Week 3: Exploratory Analysis

From the Introductory Video:

"Exploratory data analysis is a time to investigate suspicions, create simple visuals, and generally play around with the data to better understand it.

Evaluation questions may change slightly, and exploratory data analysis can help guide those changes as they happen. You may want to use Tableau for some/all of exploration as data can easily be manipulated there."

Exploratory Models Already Completed in Week 2:

- **Boxplots-** looking for outliers
- **Correlation Matrix**
- **Scatter/Normal Probability Plots**

Data Exploration by Data Type

***Categorical Data* (quality)**

Pie chart: single categorical variable

Bar graph: single categorical variable

Stacked bar graph: two or more categorical variables

***Continuous Data* (fixed acidity, volatile acidity, citric acid, residual sugars, chlorides, sulphates, free sulfur dioxide, total sulfur dioxide, pH, density, alcohol)**

Histogram: single quantitative variable

Boxplot: single quantitative variable

Scatterplot: two quantitative variables

Line graph: two quantitative variables, one of which is time or distance

Categorical and Continuous Data

Histogram with multiple groups

Side-by-side boxplots: one quantitative variable + one categorical variable with one categorical variable with two or more groups

Tree map: one quantitative + one or more categorical variables

Heat map: one (+) quantitative variable + one or more categorical variables