M6 Project Deliverable 3
Information Visualization Report

Team
Snehita Moturu - G01388464
Pravallika Avula - G01388664
Sai Roopesh Diddi- G01353614

Prof. Charles Lynch

AIT664 - 001: Information: Representation, Processing & Visualization

George Mason University

05 November 2023

Abstract:

The automobile industry has had a huge impact on many countries since its introduction in the United States in 1895. There have been numerous improvements and diverse car models produced over the years. When consumers consider purchasing a car, they typically consider affordability, luxury, and safety.

Hypothesis:

We want to find patterns and relationships that can be used to anticipate a car's performance and fuel efficiency by analyzing these attributes collectively. The year, place of manufacture, and particular model name of an automobile all have an impact on its performance and fuel consumption, which are determined by variables like displacement, mpg, cylinders, horsepower, weight, and acceleration. We hypothesize that increased fuel economy, more horsepower, and less weight are related to newer model years, particular countries of origin, and particular car models.

Data Source:

This dataset consists of data by Ross Quinlan which covers from the year 1970 to 1982. This dataset was taken from UCI’s Machine learning repository.

Attributes:

MPG: The fuel economy of an automobile is the relationship between the distance traveled and the amount of fuel consumed by the vehicle. MPG is continuous from 9 to 48 (Continuous).

Cylinders: A cylinder is the power unit of an engine; it's the chamber where the gasoline is burned and turned into power. Number of cylinder continuous from 3 to 8 (Multi-Valued Discrete)

Displacement: Engine displacement is the measure of the cylinder volume swept by all of the pistons of a piston engine, excluding the combustion chambers. Displacement is continuous from 68 to 455 (Continuous).

Note: This displacement column(cubic inches) has been converted to cubic centimetres using a conversion factor as everyone is familiar with cubic centimeters(cc).

Horsepower: Horsepower is a unit of power used to measure the forcefulness of a vehicle's engine. Horsepower continuous from 46 to 230 (Continuous).

Weight: The weight of an object is related to the amount of force acting on the object, either due to gravity or to a reaction force that holds it in place. The weight of the car continues from 1613 to 5140 (Continuous).

Acceleration: Is the rate of change of velocity of an object with respect to time. It continuous from 8 to 25 (Continuous).

Model Year: Is the year in which a product is manufactured. The year ranges from 1970 to 1982 (Multi-Valued Discrete)

Origin: The Country which manufactures the automobile. (Multi-Valued Discrete)

