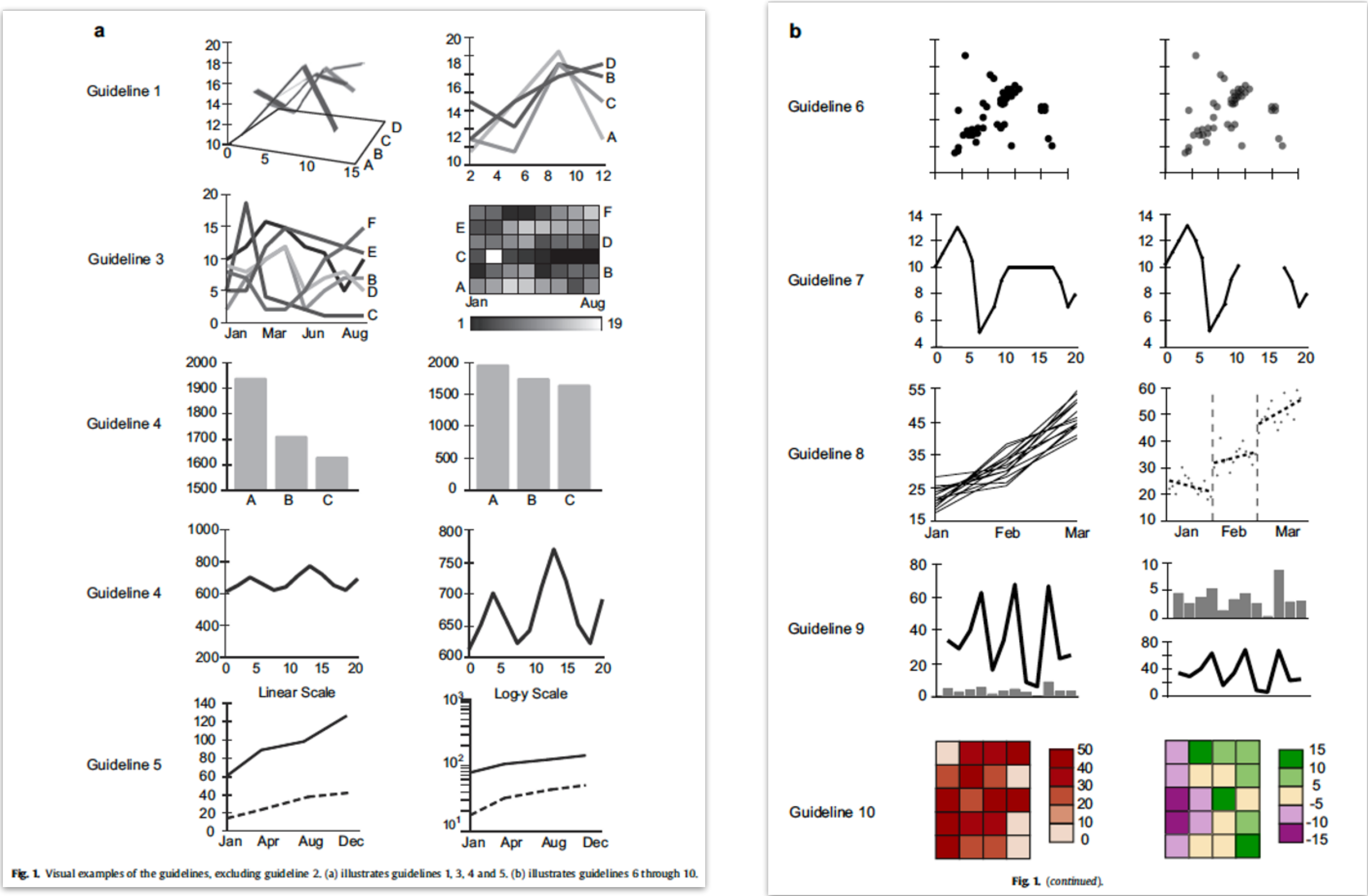


# Visualization 2



**Things that came up ...**

# 10 guidelines for effective data visualization (Keheller & Wagner, 2011)



# Nice interactive visualization

The **React** Graph Gallery

[LEARN](#) [ALL](#) [CHART TYPES](#) [ABOUT](#) [SUBSCRIBE](#)

## Why $R^2$ Alone Fails

---

**$R^2$**  and **correlation** are often seen as definitive measures to validate the relationship between two variables.

This post features an interactive sandbox that explores several edge cases, demonstrating how relying on these summary statistics without visualizing the data can be **dangerously misleading**.

Useful links

[inspiration](#) [d3 gallery](#) [About this chart](#)

<https://www.react-graph-gallery.com/example/scatterplot-r2-playground>

# Homework

# Homework

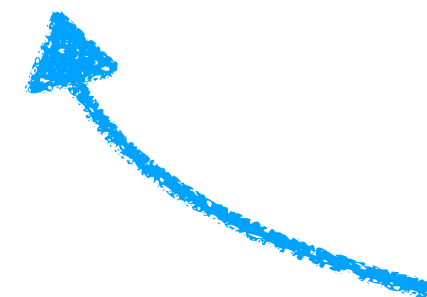
In this homework, **you'll write a short blog post** about a data set. Your goal is to tell us something interesting using a well-crafted, thoughtfully-prepared data graphic.

## Grading Rubric

There are 15 possible points for this homework.

