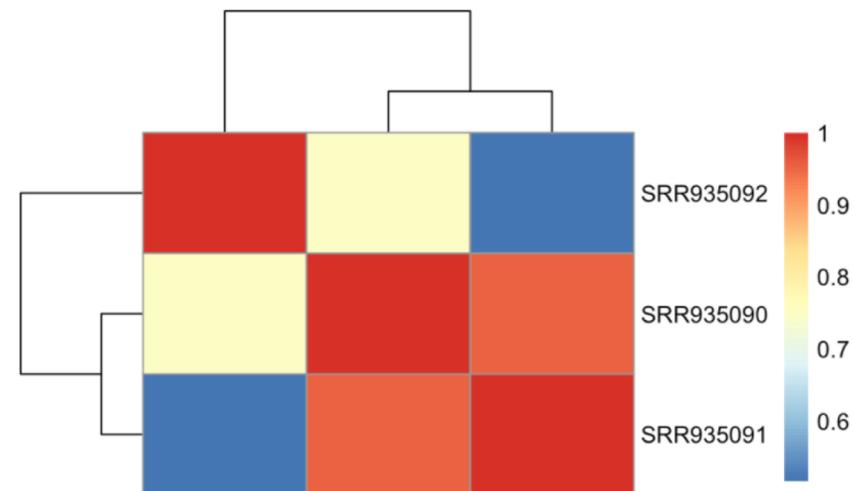
- SRR935090
- SRR935091
- SRR935092

```
# Read counts:
counts_file <- "results/tables/counts.tsv"
counts <- read.delim(counts_file, skip=1, header=F, row.names=1)
sample_names <- t(read.delim(counts_file, nrow=1, header=F))
colnames(counts) <- c("description", gsub(".*(SRR[0-9]*)\\.\\.\\.","\\1", sample_names))
```

## Plot sample correlation

Next, we will plot the sample correlation of the *count data*:

```
pheatmap(cor((counts[,-1])), show_colnames=F)
```



```
---  
title: "Supplementary material"  
author: John Doe, Joan Dough, Jan Doh, Dyon Do  
date: "`r format(Sys.time(), '%d %B, %Y')`"  
output: html_document
```

```
```{r, include=FALSE}  
library("pheatmap")  
```
```

```
# Read in the data
```

We have *count data* for three samples:

- SRR935090
- SRR935091
- SRR935092

```
```{r}  
# Read counts:  
counts_file <- "results/tables/counts.tsv"  
counts <- read.delim(counts_file, skip=1, header=F, row.names=1)  
sample_names <- t(read.delim(counts_file, nrow=1, header=F))  
colnames(counts) <- c("description",  
                    gsub(".*(SRR[0-9]*)\\.\\.\\.", "\\1", sample_names))  
```
```

```
# Plot sample correlation
```

Next, we will plot the sample correlation of the *count data*:

```
```{r, fig.height=3, fig.width=5}  
pheatmap(cor((counts[,-1])), show_colnames=F)  
```
```

## Header in YAML format

- Document-wide options
- Output format
- Parameters

```
---
title: "Supplementary material"
author: John Doe, Joan Dough, Jan Doh, Dyon Do
date: "`r format(Sys.time(), '%d %B, %Y')`"
output: html_document
---
```

```
```{r, include=FALSE}
library("pheatmap")
```
```

```
# Read in the data
```

```
We have count data for three samples:
```

```
- SRR935090
- SRR935091
- SRR935092
```

```
```{r}
# Read counts:
counts_file <- "results/tables/counts.tsv"
counts <- read.delim(counts_file, skip=1, header=F, row.names=1)
sample_names <- t(read.delim(counts_file, nrow=1, header=F))
colnames(counts) <- c("description",
                    gsub(".*(SRR[0-9]*)\\.\\.\\.","\\1", sample_names))
```
```

```
# Plot sample correlation
```

```
Next, we will plot the sample correlation of the count data:
```

```
```{r, fig.height=3, fig.width=5}
pheatmap(cor((counts[,-1])), show_colnames=F)
```
```

## Header in YAML format

- Document-wide options
- Output format
- Parameters

## Code chunks

- Evaluate R code and show output
- Also Bash, Python, Rcpp, SQL, Stan
- Chunk options

```
--  
title: "Supplementary material"  
author: John Doe, Joan Dough, Jan Doh, Dyon Do  
date: "`r format(Sys.time(), '%d %B, %Y')`"  
output: html_document  
---
```

```
```{r, include=FALSE}  
library("pheatmap")  
```
```

```
# Read in the data
```

We have *count data* for three samples:

- SRR935090
- SRR935091
- SRR935092

```
```{r}  
# Read counts:  
counts_file <- "results/tables/counts.tsv"  
counts <- read.delim(counts_file, skip=1, header=F, row.names=1)  
sample_names <- t(read.delim(counts_file, nrow=1, header=F))  
colnames(counts) <- c("description",  
                      gsub(".*(SRR[0-9]*)\\.\\.\\.", "\\1", sample_names))  
```
```

```
# Plot sample correlation
```

