W203 Lab 2 Research Proposal

Ed Brown, Daphne Lin, Linh Tran, Lisa Wu

2022-07-18

Research Question

Gas prices are currently at an all time high.[1] The Principal Components Consulting Group has been contracted by a big three automotive company to identify key factors contributing to fuel economy in automobile design. Specifically, we are examining combined fuel economy data and engine variables such as displacement, induction type and vehicle size.

Data Source

Our data source is from fueleconomy.gov and utilizes 2018 model year data [2] and a description of the data is located here[3]. We have chosen the outcome variable of comb08U, which is the unrounded combined fuel economy for a vehicle. We are currently evaluating several input variables to operationalize this test. Based on our background research, preliminary exploratory data analysis and statistical testing, we have chosen two primary variables - displ (engine displacement in liters), and startStop (a boolean representing if the engine in equipped with start/stop technology). High displacement and lack of start stop technology should be indicative of low fuel economy and low displacement and start stop technology should indicate higher fuel economy. We have chosen to work with the 2018 model year data only to keep technology consistent across vehicle comparisons.

Unit of Observation

Each row of the 2018 data set represents cross sectional data and is a vehicle available for purchase in that model year. There is one observation per vehicle. For the 2018 model year we have 1348 vehicles.

[1] https://www.forbes.com/sites/dereksaul/2022/06/09/5-milestone-gas-prices-hit-an-all-time-national-high/?sh=1c445b9a654b

[2] https://www.fueleconomy.gov/feg/epadata/vehicles.csv.zip

[3] https://www.fueleconomy.gov/feg/ws/index.shtml#vehicle