Here are some of the things we care about:

- include all the code that you used to generate the plot (3 points)
- consistent coding style (2 points)
- all the code can be seen in the knitted pdf document (1 point)
- an interesting plot that demonstrates what you've learned in class (4 points)
- a figure caption that is sufficient to understand the plot (2 points)
- a succinct blog post to go with the plot (3 points)

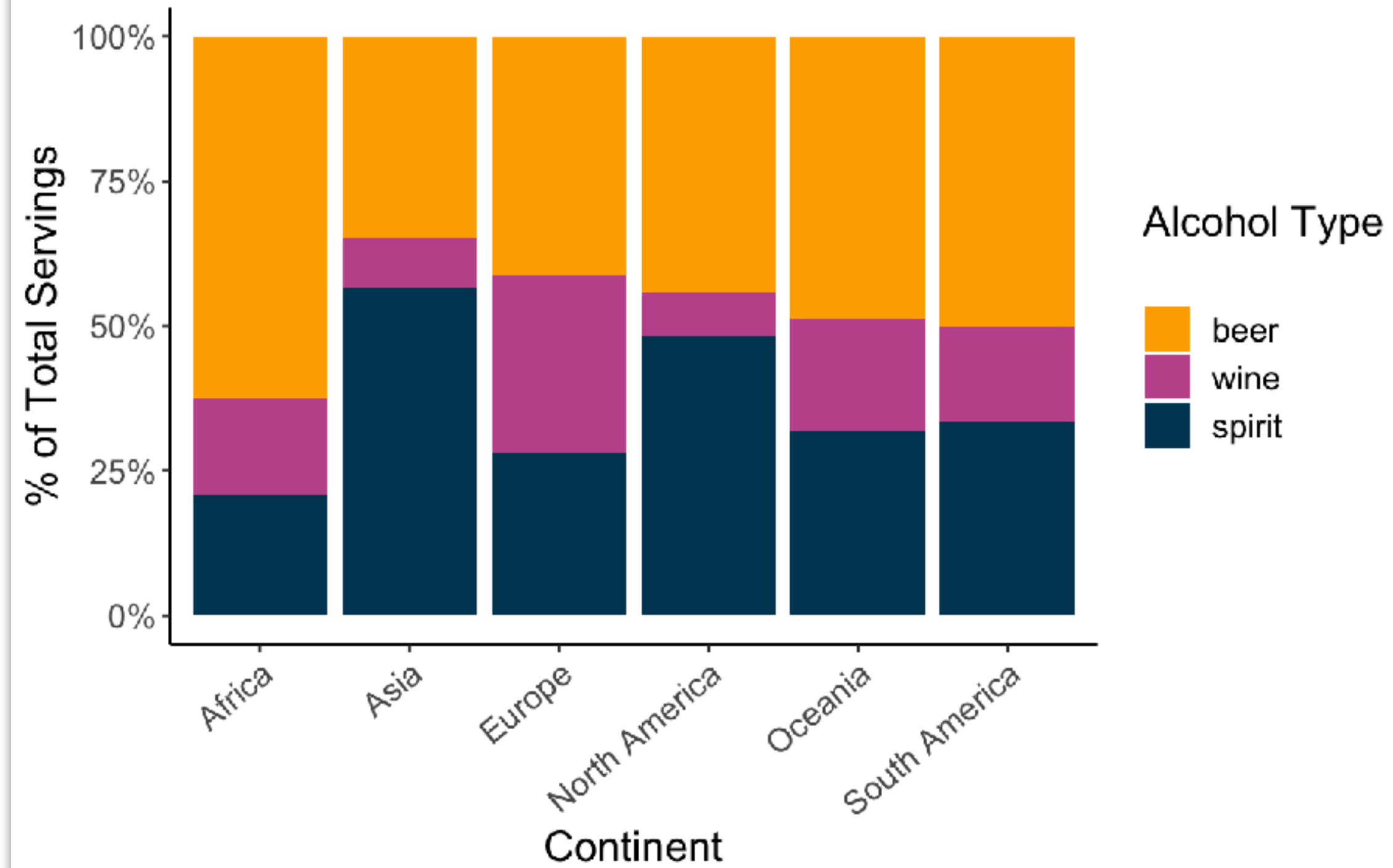


**is at the bottom of the RMarkdown file**

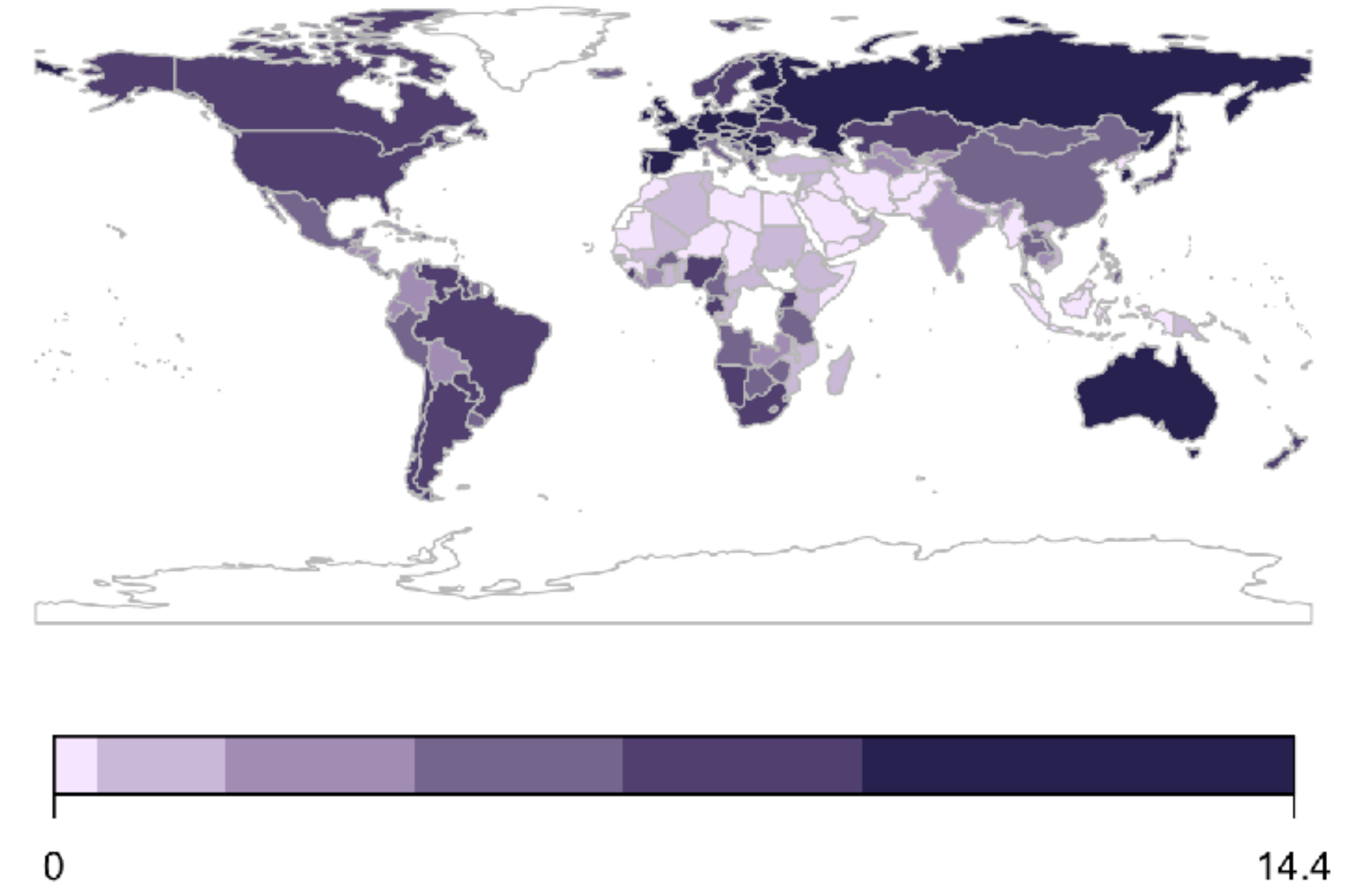


# Homework

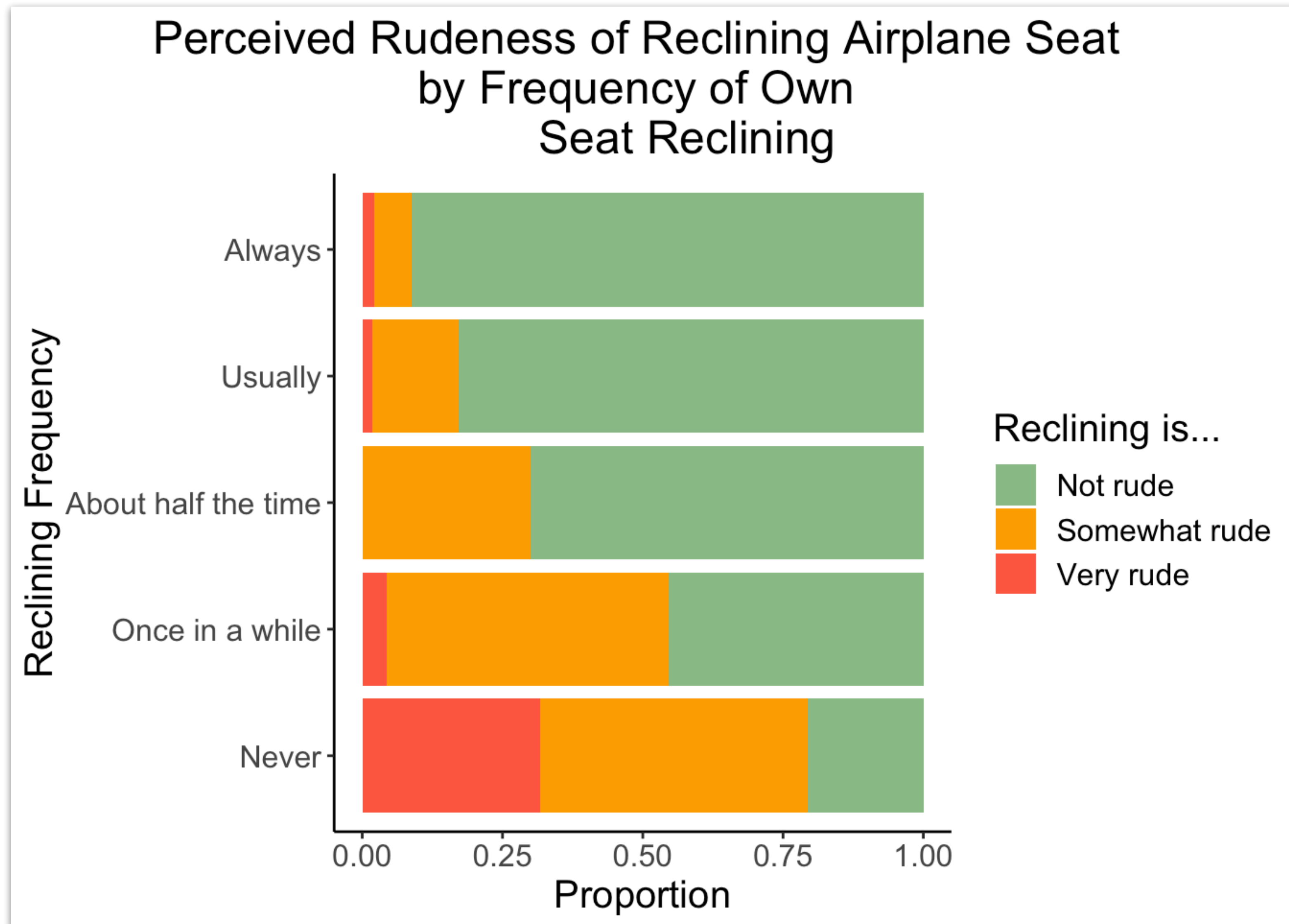
Alcohol Consumed per Person in 2010



Total Liters of Alcohol Consumed per Person in 2010



# Homework



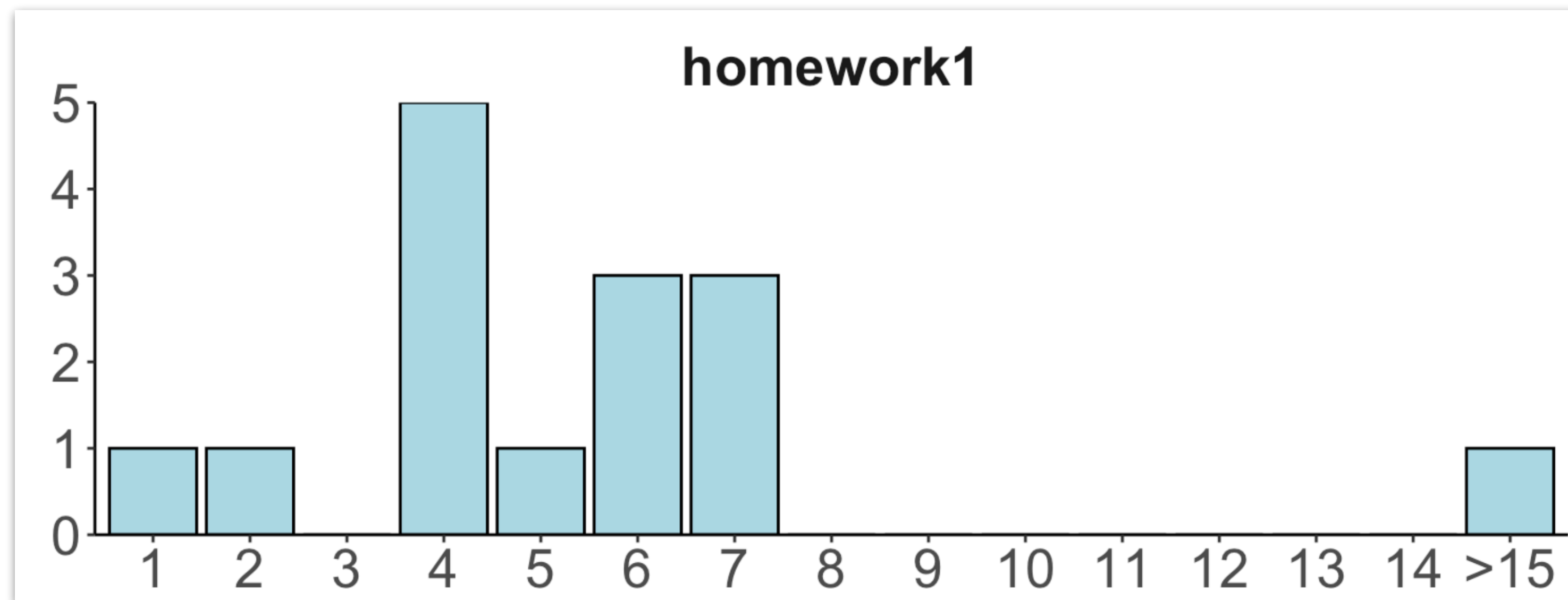


# Homework

Homework is due by **Thursday 15th, 8pm**

Submit **two files**:

- the raw **.Rmd** file
- the rendered **.pdf** file that contains the code as well as the figure.



how long it took people in the past

# Homework

1\_visualization\_homework.pdf (page 3 of 4)

{Your blog post title goes here ...}

**Load packages**

Add the package with the data set that you'd like to load below.

```
library("knitr")
library("tidyverse")
```

