ISTA 116: Lab Assignment #1 (50 pts)

Your Name Here

Due by the end of lab on Aug. 30-31 (as appropriate)

1 (Textbook 1.4) Calculation (8 pts)

Use `R` as you would a calculator to find numeric answers to the following (each lettered subpart worth 2 pts):

a. $1 + 2(3 + 4)$

b. $4^3 + 3^{2+1}$

c. $\sqrt{(4 + 3)(2 + 1)}$

d. $(\frac{1+2}{3+4})^2$

2 (Text 1.8) Arithmetic With Vectors (8 pts)

Let our small data set be:

```
2 5 4 10 8
```

a. Enter this data into a data vector `x`.

b. Find the square of each number.

c. Subtract 6 from each number.

d. Subtract 9 from each number and then square the resulting values.

Use a single line of code for each letter (2 pts each).

3 (Text 1.14) Summarizing a Vector (9 pts)

You track your commute times for ten days, recording the following times in minutes:

```
17 16 20 24 22 15 21 15 17 22
```

a. (4 pts) Enter these into `R`. Use the function `max()` to find the longest commute time, the function `mean()` to find the average, and the function `min()` to find the minimum.

b. (5 pts) How many times was your commute 20 minutes or more? What percentage of your commutes are less than 18 minutes long? Use `R` to find these answers (Hint: Extract logical subsets, and use the `length()` function)

4 (Modified from Text 1.25) Accessing Parts of a Data Frame (14 pts)

The data set `nym.2002` in the `UsingR` package contains data about participants in the 2002 New York City Marathon. Use `R` commands to answer the following questions (include your code as well as the answer, where applicable).

a. (2 pts) Load the `UsingR` library into the workspace.

b. (2 pts) How many participants are recorded in this data set (Hint: you can either use `length()` with an individual variable, or use the `nrow()` function on the entire data frame)?

c. (3 pts) Create a variable called `time.hrs` that contains times converted to hours (you can leave the result as a decimal value).

d. (4 pts) Create a new data frame that contains the data for only those runners from New York State (Hint: select a subset of rows using the `home` variable).

e. (3 pts) What percentage of the runners came from within the state?

5 Reading in Data from a File (11 pts)

There is a data set available on d2l called `BrainBodyWeight.csv`, containing brain weights and body weights for various terrestrial (land) mammals.

a. (2 pts) Create a directory somewhere on your computer for `R` data sets in this class (you don't need to use `R` to do this). Download the data file from d2l into that folder, and in `R`, set your working directory there.

b. (2 pts) Read the file into `R` using `read.csv()`, and save it as a data frame (i.e., assign the output of `read.csv()` to a variable). Be sure to use the option `header = TRUE` so that `R` interprets the first row of the data as variable names.

c. (2 pts) Have `R` print out the variable names.

d. (2 pts) Use `attach()` to make the individual variables directly accessible.

e. (2 pts) How much does the heaviest brain weigh?

f. (1 pt) When you're finished, `detach()` the data frame, to keep the workspace clean.