1: USA
2: EUROPE
3: JAPAN

1. Visualizations

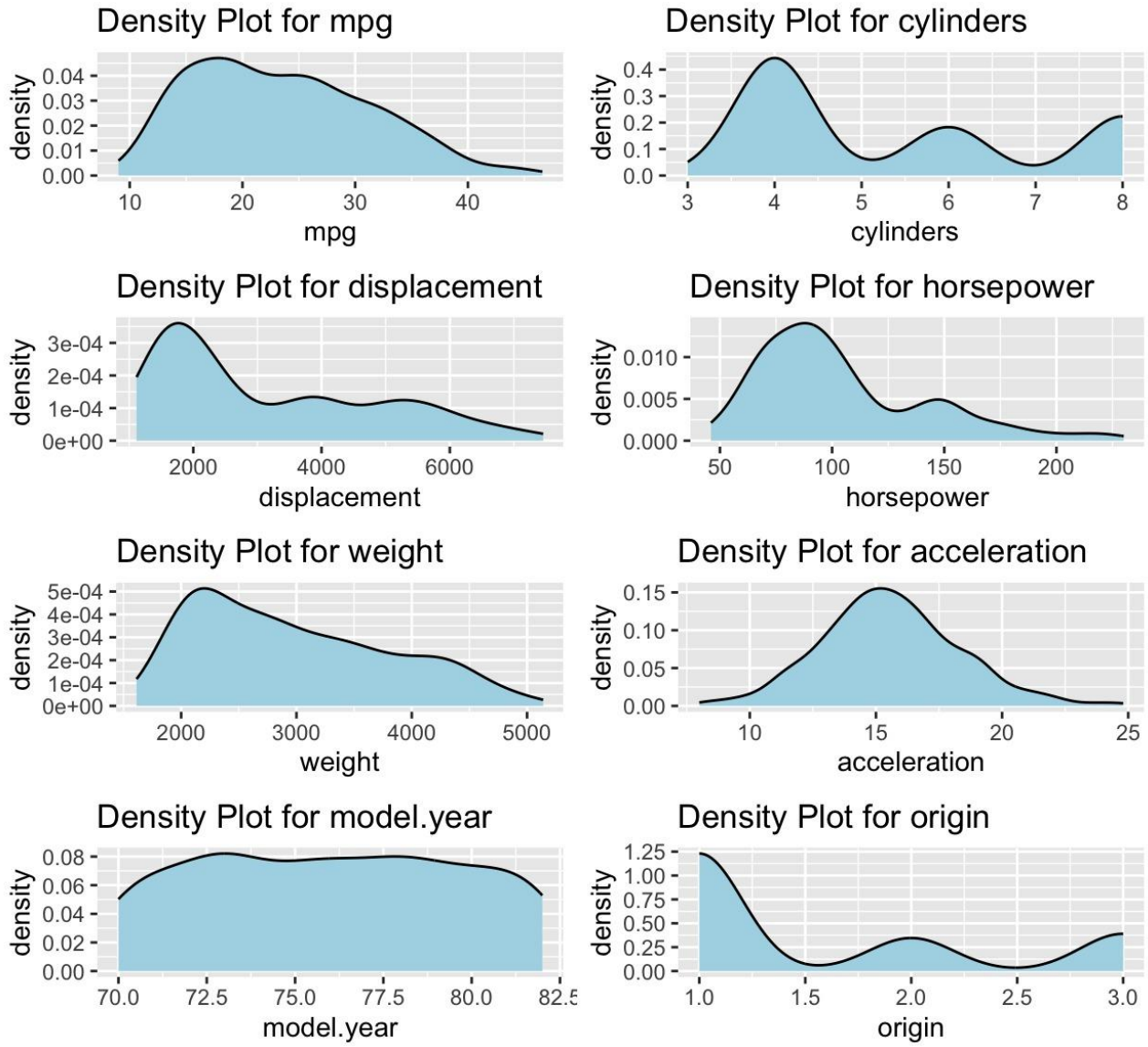

Figure1: The density plot showing the distribution of different car variables.

The mpg density plot reveals that a majority of these cars prefer to sip fuel, with most achieving between 15 and 30 miles per gallon. On the other hand, the displacement density plot suggests that these cars are not shy about packing some muscle, with most boasting a displacement between 100 and 200 cubic inches.

When it comes to horsepower, these cars seem to strike a balance between efficiency and performance, with most falling within the 100 to 150 horsepower range. As for their cylinder count, these cars favor practicality, with most opting for either 4 or 6 cylinders.

Acceleration is another area where these cars show their versatility. The acceleration density plot reveals that most can get from 0 to 60 mph in a respectable 10 to 15 seconds. Finally, the origin density plot suggests that these cars come from diverse backgrounds, with most hailing from either the United States or Japan.

Now, we determine how mpg is changing over time by plotting a scatterplot between mpg and model years. We can see that the mpg has increased from 1970 to 1982.