**Load the data set**

```
# load the data set here
```

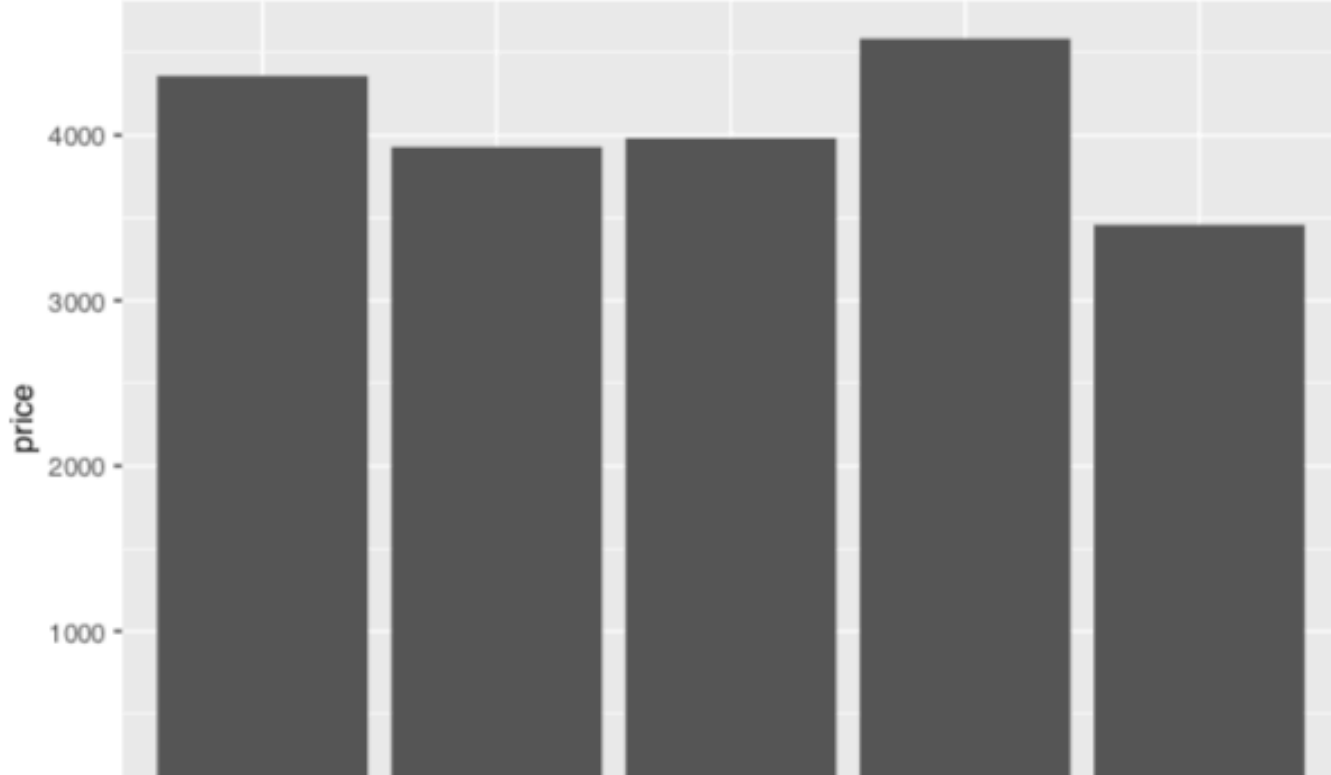
2

**Description**

Write a short text describing the data, and motivating your question here.

**Figure**

```
# replace this figure with an interesting one
ggplot(data = diamonds,
       mapping = aes(x = cut, y = price)) +
  stat_summary(fun.y = "mean", geom = "bar")
```



cut	price
Very Good	4500
Good	3900
Fair	3900
Very Good	4500
Fair	3500

should look sort of like this ...

# Homework

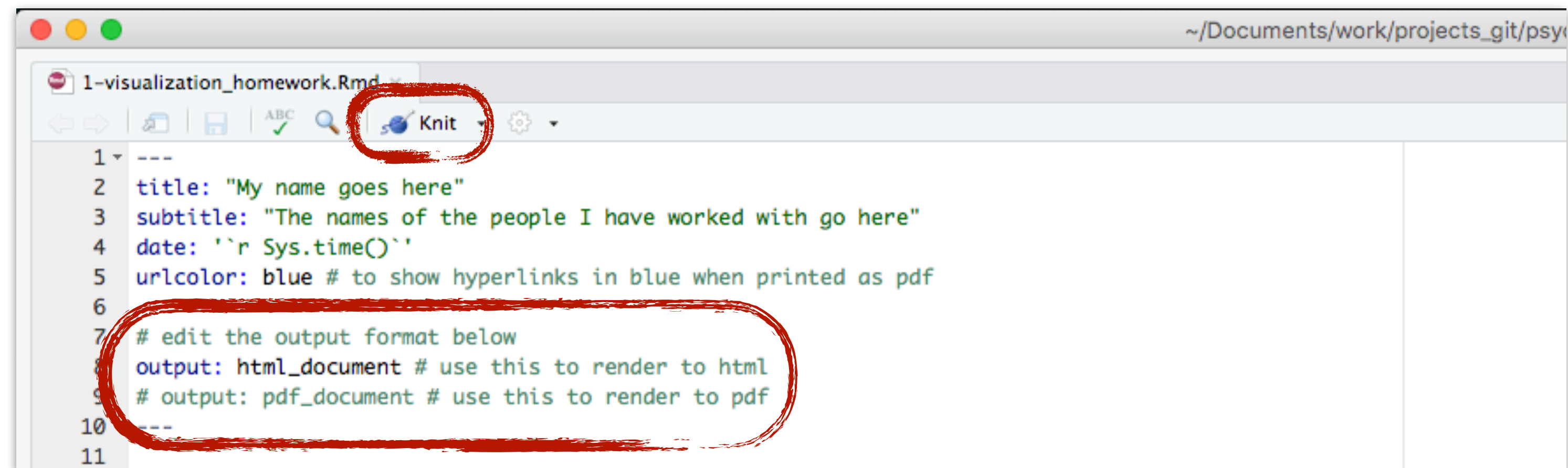
- install `tinytex` (<https://yihui.name/tinytex/r/>)
  - open `1-visualization.Rproj`
  - open `1-visualization_homework.Rmd` within RStudio

```
30 ▾ ### Install tinytex
31
32 In order to knit an RMarkdown document to a pdf file, you have to install LaTeX on your computer. The
33 easiest way of doing so is via the `tinytex` package. Run the code in the following code chunk to do so: |
34 ▾ ```{r, eval=F}
35 install.packages("tinytex")
36 tinytex::install_tinytex()
37
38 # If you experience an error like the following when trying to knit to pdf:
39 # !LaTeX Error: File `xcolor.sty' not found.
40 # then run the following command: tinytex::tlmgr_install("xcolor")
41 # and try to knit again.
42 ```
43
44 You can find out more about the `tinytex` package [here](https://yihui.org/tinytex/).
```

