Lab 1

MTH 3220

7 September 2017

1 Part A: Graphing and Summarizing Data

1.1 Murder was the Case

1. The following are the murder rates (per 100,000 people) for $n = 10$ U.S. counties:

$$15.1, \quad 11.3, \quad 7.8, \quad 10.1, \quad 10.3, \quad 6.8, \quad 3.1, \quad 6.2, \quad 10.7, \quad 13.9$$

   Use the `c()` function to create a *vector* called `murder` containing the above data. Also create another *vector* called `illit` containing the following illiteracy rates for the same counties:

$$2.1, \quad 1.5, \quad 1.8, \quad 1.9, \quad 1.1, \quad 0.7, \quad 1.1, \quad 0.9, \quad 1.3, \quad 2.0$$

2. Check that your objects `murder` and `illit` are indeed *vectors* by typing

```
is.vector(murder)
is.vector(illit)
```

3. Make a histogram of the murder rates using the function `hist()`. The first argument of `hist()` should be the variable `murder`. Choose a "pretty" color for the histogram using the `col=` argument as your second argument in the call to `hist()`. To see a list of your color choices, type

```
colors()
```

4. Make a boxplot of the murder rates using `boxplot()`. Use the argument `col=` to change the color of the boxplot.

5. Use `plot()` to make a scatterplot of `x=illit` and `y=murder`. In the call to `plot()`, do the following to enhance the plot's appearance:

   - Use the arguments `xlim=` and `ylim=` to make the x axis range from 0.5 to 2.3 and the y axis from 0 to 18. Note that these intervals should be supplied as vectors, e.g. `xlim = c(0.5, 2.3)`.

   - Include a title in your plot using the `main=` argument, e.g. `main = "Your Plot Title"`.

   - Include include "nice" x and y axis labels using the `xlab=` and `ylab=` arguments, again supplying the labels between quotes (single or double).

   - Change the plot character using the `pch=` argument. To see a list of your choices, run

     ```
     ?points
     ```

   - Choose a "pretty" color for the points in your plot using the `col=` argument.

6. Compute the mean $\bar{x}$ and standard deviation s_x of the illiteracy rates and the mean $\bar{y}$ and standard deviation s_y of the murder rates using the functions `mean()` and `sd()`.

2 Part B: Hypothesis Testing

2.1 Today was a Good Day

1. The following data were collected in an effort to discover any relationship between marijuana smoking and short term memory (source). To measure short term memory, the digit span task from the Wechsler Adult Intelligence Scale was used. Two groups of ten subjects were tested. One group, the "nonsmokers," claimed not to smoke marijuana. A second group, the "smokers," claimed to smoke marijuana regularly. Enter the data by running the following code:

   ```
   nonsmoke <- c(18, 22, 21, 17, 20, 17, 23, 20, 22, 21)
   smoke <- c(16, 20, 14, 21, 20, 18, 13, 15, 17, 21)
   ```

2. Assuming the data were sampled from populations with normal distributions, use the function `t.test()` to perform a two-sample t-test to see if there is a difference in the task measurements between the smokers and non-smokers at the 0.05 level.