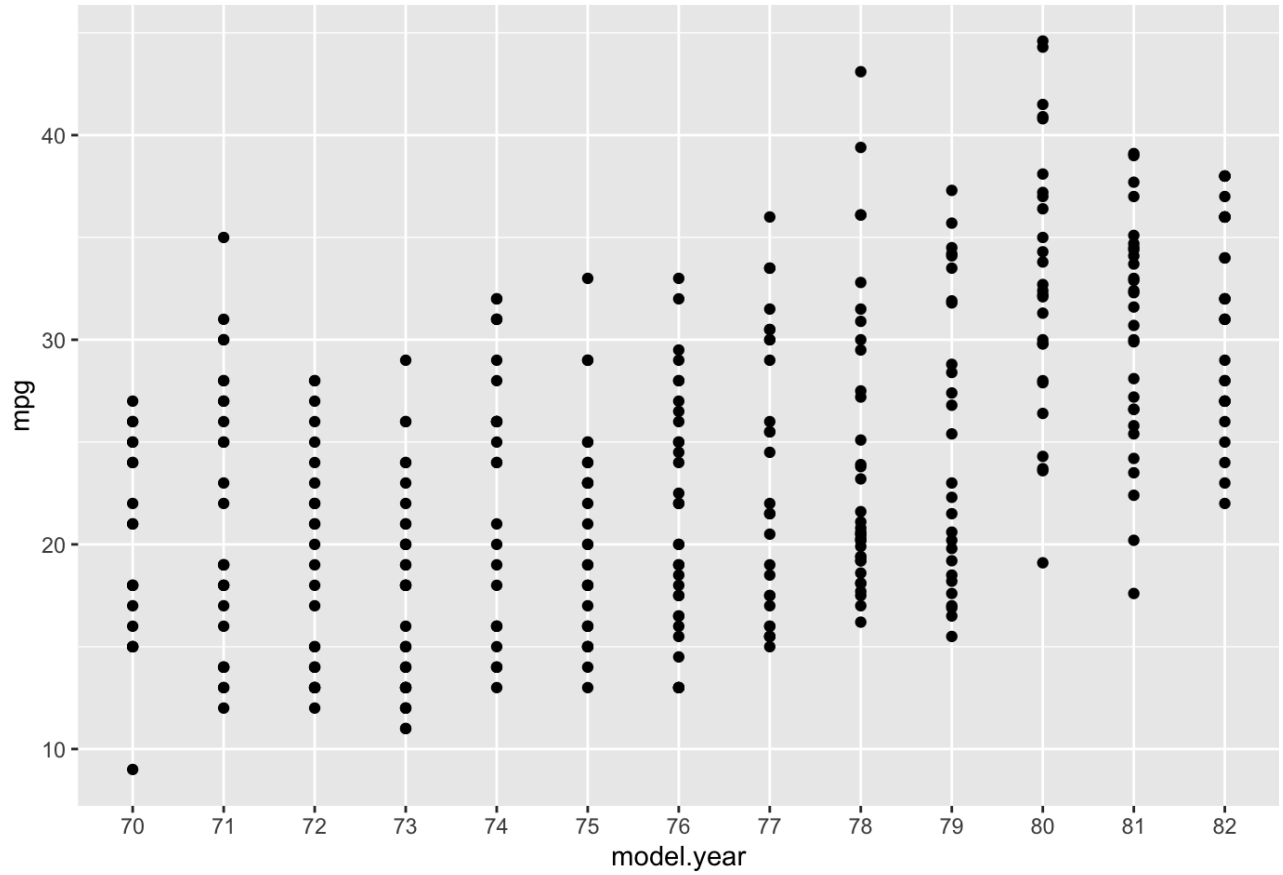

Figure2: Tracking MPG Trends Over Time: A Scatterplot of MPG and Model Years.