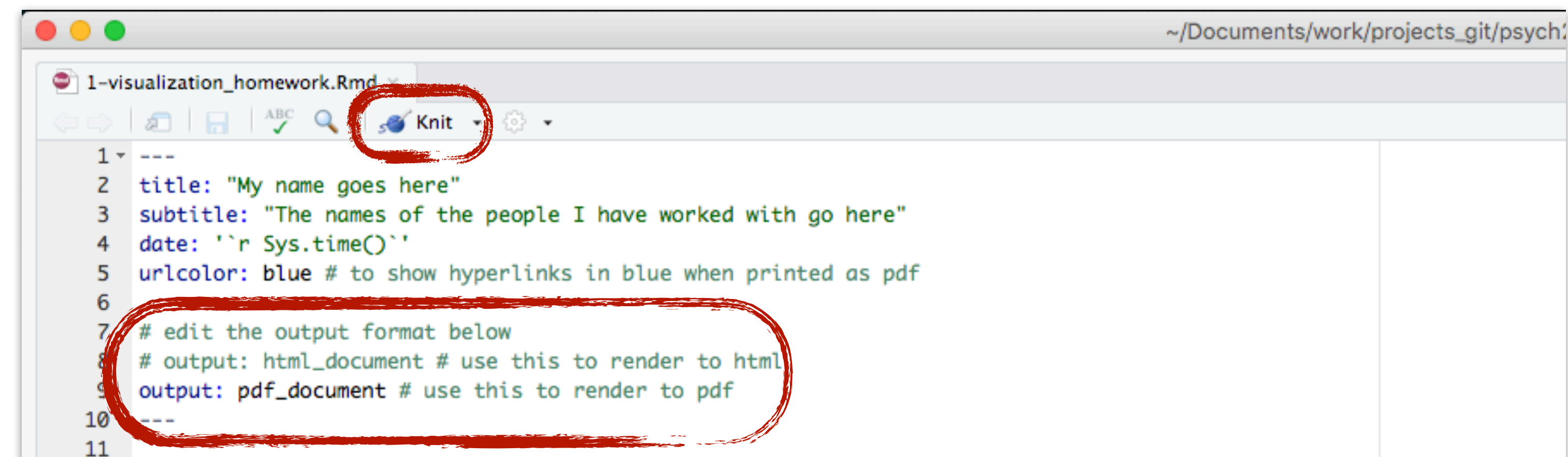
**run this code**

# Homework

- you can change the output format from html to pdf like so ...



```
1 ---
2 title: "My name goes here"
3 subtitle: "The names of the people I have worked with go here"
4 date: ``r Sys.time()``
5 urlcolor: blue # to show hyperlinks in blue when printed as pdf
6
7 # edit the output format below
8 output: html_document # use this to render to html
9 # output: pdf_document # use this to render to pdf
10 ---
11
```

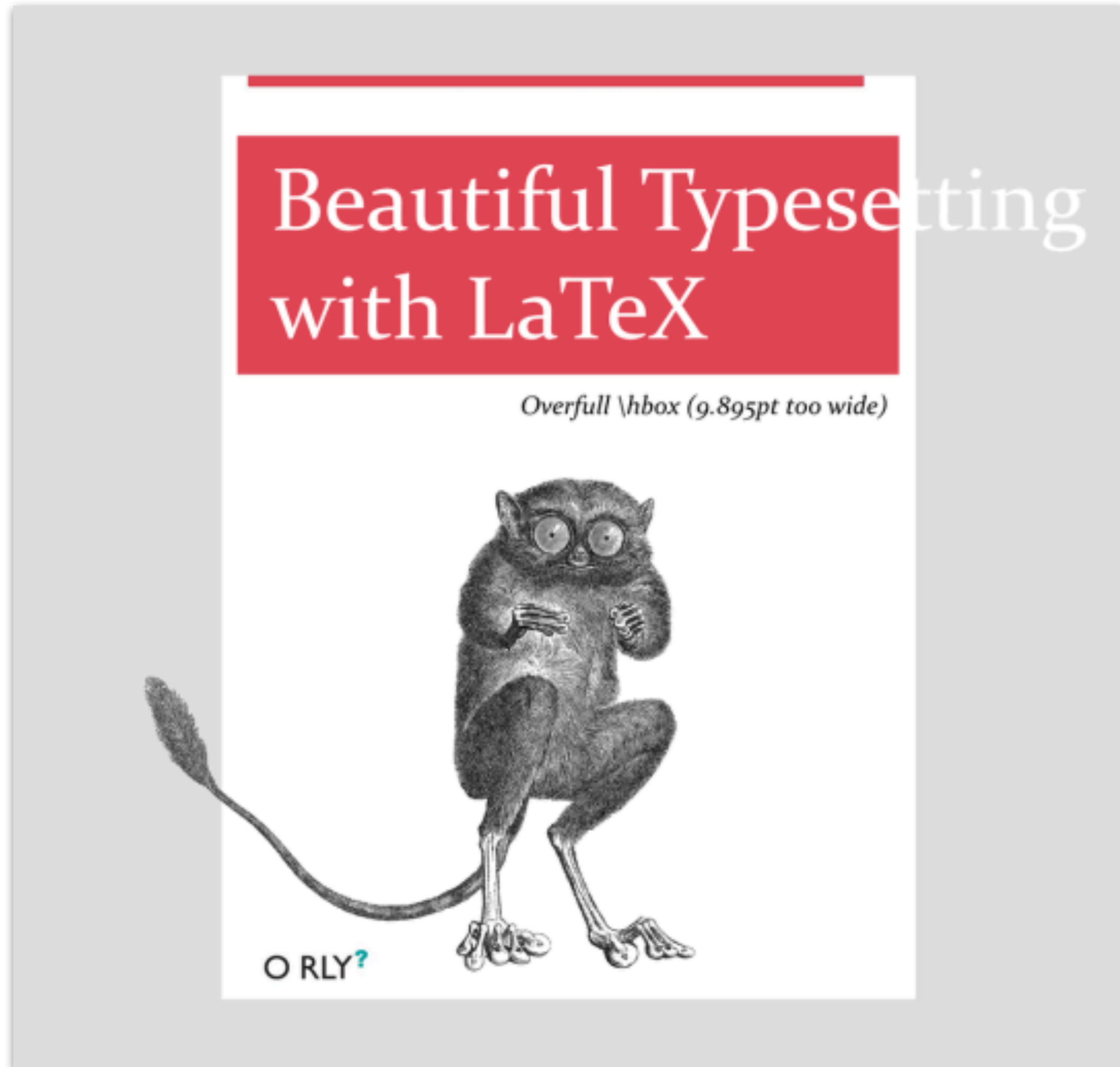


```
1 ---
2 title: "My name goes here"
3 subtitle: "The names of the people I have worked with go here"
4 date: ``r Sys.time()``
5 urlcolor: blue # to show hyperlinks in blue when printed as pdf
6
7 # edit the output format below
8 # output: html_document # use this to render to html
9 output: pdf_document # use this to render to pdf
10 ---
11
```

