**Introduction to R - assignment**

**Due**:

**Task**: upload this document in the assignment on …

**Points**:

**Objective**: In this exercise you will work through a tutorial in R. The goal of this exercise is to familiarize you with the basics of R and R Studio as well as introductory data manipulation and visualization.

**Instructions:**

The tutorial contains 19 exercises. You can begin reading in the tutorial and refer back to this sheet when you are ready to answer a question or complete one of the exercises.

Open the R script and start reading.

**#Opening a script and saving**

Exercise 1 (~Line 24):

Question 1: What do you notice about the initial appearances of the R script vs. the R notebook?

Answer:

Question 2: What is different between the two file types with respect to where the figure appears?

Answer:

**#Installing a package**

Exercise 2 (~Line 53)

Question 1: What can you do with this package?

Answer:

**#R as a calculator**

Exercise 3: (~Line 79)

Question 1: Using R find the answer to "(9\*7)/15 + 6". Assign that to a value called "answer1" and multiple it by 15. What is your final product?

Answer:

Copy the code you used to do this from R into this Word Document here:

**#Familiarizing yourself with your working environment.**

Exercise 4 (~Line 105)

Code used to change your working directory to your course (e.g. MOLB4540) folder on your computer and display what your current working directory is:

**#Loading data**

Exercise 5 (~Line 128)

Question 1: What is the difference between the 2 codes and what does it do?

Answer:

Question 2: What happened when you forget to include the "" around the file name?

Answer:

Exercise 6 (~Line 135)

Error message from code in lines 136-138 and your code if you were able to get the .csv read in successfully.

Answer:

Correct code to read in the Pima.csv file using the entire pathway to your folder in your Document/course folder: “/Users/….”

Exercise 7 (~Line 136):

Description of "read.csv" in help window:

Answer:

**#Familiarizing yourself with different data types and structures**

Exercise 8 (~Line 167):

Question 1: What are some other types of data in R according to your Google search?

Answer:

Exercise 9 (~Line 171):

Question 1: What does the structure looks like when it is a matrix?

Answer:

**#Creating your own data frame**

Exercise 10 (~Line 215):

Create your own data frame. Copy code below.

Answer:

**#Working with data**

Exercise 11 (~Line 231):

Code to create a dataset with a subset of columns (columns 1-5):

Answer:

First line of the resulting display of the data structure using the 'str' command:

Exercise 12 (~Line 249):

Question 1: What does the code “subsetPima<-Pima[,colnames(Pima) %in% columns\_to\_keep]” do? Describe in your own words.

Answer:

Exercise 13 (~Line 251):

Code that will pull only the first two rows of the columns "age" and "triceps" from the Pima data set and name it subsetPima3:

Answer:

Exercise 14 (~Line 262):

Calculate the mean and standard deviation individually for each of the three columns of "subsetPima2".

Answer:

Exercise 15 (~Line 269):

Code you used to load the data “metadata” into R:

Answer:

Exercise 16 (~Line 296):

Code you used to create a new column "Infestation\_Stage", and add values to this column for each sample based on the sample ID:

Answer:

Exercise 17 (~Line 306):

Code you used to save the metadata file as a new file in your folder on the Desktop:

Answer:

Exercise 18 (~Line 314): Empty and reload of your environment, and add the name of your environment file here:

**#Final exercise:**

Paste the code you used to make the graph as well as the graph itself here: