E1: Collecting and Entering Data into R

This exploration group assignment will cover the following:

1. How do you collect data and enter it into a spreadsheet program in a format that can be inputted into R easily.
2. Types of data
3. How to import data into a spreadsheet that can be imported ino R

The goal of this lab is to teach you how to collect data. This is a valuable skill if you want to conduct your own research, as you will have to collect and enter data correctly before doing any statistical analysis.

At the end of this lab, each person will upload a file for your group. Make sure your file is named with all your names in the file name, like "Hannah_Marcus_Tara.csv". If you use google sheets, make sure you output your data as a csv.

Step 1: collecting data using a spreadsheet program

The best way to enter large amounts of data is to use a spreadsheet program, like Excel or Google sheets. When you enter data into a spreadsheet program, it is very important to make sure your data are entered correctly. As we discussed in class, this means that you should do the following:

1. Your data start at the very top of the spreadsheet in Row 1, Column A. There should be no blank rows or columns.
2. Each observation gets a row
3. Each variable gets a column.

Set up your spreadsheet with the following variables. Note there should be no spaces. When you enter your variable names, you can capitalize them however you want, but note that R is case-sensitive. This means that R will give you an error if you use the wrong case. The variable `name` is different than `Name` or `NaMe`

1. Name: Each entry in this variable wil be your first name, using a capital letter. If two people in your group have the same name, add the last initial with no space (e.g., HannahB)
2. Trial: This will be where you list which trial is happening. There will be 10 trials for each person, from 1 to 10
3. TrialType: This is where you will list whether the trial is an "Attention" trial or a "Distraction" Trial
4. RT: This is where you enter the data value you collect, or the reaction time score.

At this point, you should have a spreadsheet with four columns labeled as below.

Step 2: Collecting your data

We are going to use the following website to collect reaction time data: https://humanbenchmark.com/tests/reactiontime. When you go to this website, I want you to practice a couple times to get the hang of how it works. Basically, you have to click to do the reaction time. After you are ready, then you will do the experiment.

You will do 10 trials for the experiment. For 5 trials, you should try to do your reaction time as fast as you can. Those are called **Attention** trials. For the other 5 trials, you will try to count backward by 2s aloud while doing the test. This is to examine how distractions affect your reaction time. These are **Distraction** trials.

One person should enter your data, or you can set up a shared google sheet to hold your data. The data should be set up like follows. The reaction time should be in milliseconds.

Name	Trial	TrialType	RT
Amber	1	Attention	250
Amber	2	Attention	233
Amber	3	Distraction	240
Jasmine	1	Distraction	260
Jasmine	2	Distraction	255

Make sure you spell the name the same way each time and have the same capitalization. It is important to enter your data correctly, like we discussed in class.

Step 3: Saving your data and uploading to Canvas

At this point, you should make sure your data are formatted. Make sure you have the data formatted correctly and you follow the best practices in data formatting. Ensure that you don't make these mistakes:

1. Make sure each variable has the same type of data and there is no text in the numerical columns
2. Make sure no columns are incomplete
3. Make sure all the column names are correct and on the first row
4. Make sure all the values in the Name column have the same capitalization

After you do this, save the file as a .csv file and then upload the .csv file to Canvas in your submission.

Step 4: Questions to answer

In this part, I want your group to answer the following questions. When you are done with this, show it to me. This lab will become part of your notes.

1. How many variables do you have in your data? How many observations?

2. What type of data is each variable? (categorical, ordinal, interval, or ratio). Why?

3. What was the hardest or most confusing part of this assignment?

4. If you had to collect data for a research project, do you think you could do it based on these instructions? If not, what other notes do you need to take now so you can do that?