S&DS 220: Homework 2

Due Friday January 26th, 11:59 pm

Braeden Cullen

Instructions

Complete the questions below. Upload your knitted PDF solutions to Gradescope by January 26th, 11:59PM.

Question 1: Exercise 1.4

In this exercise, you will graph the function $f(p) = p(1-p)$ for $p \in [0, 1]$.

(a) Use seq to create a vector `p` of numbers from 0 to 1 spaced by 0.2.

```
# your code here
#use seq to create a vector `p` of numbers from 0 to 1 spaced by 0.2
seq()
```

```
## [1] 1
```

(b) Use plot to plot `p` in the `x` coordinate and `p(1-p)` in the `y` coordinate. Read the help page for `plot` and experiment with the `type` argument to find a good choice for this graph.

```
# your code here
```

(c) Repeat, but with creating a vector `p` of numbers from 0 to 1 spaced by 0.01.

```
# your code here
```

Question 2: Exercsie 1.7

R has a built-in vector `rivers` which contains the lengths of major North American rivers.

(a) Use `?rivers` to learn about the data set.

```
# your code here
```

(b) Find the mean and standard deviation of the rivers data using the base R functions `mean` and `sd`.

```
# your code here
```

(c) Make a histogram (`hist`) of the rivers data.

```
# your code here
```

(d) Get the five number summary (`summary`) of rivers data.

```
# your code here
```

(e) Find the longest and shortest lengths of rivers in the set.

```
# your code here
```

(f) Make a list of all (lengths of) rivers longer than 1000 miles.

```
# your code here
```

Question 3: Exercise 1.9

There is a built-in data set `state`, which is really seven separate variables with names such as

`state.name`, `state.region`, and `state.area`.

(a) What are the possible regions a state can be in? How many states are in each region?

```
# your code here
```

(b) Which states have area less than 10,000 square miles?

```
# your code here
```

(c) Which state's geographic center is furthest south? (Hint: use `which.min`)

```
# your code here
```

Question 4: Exercise 1.11

Consider the `mtcars` data set.

(a) Convert the am variable to a factor with two levels, auto and manual, by typing the following:

`mtcars$am <- factor(mtcars$am, levels = c(0, 1), labels = c("auto", "manual"))`.

```
# your code here
```

(b) How many cars of each type of transmission are there?

```
# your code here
```

(c) How many cars of each type of transmission have gas mileage estimates greater than 25 mpg?

```
# your code here
```

Question 5: Exercise 1.12

This problem uses the data set `hot_dogs` from the package `fosdata`. See the section called **Libraries** in the Preface of the text under "Software Installation" (page xii).

> *Important: never install a package in an R script or R Markdown document. Always use the console!*

(a) How many observations of how many variables are there? What types are the variables?

```
# your code here
```

(b) What are the three kinds of hot dogs in this data set?

```
# your code here
```

(c) What is the highest sodium content of any hot dog in this data set?

```
# your code here
```

(d) What is the mean calorie content for Beef hot dogs?

```
# your code here
```