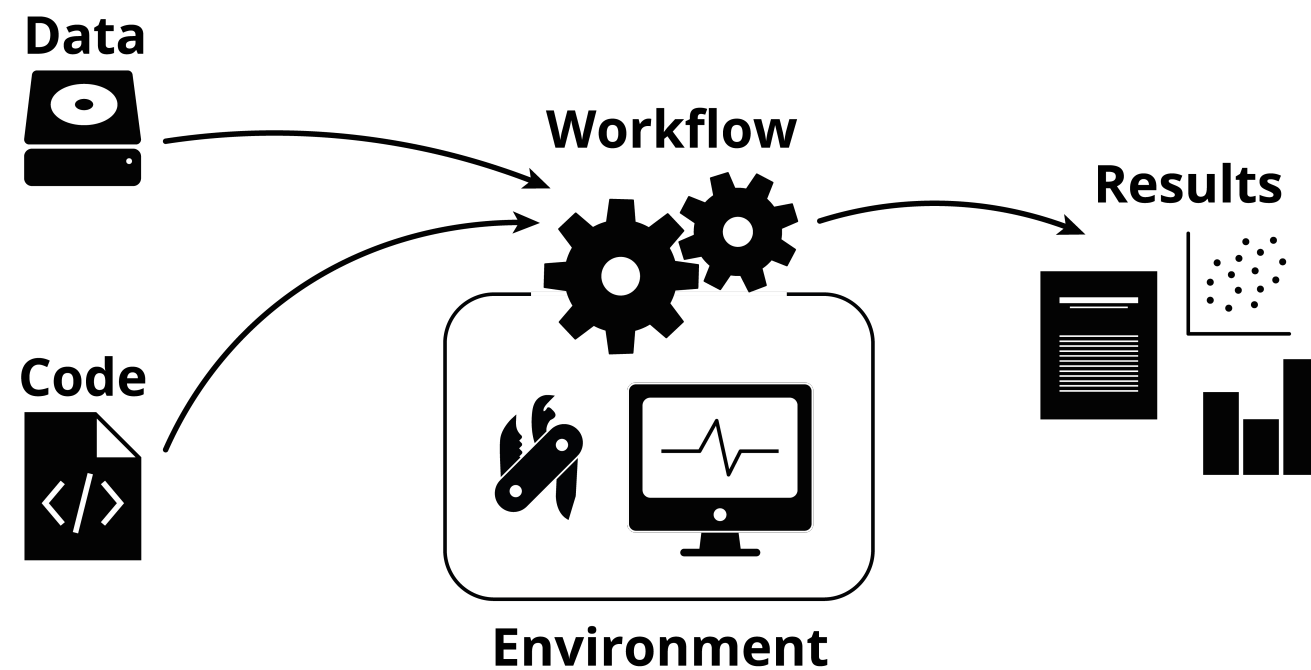Making reports with

R Markdown connects code with results

R Markdown connects code with results

Supplementary material

John Doe, Joan Doe, Jan Doh, Dyon Do
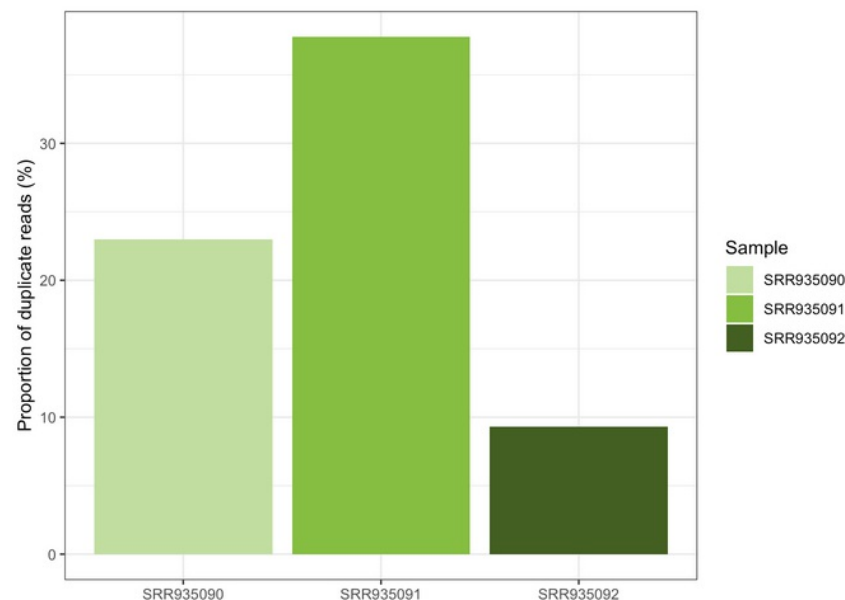
08 April, 2020

