R Lab 1: Introduction to Data in R

Team Captain: NAME HERE Facilitator: NAME HERE
Recorder: NAME HERE Resource Manager: NAME HERE

Invalid Date

Remember – Question the Arson

You can load data files into R using functions from different packages. Typically, we will work with data that are stored in a Comma Separated Values file (a .csv). For this format of data, we will use the `read_csv()` function to read in our data.

The data are located in the *data* folder of your project (see bottom right files tab). To load in the data, we need to specify the "path" to the data, or where R needs to look for the data. I'll walk you through what the code below is doing.

- `arson_data` is the name of the object we are storing the data set into. This is the name R knows the data set by.
- The `<-` "assigns" the output of the `read_csv()` function into the `arson` object
- `read_csv()` is a function that reads in our csv data file
- `"data"` is the name of the sub-folder the data live in (look in the bottom right of your RStudio IDE)
- `"arson.csv"` is the name of the data set. This is the name Posit Cloud's data sub-folder knows the data set by.

1. **Write code to preview the top 6 rows of the arson data set.**

   The recorder should type the answer here.

```
# To run the code, hit the green "play" button in the top right of this code chunk.
# You should type code for question 5 below this line.
```

2. **What are the observations in this data set?**

   The recorder should type the answer here.

3. **How many observations are there in the data set?** *Hint: you can use code to find this or look in the top right environment.*

   The recorder should type the answer here.

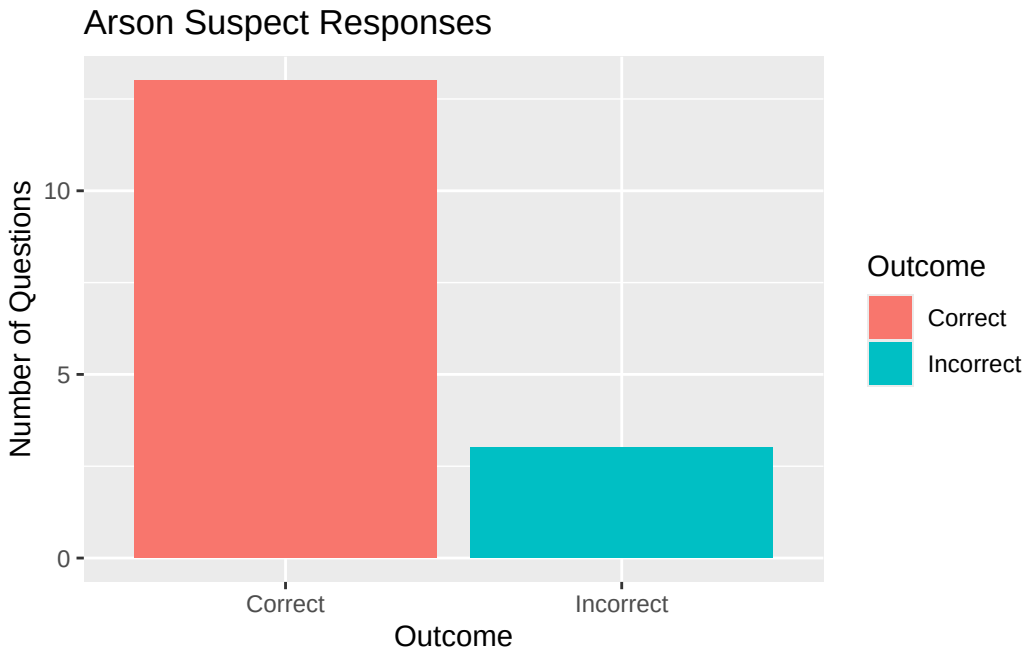
4. **What are the variables in this data set? What are their data types?**

   The recorder should type the answer here.

5. **How many variables are there in the data set?** *Hint: you can use code to find this or look in the top right environment.*

   The recorder should type the answer here.

You can also embed plots in the rendered document. Here is an example of a plot:



Childcare Costs in San Luis Obispo

6. **Load in the "ca_childcare.csv" data set in the following code chunk.**

```
# Write your code below this line.


# You may choose to write more code below this line to help you answer the next couple of
```

7. **What are the observational units in this data set?**

   The recorder should type the answer here.

8. **What are the variables in this data set? Also describe what information the variables contain, don't just list the names of the variables.**

   A description of the data can be found here.

   The recorder should type the answer here.

9. **Suggest one categorical variable that you could measure about the observational units.**

   It is worth it to note that we don't have data for individual households. Carefully consider your observational unit here.

   The recorder should type the answer here.

10. **Could we use this data set to say anything about the cost of childcare in Santa Barbara County? Explain your answer.**

    The recorder should type the answer here.