Troubleshooting / Review

STA 199

Bulletin

- this `ae` is **not due for grade**. No need to push your solutions.
- lab-3 due today at 5:00pm
- exam 1 released today at 5:00pm
- checkout the cheat sheets on the website

Getting started

Clone *your* `ae9-username` repo from the GitHub organization.

Today

By the end of today you will...

- debug code

Data

In this `ae` we will work within `mpg` data set within the `ggplot2` package.

Notes

Almost none of the code below runs. It's your job to figure out why.

Exercise 1

Let's load the library that contains the data:

```
library(ggplot2)
```

and then explicitly place the data into our environment:

```
data(mpg)
```

and finally glimpse the data:

```
glimpse(mpg)
```

```
Error in glimpse(mpg): could not find function "glimpse"
```

write here what went wrong

Next we'll mutate a column called `2avg` that reports the average of both city and highway miles per gallon.

```
library(dplyr)
mpg |>
  mutate(2avg = (cty + hwy) / 2)
```

```
Error: <text>:3:11: unexpected symbol
2: mpg |>
3:   mutate(2avg
             ^
```

what went wrong?

Next, before we continue further, we want to make sure our document renders to "pdf". Change the `format` to PDF in the YAML at the top of the document and render. What goes wrong?

what went wrong?

We next decide to print the five most fuel efficient highway cars to the screen.

```
mpg2 = mpg |>
  arrange(desc(hwy)) |>
  head(5)
```

```
Error in arrange(mpg, desc(hwy)): could not find function "arrange"
```

what went wrong?

Next, we want to plot `cty` vs `hwy` fuel efficiency. If you run into something you haven't encountered, read the documentation with `?` e.g. `?geom_abline` and scroll to the bottom to find an example of the code in action.

```
mpg = mpg |>
  ggplot(aes(x = cty, y = hwy)) +
  geom_point()

mpg |>
  ggplot(aes(x = cty, y = hwy)) +
  geom_point() +
  geom_abline(y = x)
```

what went wrong?

Debug

Debug the following code chunks.

Exercise 2

```
mpg |>
  ggplot(aes(x = hwy, y = count)) +
  geom_histogram() +
  labs(x = "Highway MPG", title = "Histogram of fuel economy")
```

```
mpg
  ggplot(aes(x = as.factor(cyl), y = cty))
  geom_boxplot() +
  labs(x = "Cylinders", title = "Box plot of city MPG by engine size")
```

- why doe we need `as.factor()` above?

Exercise 3

- When in doubt, comment out! A trick to figuring out what went wrong with your code is to try running it line by line, commenting out the rest (CMD + SHIFT + C : mac), (CTRL + SHIFT + C : windows).

```
mpg |>
  ggplot(aes(x = year, cty))) |>
  geom_boxplot()
  labs(x = Year,
       y = "City MPG" title = "Fuel efficiency by year"} +
    "theme_bw"()
```

```
df = data.frame(x = c(1, 2, NA, 5, 55, 32, 19, 12, 43, 6, 41, 3),
                y = rep(c("A", "B"), 6))

glimpse(df)
```

```
Error in glimpse(df): could not find function "glimpse"
```

```
df |>
  summarize(mean(x))
```

```
Error in summarize(df, mean(x)): could not find function "summarize"
```

YAML errors

The YAML at the top of the document and the **code chunk specific** YAML matters.

YAML below

```
#| label:codeChunk1
```

will result in an error due to incorrect spacing around :

```
ERROR: Validation of YAML cell metadata failed.
```

```
ERROR: Render failed due to invalid YAML.
```

Also, you can't have two code chunks with the same label.

Directories

Where is the data?

R has a host of functions to read in various types of data. From JSON to CSV to XML, there's a function to load it and a package to interact with it.

Most commonly, in this course, we will encounter comma separated values (CSV) files and excel files (e.g. with extension `.xslx`). A couple easy ways to read these in:

- To read `.csv` files, `readr::read_csv()`
- To read `.xlsx` files, `readxl::read_xlsx()`

The notation `x::y()` means that the function `y()` is found within the package `x` and can be loaded with `library(x)`. `readr` is one of the core sub-packages of the `tidyverse`.

Most often, you will open a project and the path to the data will be **relative** to the project folder.

Why can't I push my `.qmd` file to GitHub?

- Are you in the appropriate directory? AKA does the project you have open (upper right) match the quarto file you are working in?
- Did you save the changes in your file? If you haven't changed the file, you can't commit changes.