Load the data

We have three different samples and several quality metrics for each, originating in publicly available data.

```r
data <- read.table("rmarkdown/intermediate/multiqc_general_stats.txt",
                   sep             = "\t",
                   header          = TRUE,
                   stringsAsFactors = FALSE)
```

Plot quality metrics

We will now plot the proportion of read duplications in each of the samples, to see if there are any problems with data quality.

```r
ggplot(data, aes(x    = Sample,
                 y    = FastQC_mqc.generalstats.fastqc.percent_duplicates,
                 fill = Sample)) +
    geom_bar(stat = "identity") +
    theme_bw() +
    labs(x = NULL,
         y = "Proportion of duplicate reads (%)") +
    scale_fill_manual(values = c("#c1dea0", "#85be42", "#425f21"))
```



```
 1  ---
 2  title: "Supplementary material"
 3  author: "John Doe, Joan Doe, Jan Doh, Dyon Do"
 4  date: "`r format(Sys.time(), '%d %B, %Y')`"
 5  output: html_document
 6  ---
 7
 8  ```{r Load packages, include = FALSE}
 9  library("ggplot2")
10  ```
11
12  # Load the data
13
14  We have three different samples and several quality metrics for each,
15  originating in publicly available data.
16
17  ```{r Load data}
18  data <- read.table("rmarkdown/intermediate/multiqc_general_stats.txt",
19                     sep             = "\t",
20                     header          = TRUE,
21                     stringsAsFactors = FALSE)
22  ```
23
24  # Plot quality metrics
25
26  We will now plot the proportion of read duplications in each of the samples, to
27  see if there are any problems with data quality.
28
29  ```{r Plot data}
30  ggplot(data, aes(x    = Sample,
31                   y    = FastQC_mqc.generalstats.fastqc.percent_duplicates,
32                   fill = Sample)) +
33      geom_bar(stat = "identity") +
34      theme_bw() +
35      labs(x = NULL,
36           y = "Proportion of duplicate reads (%)") +
37      scale_fill_manual(values = c("#c1dea0", "#85be42", "#425f21"))
38  ```
```

R Markdown connects code with results

A header in YAML-format

- Defines document-wide options
- Specifies the output format
- Can include several parameters

```r
1  ---
2  title: "Supplementary material"
3  author: "John Doe, Joan Doe, Jan Doh, Dyon Do"
4  date: "`r format(Sys.time(), '%d %B, %Y')`"
5  output: html_document
6  ---
7
8  ```{r Load packages, include = FALSE}
9  library("ggplot2")
10 ```
11
12 # Load the data
13
14 We have three different samples and several quality metrics for each,
15 originating in publicly available data.
16
17 ```{r Load data}
18 data <- read.table("rmarkdown/intermediate/multiqc_general_stats.txt",
19                    sep              = "\t",
20                    header           = TRUE,
21                    stringsAsFactors = FALSE)
22 ```
23
24 # Plot quality metrics
25
26 We will now plot the proportion of read duplications in each of the samples, to
27 see if there are any problems with data quality.
28
29 ```{r Plot data}
30 ggplot(data, aes(x    = Sample,
31                  y    = FastQC_mqc.generalstats.fastqc.percent_duplicates,
32                  fill = Sample)) +
33     geom_bar(stat = "identity") +
34     theme_bw() +
35     labs(x = NULL,
36          y = "Proportion of duplicate reads (%)") +
37     scale_fill_manual(values = c("#c1dea0", "#85be42", "#425f21"))
38 ```
```