Next, we will plot the sample correlation of the *count data*:

```
```{r, fig.height=3, fig.width=5}  
pheatmap(cor((counts[,-1])), show_colnames=F)  
```
```

## Header in YAML format

- Document-wide options
- Output format
- Parameters

## Code chunks

- Evaluate R code and show output
- Also Bash, Python, Rcpp, SQL, Stan
- Chunk options

## Markdown text

- Freely add and format text using markdown

```
---
title: "Supplementary material"
author: John Doe, Joan Dough, Jan Doh, Dyon Do
date: "`r format(Sys.time(), '%d %B, %Y')`"
output: html_document
---
```

```
```{r, include=FALSE}
library("pheatmap")
```
```

```
# Read in the data
```

We have *count data* for three samples:

```
- SRR935090
- SRR935091
- SRR935092
```

```
```{r}
# Read counts:
counts_file <- "results/tables/counts.tsv"
counts <- read.delim(counts_file, skip=1, header=F, row.names=1)
sample_names <- t(read.delim(counts_file, nrow=1, header=F))
colnames(counts) <- c("description",
                     gsub(".*(SRR[0-9]*)\\.\\.\\.","\\1", sample_names))
```
```

```
# Plot sample correlation
```

Next, we will plot the sample correlation of the *count data*:

```
```{r, fig.height=3, fig.width=5}
pheatmap(cor((counts[,-1])), show_colnames=F)
```
```

# Supplementary material

R Markdown

from  Studio

John Doe, Joan Dough, Jan Doh, Dyon Do

18 March, 2018

## Read in the data

We have *count data* for three samples:

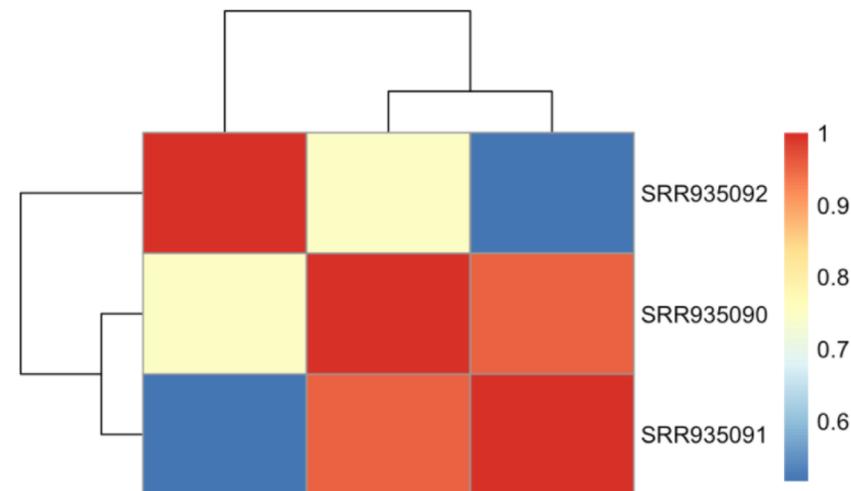
- SRR935090
- SRR935091
- SRR935092

```
# Read counts:
counts_file <- "results/tables/counts.tsv"
counts <- read.delim(counts_file, skip=1, header=F, row.names=1)
sample_names <- t(read.delim(counts_file, nrow=1, header=F))
colnames(counts) <- c("description", gsub(".*(SRR[0-9]*)\\.\\.\\.","\\1", sample_names))
```

## Plot sample correlation

Next, we will plot the sample correlation of the *count data*:

```
pheatmap(cor((counts[,-1])), show_colnames=F)
```



# Output formats

- Documents/reports (HTML, PDF, MS Word, Tufte handouts)
- Presentations (Powerpoint, Beamer, Slidy, ioslides, reveal.js)
- Interactive documents and dashboards (HTML widgets, Shiny)
- Books and websites
- Other templates...

Can require different markdown syntax depending on output!



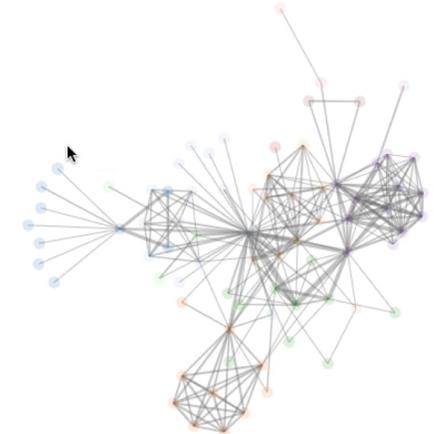
|                 |                                      |
|-----------------|--------------------------------------|
| Leaflet         | Geo-spatial mapping                  |
| dygraphs        | Time series charting                 |
| Plotly          | Interactive graphics with D3         |
| rbokeh          | R interface to Bokeh                 |
| Highcharter     | R interface to Highcharts            |
| visNetwork      | Graph data visualization with vis.js |
| networkD3       | Graph data visualization with D3     |
| d3heatmap       | Interactive heatmaps with D3         |
| DataTables      | Tabular data display                 |
| threejs         | 3D scatterplots and globes           |
| rglwidget       | Render scenes created with rgl       |
| DiagrammeR      | Diagrams and flowcharts              |
| MetricsGraphics | Scatterplots and line charts with D3 |

