# ISTA 116: Lab Assignment #1 (50 pts)

Kyle Reese Almryde

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 _nd numeric answers to the following (each lettered subpart worth 2 pts):

a. **1 + 2 * ( 3 + 4 ) = 15**

b. **4^3 + 3^( 2 + 1 ) = 91**

c. **sqrt(( 4 + 3 ) * ( 2 + 1 )) = 4.582578**

d. **(( 1 + 2) / (3 + 4))^2 = 0.1836735**

**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.

```
x <- c(2, 5, 4, 10, 8)
```

b. Find the square of each number.

```
x^2 = 4 25 16 100 64
```

c. Subtract 6 from each number.

```
x – 6 = -4  -1  -2  4  2
```

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

```
(x-9)^2 = 49   16   25  1   1
```

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 _nd the longest commute time, the function mean() to _nd the average, and the function min() to _nd the minimum.

**x <- c(17, 16, 20, 24, 22, 15, 21, 15, 17, 22)**

**max(x) = 24**

**mean(x) = 18.9**

**min(x) = 15**

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)

**length(x[x > 19])** = **5** (20, 21, 22, 22, 24)

**length(x[ x < 18])/length(x)** = **0.5** (15 15 16 17 17)

**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.

**Library(UsingR) data(nym.2002)**

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)?

**length(nym.2002$place) = 1000**

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

**time.hrs = with(nym.2002, table(time))**

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.

**read.csv(BrainBodyWeight.csv)**

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 _le from d2l into that folder, and in R, set your working directory there.

**setwd("/Users/kylealmryde/Desktop/Rstuff")**

b. (2 pts) Read the _le 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 1rst row of the data as variable names.

**read.csv("BrainBodyWeight", header=TRUE)**

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

**names(BrainBodyWeight)**

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

**attach(BrainBodyWeight)**

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

**.14 grams**

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

**detach(BrainBodyWeight)**