Group 1

Brandon Strong, Emma Doyle, Matthew Lane, Mikhail Mikhailov, Reid Hoffmier

Dr. Zakariya Qawaqneh

CSC 435

Dataset: Used cars Market in Belarus ([kaggle dataset here](https://www.kaggle.com/lepchenkov/usedcarscatalog))

Description: The dataset we have chosen will be utilized in this project to understand what variables have an effect on the price point which cars are sold at. Analyzing cars that are for sale therefore would contribute to gaining understanding on what makes a car being sold get priced and/or sold in the way it does.

Specification: Data from the dataset was collected December 2nd of 2019 from a popular online catalog in Belarus that contained entries parsed from used car ads.

Questions:

* What is the distribution of manufacturers for each region and whether manufacturer has a significant impact on the asking price of a vehicle?
  + Pie Graph
* What is the most popular model by region and whether we can conclude that the popularity of a model has a direct impact on the price of a vehicle? (First 2 questions can be combined)
  + Bar Graph
* What is the average age of each vehicle manufacturer and whether manufacturer has an impact on the impact of production year on the selling price?
  + Bar graph with data grouped and summarized
* What is the average asking price for each region and what impact does a region have on price?
  + Scatter plot
* Does body type make a difference on the asking price for a region?
  + Bar graph
* What is the relationship between odometer and price and whether sellers are more likely to accept exchanges as the odometer value increases (Fill in with whether its exchangeable)?
  + Scatter Plot
* What is the relationship between engine type and engine capacity and what is the impact of this on the selling price?
  + Scatter Plot
* Manufacturer Origin Distribution
  + (Pie graph)
* Distribution of Engine type for each region and by car type?
  + Mosaic Plot
* Distribution of colors to body type.
  + Bar Graph
* Which variable has the largest impact on the selling price of a vehicle what impacts could explain the outliers for this correlation?
  + Mosaic plot, tying together previous data