Alejandro Hohmann

General Assembly Final Project Problem Statements

02-15-2018

**Problem Statement 1:**

* **Dataset:**
  + Option 5: Fuel Economy Data, 2000 - 2018
* **Question(s) seek to answer:** 
  + How does fuel efficiency compare across different car manufacturers (make)?
  + How does fuel efficiency compare across different vehicle models?
  + How does fuel efficiency compare across different vehicle classes?
  + Which manufacturers (makes) have seen the biggest improvement in fuel efficiency?
  + Which models have seen the biggest improvements in fuel efficiency?
  + Which variables have the biggest impact on fuel efficiency?
* **Possible challenges:**
  + While the brand (make) of a vehicle may stay the same, changes to the engine can lead to larger changes in fuel efficiency that history would dictate. This is especially true as manufacturers are moving towards all-electric vehicles. The predictions from this dataset could be wildly inaccurate and need to be converted to miles per gallon equivalent (MPG-e) where applicable.
  + MPG can be difficult to predict as manufacturers begin to preparing their lines to be fully autonomous. The impact on fuel efficiency from level-4 and level-5 autonomy will not be reflected in historical data.
  + Recent scandals have exposed manufacturers that have cheated on their fuel efficiency tests. The data from this set may be corrupted and lead to inaccurate conclusions.

**Problem Statement 2:**

* **Dataset:**
  + Option 4: Loan Data
* **Question(s) seek to answer:** 
  + What factors most explain on-time loan repayments?
  + As a bank, what evaluation metric would you opt for to determine if a given load should be given?
  + As a consumer, what evaluation metric would you opt for to determine if a given loan should be given?
  + How does this affect your model's performance?
* **Possible challenges:**

**Problem Statement 3:**

* **Dataset:**
  + Option 2: Health Insurance Marketplace
* **Question(s) seek to answer:** 
  + What is the relationship between a state’s political affiliation and the state’s rates and benefits?
  + How do plan rates and benefits vary across states?
  + How do plan benefits relate to plan rates?
  + How do plan rates vary by age?
  + How do plans vary across insurance network providers?
* **Possible challenges:**