networkD3

<http://christophergandrud.github.io/networkD3/>

networkD3 provides tools for creating D3 JavaScript network graphs from R.

```
library(networkD3)
data(MisLinks, MisNodes)
forceNetwork(Links = MisLinks, Nodes = MisNodes, Source = "source",
             Target = "target", Value = "value", NodeID = "name",
             Group = "group", opacity = 0.4)
```



# R Markdown in RStudio

The screenshot shows the RStudio interface with the following elements:

- Files pane:** Shows a project structure with files like `Report.html` (763.6 KB) and `Report.Rmd` (778 B).
- Code editor:** Contains R code for plotting a heatmap:
 

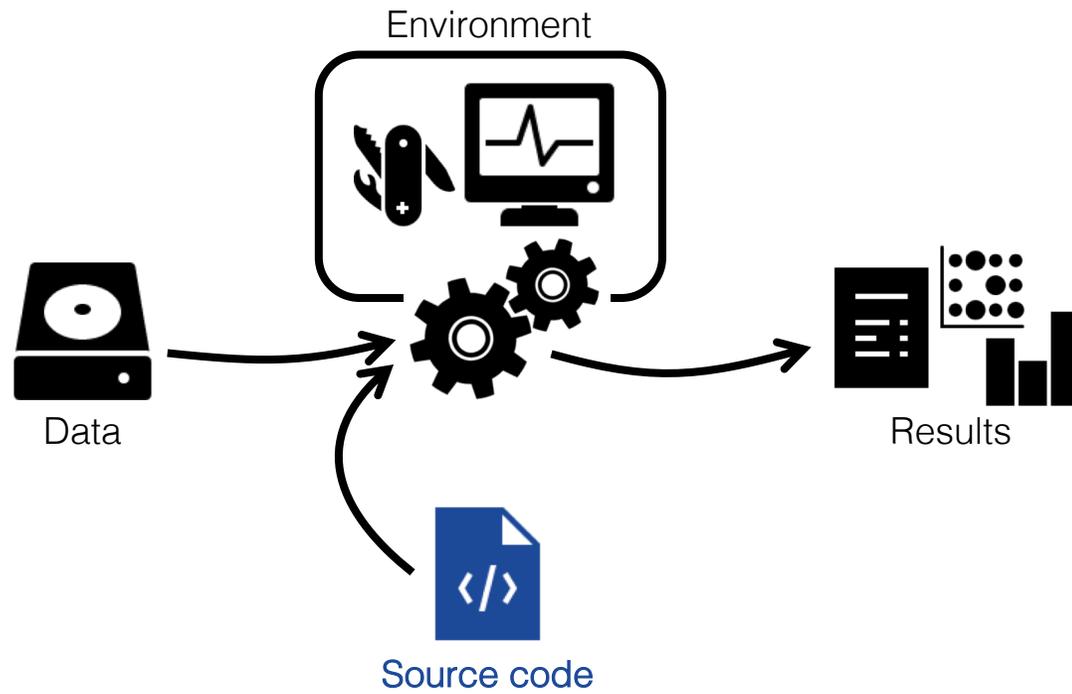
```

26 sample_names <- t
27 colnames(counts)
28 sample_names))
29
30
31 # Plot sample correlation
32
33 Next, we will plot
34
35 ```{r, fig.height=3, fig.width=5}
36 heatmap(cor((counts[,-1])), show_colnames=F)
37
38
39:1 Plot sample correlation

```
- Context menu:** Opened over the code, showing options: Knit to HTML, Knit to PDF, Knit to Word, Knit with Parameters..., Knit Directory, and Clear Knitr Cache...
- Heatmap plot:** A heatmap showing correlation between samples. The color scale ranges from 0.6 (blue) to 1.0 (red). The y-axis labels are SRR935092, SRR935090, and SRR935091.

- Evaluate inline
- Render from menu
- Render from R console or terminal

```
$ R -e "rmarkdown::render('Report.Rmd')"
```



```
project
|- doc/
|
|- data/
|   |- raw_external/
|   |- raw_internal/
|   |- meta/
|
|- code/
|- notebooks/
|
|- intermediate/
|- scratch/
|- logs/
|
|- results/
|   |- figures/
|   |- tables/
|   |- reports/
|
|- Snakefile
|- config.yml
|- environment.yml
|- Dockerfile
```