SciLifeLab

NBIS

R Markdown connects code with results

A header in YAML-format

- Defines document-wide options
- Specifies the output format
- Can include several parameters

Markdown text

- Freely add and format text using standard markdown

```
1  ---
2  title: "Supplementary material"
3  author: "John Doe, Joan Doe, Jan Doh, Dyon Do"
4  date: "`r format(Sys.time(), '%d %B, %Y')`"
5  output: html_document
6  ---
7
8  ```{r Load packages, include = FALSE}
9  library("ggplot2")
10 ```
11
12 # Load the data
13
14 We have three different samples and several quality metrics for each,
15 originating in publicly available data.
16
17 ```{r Load data}
18 data <- read.table("rmarkdown/intermediate/multiqc_general_stats.txt",
19                     sep              = "\t",
20                     header           = TRUE,
21                     stringsAsFactors = FALSE)
22 ```
23
24 # Plot quality metrics
25
26 We will now plot the proportion of read duplications in each of the samples, to
27 see if there are any problems with data quality.
28
29 ```{r Plot data}
30 ggplot(data, aes(x    = Sample,
31                  y    = FastQC_mqc.generalstats.fastqc.percent_duplicates,
32                  fill = Sample)) +
33     geom_bar(stat = "identity") +
34     theme_bw() +
35     labs(x = NULL,
36          y = "Proportion of duplicate reads (%)") +
37     scale_fill_manual(values = c("#c1dea0", "#85be42", "#425f21"))
38 ```
```

SciLifeLab

NBIS

R Markdown connects code with results

A header in YAML-format

- Defines document-wide options
- Specifies the output format
- Can include several parameters

Markdown text

- Freely add and format text using standard markdown

Code chunks

- Evaluate R code and show its output
- Specify global or local chunk options (e.g. figure dimensions)
- Also works with several other languages (e.g. Bash and Python)

```
1  ---
2  title: "Supplementary material"
3  author: "John Doe, Joan Doe, Jan Doh, Dyon Do"
4  date: "`r format(Sys.time(), '%d %B, %Y')`"
5  output: html_document
6  ---
7
8  ```{r Load packages, include = FALSE}
9  library("ggplot2")
10 ```
11
12 # Load the data
13
14 We have three different samples and several quality metrics for each,
15 originating in publicly available data.
16
17 ```{r Load data}
18 data <- read.table("rmarkdown/intermediate/multiqc_general_stats.txt",
19                    sep              = "\t",
20                    header           = TRUE,
21                    stringsAsFactors = FALSE)
22 ```
23
24 # Plot quality metrics
25
26 We will now plot the proportion of read duplications in each of the samples, to
27 see if there are any problems with data quality.
28
29 ```{r Plot data}
30 ggplot(data, aes(x    = Sample,
31                  y    = FastQC_mqc.generalstats.fastqc.percent_duplicates,
32                  fill = Sample)) +
33     geom_bar(stat = "identity") +
34     theme_bw() +
35     labs(x = NULL,
36         y = "Proportion of duplicate reads (%)") +
37     scale_fill_manual(values = c("#c1dea0", "#85be42", "#425f21"))
38 ```
```