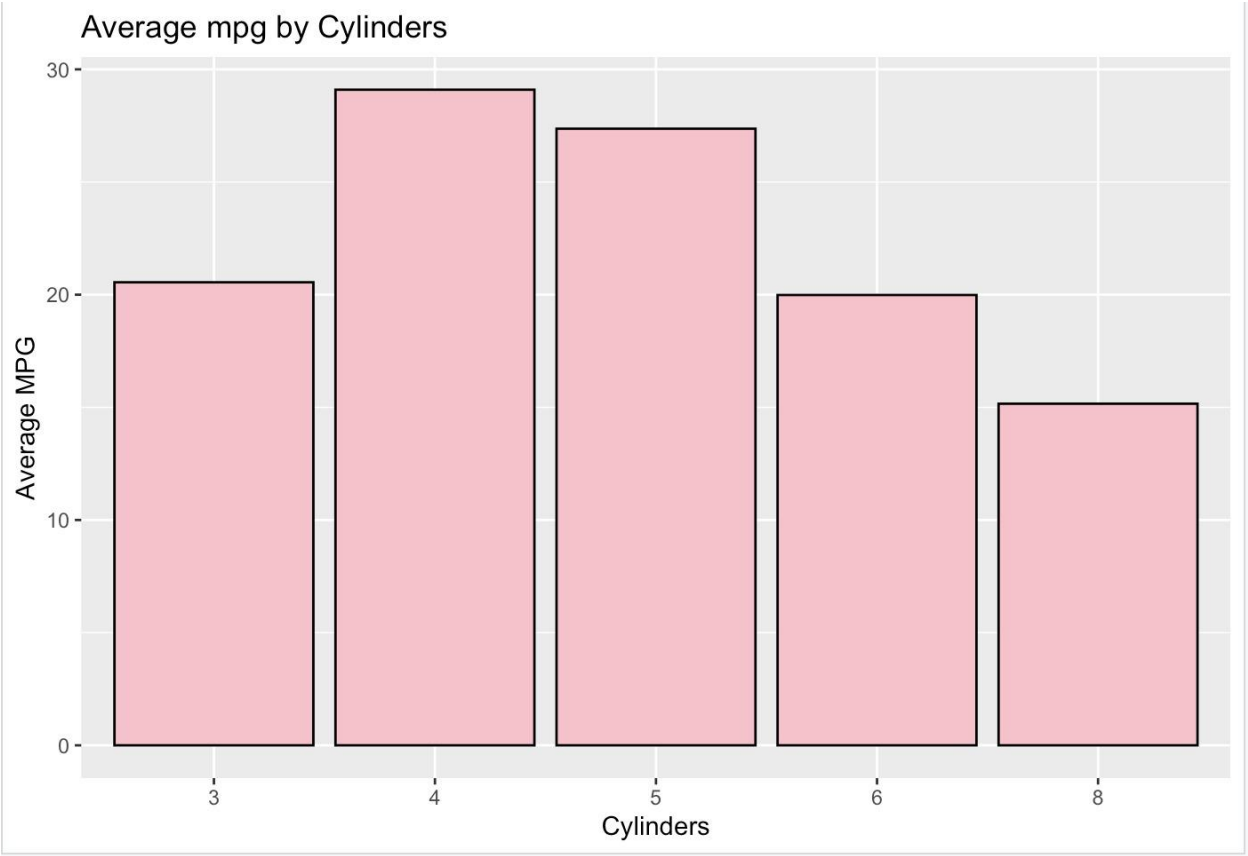

Figure 3: Fuel Efficiency Showdown - Cylinders vs. Miles per Gallon.

The graph reveals that cars with fewer cylinders tend to be more frugal, sipping less fuel per mile. For instance, cars with 4 cylinders average an impressive 28.5 miles per gallon, while their 8-cylinder counterparts manage only 19.5 miles per gallon. This difference stems from the size of the engines under the hood. Cars with more cylinders often pack larger engines, which naturally consume more fuel. Additionally, these cars often carry more weight, further diminishing their fuel efficiency.

However, there are a few exceptions to this trend. For example, cars with 3 cylinders average 24 miles per gallon, falling behind their 4-cylinder counterparts. This is likely due to the relative newness of 3-cylinder engines, which are still undergoing development. In general, the graph highlights that cars with fewer cylinders tend to be more fuel-efficient companions. This is a crucial factor for car buyers seeking vehicles that make the most of every drop of fuel.

