

ADS2 Problem Set 11: R Markdown

MI Stefan

Semester 1, 2020/21

We expect this problem set to take around an hour to complete. But professors are sometimes wrong!^[citation missing]. If this or future problem sets are too long, please let us know, so we can adjust and plan accordingly.

Examining an R markdown file

Have you ever thought about how we write your practical guides and problem sets? We use R markdown!

For this problem set, we provide the source .Rmd file as well as the pdf. Before you carry on with this problem set, read through the .Rmd file. See if from reading the .Rmd file, you can form an idea of what the .pdf will look like. Are there bits of R markdown syntax that are unfamiliar to you? Can you guess what they mean? Compare to the .pdf to verify your ideas!

A simple data analysis procedure

The laboratory technician in Prof. Xu's lab is in charge of looking after the mice. In order to monitor their health, every month she will record the age and weight of every animal by genotype, and record these in a file. The file is always structured the same, with a total of 5 columns:

- “date” holding the date (month/year)
- “ID” holding each individual animal's identifier (an alphanumeric code of length 5)
- “genotype” holding the animal's genotype (a three-letter code)
- “weight” holding the animal's weight in g
- “age” holding the animal's age in weeks

The file is always called “mouse_report.csv”. An example mouse report (for July 2019) is attached.

Prof. Xu wants to take a look at the data every month, but is very busy, and therefore just wants a short overview. This overview should

- be in pdf format
- be no longer than 2 pages
- contain a nice title/heading including the current date
- contain summary statistics about the overall weight and age of mice
- show boxplots of the mouse weight and age by genotype
- contain as much text as is needed to understand the report (this could either be in the form of section headings or short sentences or figure legends)

Note that Prof. Xu is more interested in the results, and does not need to see what commands were used to generate them.

Create an R markdown file that will allow you to generate the report for Prof. Xu with just one click every month.

Fun with maths

One lovely thing about R markdown is that it includes \LaTeX functionality. \LaTeX is a typesetting system that uses commands to typeset text in a specific way - a bit like markdown, but more complex and therefore more powerful. It is especially useful for typesetting mathematics. Here, we will look at a few examples:

- In general, you can include LaTeX code in R markdown by putting it within special brackets: They open with a backslash followed by an left round bracket. And they close with another backslash followed by a right round bracket. Like this:

```
\( LaTeX stuff \)
```

- Most commands in \LaTeX start with a backslash, followed by the name of the command, followed by the argument(s) of the command (if there are any) in curly brackets. For instance, the following can be used to type text in **bold face**:

```
\( \textbf{some text} \)
```

This is not so exciting, because there are easier ways to get bold faced text in R markdown files. Where \LaTeX shines is mathematics. For instance, the `\frac{}{}` command makes fractions:

```
\(\frac{1}{2}\)
```

will display as

$$\frac{1}{2}$$

The `^` command can be used for superscripts, `_` for subscripts, and there are straight-forward commands for greek letters.

- What does the following command display? (Guess first, then try it!)

```
\( A = \frac{d^2 \pi}{4} \)
```

- See if you can typeset the following formulas. (If you need help with how to draw a specific symbol in \LaTeX , you may look for help on the very useful Detexify Website: (<http://detexify.kirelabs.org/classify.html>))

$$E = mc^2$$

$$K_d = \frac{k_b}{k_f}$$

$$v = \frac{v_{max}[S]}{K_M + [S]}$$

$$\sigma = \sqrt{\text{var}}$$

$$\mu = \frac{1}{n} \sum_{k=1}^n x_k$$