## Abstract

Build a linear model and compare how different attributes of a car affect its fuel economy.

## Introduction

[Use a list of references at the end of the paper about the model that I used in my analysis]

## Theoretical Foundation of Model

Linear Regression is a technique to model the relationship between a dependent variable and one or more independent variables.

## Data Set

The data set was obtained from the U.S. Department of Energy’s and Environmental Protection (EPA) agency’s website for fuel economy information, [www.fueleconomy.gov](http://www.fueleconomy.gov).The EPA publishes a report on this website each year that help consumers with their car buying decisions. The goal of the report is to provide a reliable resource for consumers to compare car models based on fuel economy. The report is titled, “Model Year 2017, Fuel Economy Guide”. The report describes how the EPA conducted the fuel economy test for each vehicle and gave general information about improving fuel economy while driving. This document gives a better understanding of how the fuel economy experiment were conducted. The site provides a Microsoft excel file that contains the raw fuel economy test results for the 2017 car models.

The Microsoft excel file contains 2,500 rows of data, each row represents a different car model for the 2017 model year from a car manufacture. Each row includes the car model name, displacement of engine, transmission type, two or four wheel drive, fuel type and fuel economy data for city mpg (miles per gallon), highway mpg and combined mpg. The linear models will be built using each one of these attributes individual or in combination with another attribute.

Cars come in many different shapes a sizes. The different attributes of a car can greatly affect the fuel economy of the car. Generally cars that are bigger and weigh more will have poorer fuel economy. The way a driver drives their car can also have an effect on fuel economy. The “Model Year 2017, Fuel Economy Guide” report claims that aggressive driving (hard accelerating and braking), excessive idling, accelerating and braking in traffic and even under inflated tires can have a negative effect on fuel economy.

There are many different attributes included in the excel file. The most important attributes are the fuel economy estimates for city driving, highway driving and an estimate for both city and highway driving combined. The city mpg estimate represents the fuel economy of a car when driving in city with frequent stops and accelerations. The highway mpg estimate the fuel economy of a car that is driving on a highway with no traffic. The combined mpg fuel economy estimate was calculated by taking the fuel economy of a car when driving 55% of the time city driving and 45% of the time highway driving. When building the linear regression model I will used the combined mpg estimate attribute as the dependent variable. The combined mpg attribute is better to use since it is a combination of the city mpg and highway mpg fuel economy.

The attributes that will be used as independent variables when building the linear models are displacement, number of cylinders, transmission (automatic or manual) and drive (2 wheel drive or 4 wheel drive). I will briefly describe each attribute of the car that will be used as an independent variable. The displacement of the engine is the swept volume of all the pistons inside the cylinder of a car engine in a single stroke of an engine. Engine displacement is measured in liters. For example a Jeep wrangler has an engine displacement of 3.6 liters. Usually the higher the engine displacement, the lower the fuel economy of the car. The cylinder is a part of the engine where the piston travels. Car generally have four, six or eight cylinder engines. More cylinders usually means a lower fuel economy, four cylinder cars usually have the best fuel economy. Engine displacement and number of cylinders are closely related. Having more cylinders in a car engine will raise the displacement of the engine.

Cars can either have a manual or automatic transmission. The transmission of the car controls the amount to power that goes to a car’s wheels when driving. With a manual transmission a driver has to control the amount of power that goes to the wheel by manually switching the gears themselves. A driver who is operating a car with an automatic transmission does not have to control the power to the wheels because the transmission automatically switches gears for them. Cars with manually transmissions tend to have better fuel economy.

The last attribute that will be used as an independent variable when building the linear model is the drivetrain. The engine of a car will either deliver power to two or four wheels of a car. Two wheel drive cars tend to have better fuel economy than 4 wheel drive cars.

## Sample Run

Throughout the semester we have used SAS to analyze data and create a variety of different statistical models. One of the runs of my experiments is outlined below.

## Results and Comments

[Perform a run of displacement and number of cylinders]

## References