# Homework

```
1 ggplot(data = df.diamonds, mapping = aes(y = price, x = color, fill = color, group = cut, shape = cut, ...)) +  
2   stat_summary(fun.y = "mean", geom = "bar", color = "black") +  
3   stat_summary(fun.data = "mean_cl_boot", geom = "linerange") +  
4   facet_grid(rows = vars(cut), cols = vars(clarity))
```

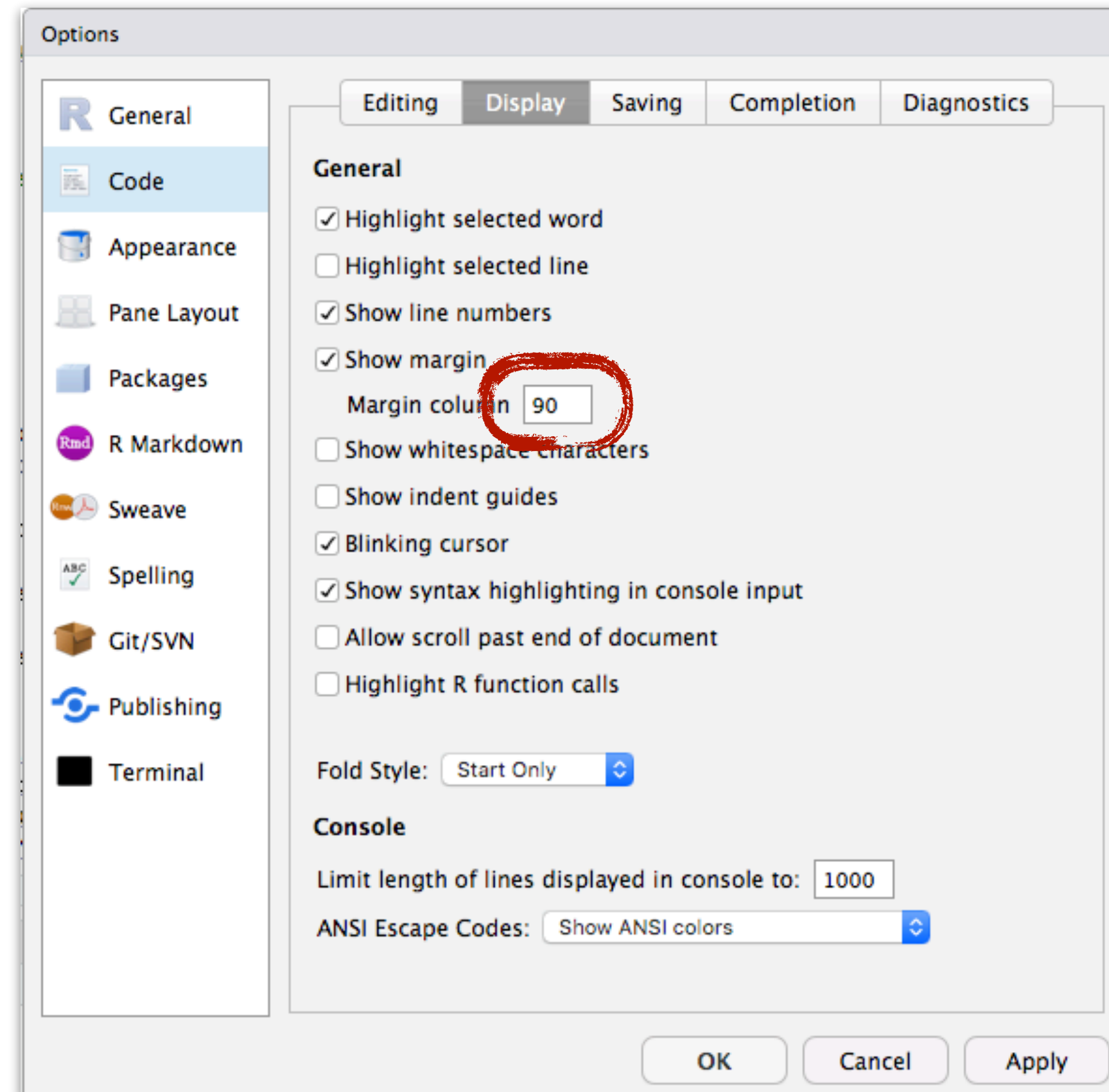
very long code without line break






# Homework

- set the margin to 90 (and make sure not to go over that margin in code blocks)
- Preferences... > Code > Display





