Homework 1

Chenxi Li & 22212393
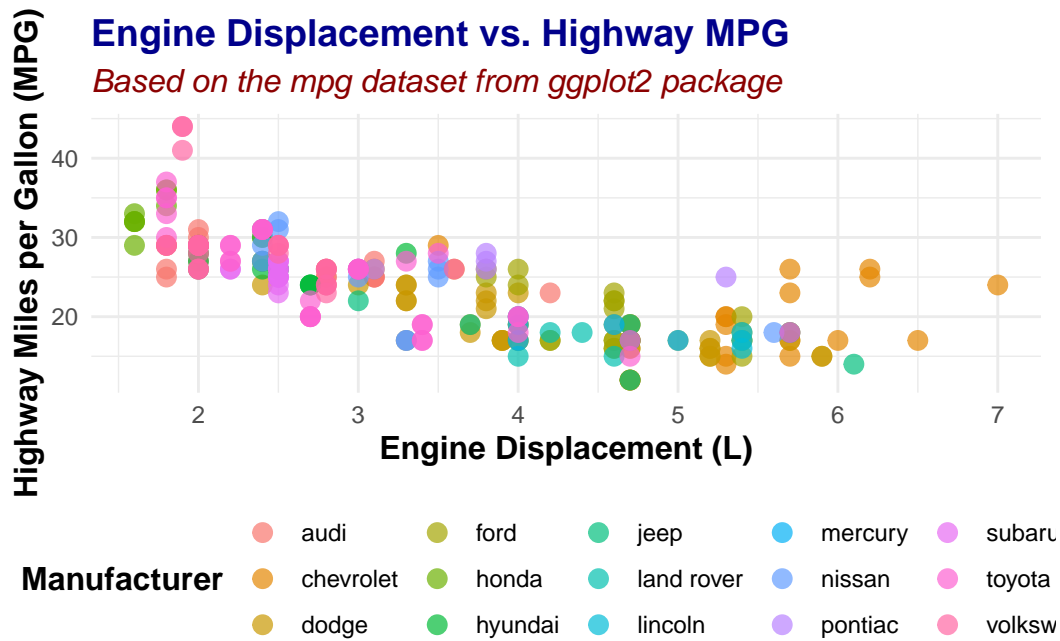
2024-06-05

Contents

Introduction

This document provides a plot based on a dataset from an R package and some descriptive statistics about the data used.

Dataset and Plot

The dataset used in this plot is the `ggplot2` package's `mpg` dataset. The `mpg` dataset contains observations collected by the US Environmental Protection Agency on 38 models of cars. It includes variables such as manufacturer, model, engine displacement, year, number of cylinders, transmission, drive type, and fuel economy (in city and highway driving).

Plot

Below is a scatter plot showing the relationship between engine displacement (in liters) and highway miles per gallon. We've added colors for each manufacturer and labels to make it visually appealing.

Engine Displacement vs. Highway MPG
Based on the mpg dataset from ggplot2 package

Manufacturer: audi, chevrolet, dodge, ford, honda, hyundai, jeep, land rover, lincoln, mercury, nissan, pontiac, subaru, toyota, volksw

Description of the Plot

The scatter plot above shows the relationship between engine displacement and highway miles per gallon. As the engine displacement increases, the highway miles per gallon tend to decrease. This is expected as larger engines typically consume more fuel. Different colors represent different manufacturers, aiding in identifying the distribution of data points for each manufacturer.

Descriptive Statistics

Here are some descriptive statistics of the variables used in the plot:

```
     displ              hwy
 Min.   :1.600    Min.   :12.00
 1st Qu.:2.400    1st Qu.:18.00
 Median :3.300    Median :24.00
 Mean   :3.472    Mean   :23.44
 3rd Qu.:4.600    3rd Qu.:27.00
 Max.   :7.000    Max.   :44.00
```