Programming IV

Statistical Programming Assignment

- Instructions
  - Download Analysis 3 Zip Folder
  - Unzip Folder
  - Open Analysis 3 Rmd File
  - Knit to HTML
  - Read Introduction

- Three Part Assignment
  - Using Climate Data from Delhi
  - Found on Kaggle
  - Split Into Train and Test
  - Daily Data Sorted by Date
  - Train: 2013-2016
  - Test: 2017

Statistical Programming Assignment

- Play Close Attention
  - Notice: DO NOT CHANGE
  - Notice: COMPLETE
  - Notice: EVAL=F
- Coding Restrictions
  - You Must Use If-Else When Asked or Lose Points
  - You Must Use Loops When Asked or Lose Points

Part 1
Cleaning Data

- Process of Programming
  - Create Code for a Task
  - Check Code Works
  - Create a Function of the Process to Generalize
- Question 1
  - Taking Working Code and Create a Function
  - Apply Function to Train
  - Apply Function to Test
  - Use Copy and Paste with Some Minor Changes

Part 1
Cleaning Data

- Question 2
  - Change Month from Numbers to Names
  - Create Function to convert a Vector of Months in Number to a Vector of Months in Name
  - Alternative for Variable Creation

```
Data$NEWVAR = CONTENT
```

- Question 3
  - Look at Use of Function
  - Rewrite Over Day Variable with Day Names
  - Data Organized by Date and Starts with a Tuesday
  - Do this For Train and Test

Part 1
Cleaning Data

- Question 4
  - Convert Celsius to Fahrenheit
  - Google Conversion Formula
  - Function that Changes Vector of Celsius Temperatures to Vector of Fahrenheit Temperatures

Part 2
Summarizing Data

- Question 1
  - Use a Double Loop
  - Building a 12x7 Table
  - Start with NA's
  - Imputing the Mean Temperature for Each Month and Day
  - Run This and Think:

```
rownames(MEAN.TEMP.TRAIN)[4]
```

```
colnames(MEAN.TEMP.TRAIN)[6]
```

- Question 2
  - Create Function that Computes Range
  - Do Not Use print() in Function

Part 2
Summarizing Data

- Question 3
  - Apply() Function to Matrix
    - To Rows

```
apply(Matrix, 1, Function)
```

    - To Columns

```
apply(Matrix, 2, Function)
```

- Question 4
  - Loop Through All Observations
  - Start with NULL to Create Vector of Humidity
  - Explain Results from t-Test in the Context of the Problem

Part 3
Time Series Variables

- Question 1
  - Create Percent Change

$$\text{Pressure Percent Change Today} = \frac{\text{Pressure Today} - \text{Pressure Yesterday}}{\text{Pressure Yesterday}}$$

  - Use a Loop To Create
  - Look Back at Original Lag Function from Lecture

- Question 2
  - Create Moving Average

$$\text{Pressure 3-Day Average} = \frac{\text{Pressure Yesterday} + \text{Pressure Today} + \text{Pressure Tomorrow}}{3}$$

  - Similar Process to First Question

Closing

Disperse and Make Reasonable Decisions