# Homework



```
1 ---
2 title: "Class 3"
3 author: "Tobias Gerstenberg"
4 date: "January 11th, 2019"
5 output:
6   bookdown::html_document2:
7     toc: true
8     toc_depth: 4
9     theme: cosmo
10    highlight: tango
11 ---
12 |
13 ```{r setup, include=FALSE}
14 # these options here change the formatting of how comments are rendered
15 knitr::opts_chunk$set(
16   collapse = TRUE,
17   comment = "#>")
18 ```
19
20 # Visualization 2
21
22 In this lecture, we will lift our `ggplot2` skills to the next level!
23
24 ## Learning objectives
25
26 - Deciding what plot is appropriate for what kind of data.
27 - Customizing plots: Take a sad plot and make it better.
28 - Saving plots.
29 - Making figure panels.
30 - Debugging.
31 - Making animations.
32 - Defining snippets.
33
```

margin column

Visualization 2

- Learning objectives
- Install and load pack...
- Overview of different...
- Proportions
  - Stacked bar charts
  - Pie charts
- Comparisons
  - Boxplots
  - Violin plots
  - Joy plots
- Practice plot 1
- Relationships
  - Scatter plots
  - Raster plots
- Temporal data
- Customizing plots
  - Changing the order...
  - Dealing with legends
  - Choosing good colors
  - Customizing themes
- Saving plots
- Creating figure panels
- Peeking behind the ...
- Making animations
- Shiny apps
- Defining snippets
- Additional resources
  - Cheatsheets
  - Data camp courses
  - Books and chapters
  - Misc
- Session info

# Homework

- set the margin to 90 (and make sure not to go over that margin in code blocks)
- Preferences... > Code > Display

```
# take a look at the data sets that come with the package
data(package = "fivethirtyeight")

# take a look at the help file to get more information about the different data sets (not all packages
help("fivethirtyeight")

# the "fivethirtyeight" provides a detailed overview over the different data sets with this command
vignette("fivethirtyeight", package = "fivethirtyeight")

# to load a particular data set (e.g. US_births_2000_2014, replace with the name of the data set you'd
df.data = US_births_2000_2014
```

not good

only important in  
code chunks!

```
# take a look at the data sets that come with the package
data(package = "fivethirtyeight")

# take a look at the help file to get more information about the different data sets (not
# all packages have help files)
help("fivethirtyeight")

# the "fivethirtyeight" provides a detailed overview over the different data sets with
# this command
vignette("fivethirtyeight", package = "fivethirtyeight")

# to load a particular data set (e.g. US_births_2000_2014, replace with the name of the
# data set you'd liked to load) into your environment, run the following
df.data = US_births_2000_2014
```

good!

# Homework

The screenshot shows the Ed Discussion interface for the course 'psych252'. The left sidebar contains a 'COURSES' section with 'psych252' selected and a 'CATEGORIES' section with 'Homework' selected, and 'HW1' highlighted. The main area displays 'No threads' with the prompt 'Be the first to create a thread!'. The 'New Post' form is open, showing options for 'Question', 'Post', and 'Announcement'. The 'Post' option is selected. The 'Category' dropdown is set to 'Homework', and the 'Subcategory' dropdown is set to 'HW1'. The form includes a title field, a rich text editor with various formatting options (bold, italic, underline, link, unlink, list, ordered list, image, video, link, unlink, link, unlink, link, unlink), and a 'Post' button at the bottom right. Below the rich text editor, there are five checkboxes: 'Pinned' (Keep at top of thread list), 'Private' (Visible to you and staff only), 'Anonymous' (Hide your name from students), 'Anonymous Comments' (Allow anonymous comments), and 'Megathread' (Resolvable comments).

post on Ed Discussion if you have any questions  
about the homework

**RStudio & visualization time!**

05:00



**blue**



**pink**



# Anatomy of a nice ggplot

```
1 # ggplot call with global aesthetics
2 ggplot(data = data,
3       mapping = aes(x = cause,
4                     y = effect)) +
5   # add geometric objects (geoms)
6   geom_point() +
7   stat_summary(fun.y = "mean", geom = "point") +
8   ... +
9   # add text objects
10  geom_text() +
11  annotate() +
12  # adjust axes and coordinates
13  scale_x_continuous() +
14  scale_y_continuous() +
15  coord_cartesian() +
16  # define plot title, and axis titles
17  labs(title = "Title",
18       x = "Cause",
19       y = "Effect") +
20  # change global aspects of the plot
21  theme(text = element_text(size = 20),
22        plot.margin = margin(t = 1, b = 1, l = 0.5, r = 0.5, unit = "cm")) +
23  # save the plot
24  ggsave(filename = "super_nice_plot.pdf",
25         width = 8,
26         height = 6)
```

what?

how?

add some text?

"local" adjustments

"global" adjustments

save the beauty!



# Feedback

# How was the pace of today's class?

much  
too  
slow

a little  
too  
slow

just  
right

a little  
too  
fast

much  
too  
fast

# How happy were you with today's class overall?



**What did you like about today's class? What could be improved next time?**

**Thank you!**