SciLifeLab

NBIS

Rendering R Markdown documents

- Render from the command line (or R terminal in RStudio):
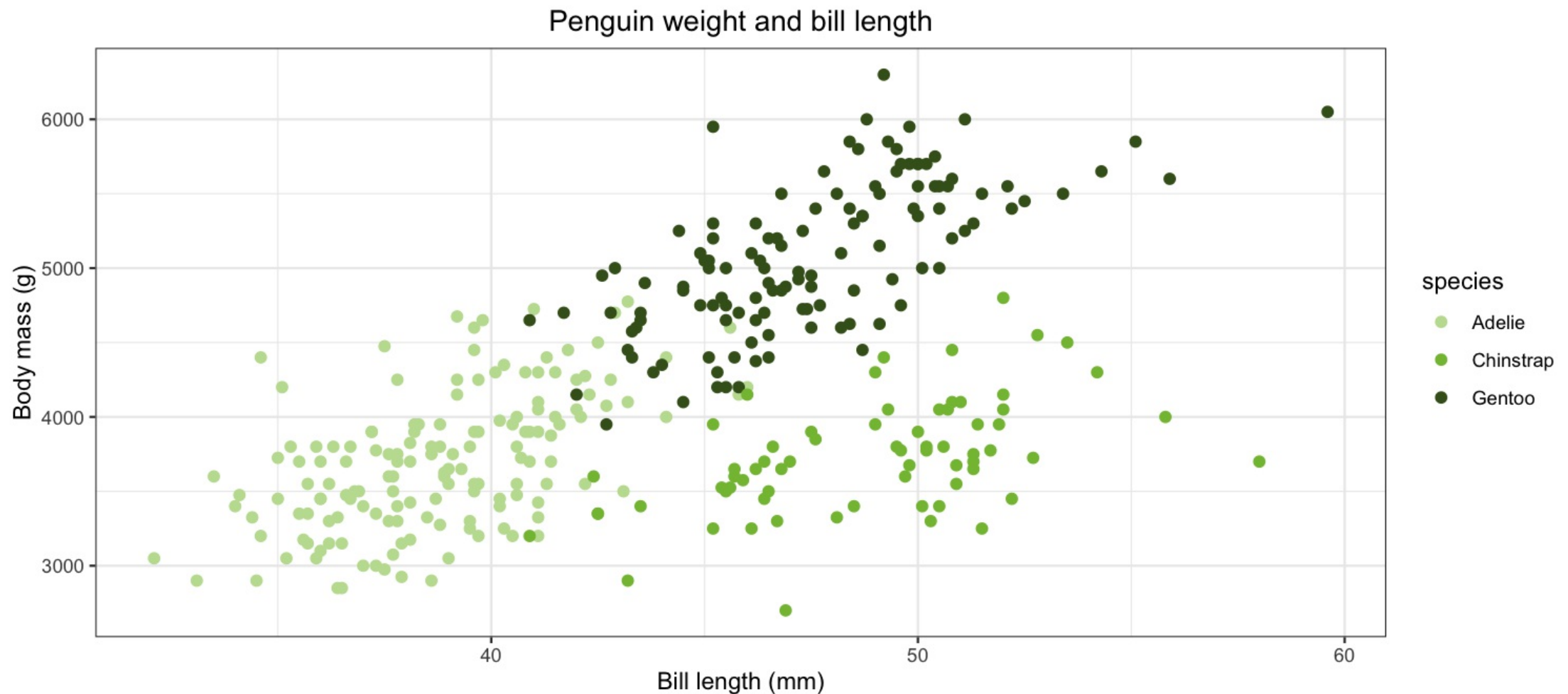
  `R -e "rmarkdown::render('Report.Rmd')"`

- Render from the RStudio menu

SciLifeLab

NB IS

Output formats

- Reports and general documents (HTML, PDF, Microsoft Word)

- Presentations (Xaringan, reveal.js, PowerPoint, Beamer)

- Interactive documents and dashboards (HTML widgets, Shiny)

- Books and websites (Bookdown)

SciLifeLab

NBIS

Presentations with R Markdown

```r
library("ggplot2")
library("palmerpenguins")
data(penguins, package = "palmerpenguins")
ggplot(penguins, aes(x     = bill_length_mm,
                     y     = body_mass_g,
                     colour = species)) +
  geom_point(size = 2) +
  theme_bw() +
  labs(x = "Bill length (mm)",
       y = "Body mass (g)") +
  ggtitle("Penguin weight and bill length") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_colour_manual(values = c("#c1dea0", "#85be42", "#425f21"))
```


Penguin weight and bill length

Questions?