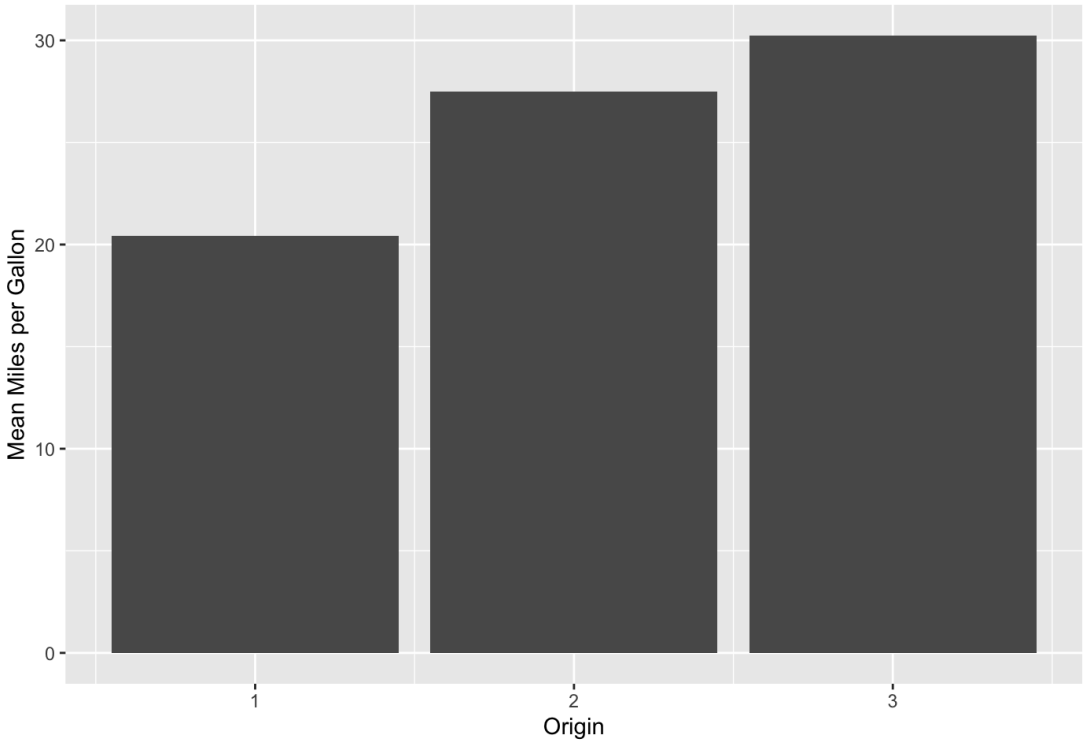

Figure4: MPG Vs Origin

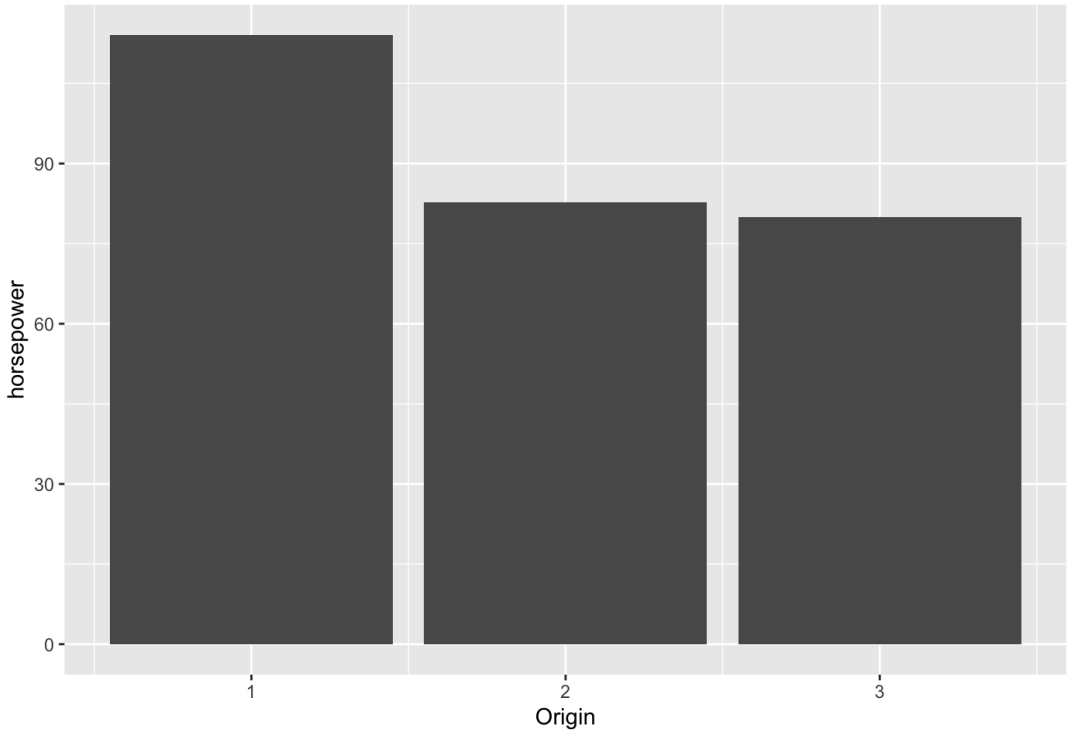

Figure5: Horsepower Vs Origin

It is observed that the mpg is less and horsepower is more for USA-originated cars. Whereas, mpg is higher and horsepower is lesser for Japanese-originated cars.
Below is the scatterplot between weight and mpg to determine the relationship between the both. We can observe that the mpg is decreases as the weight increases.

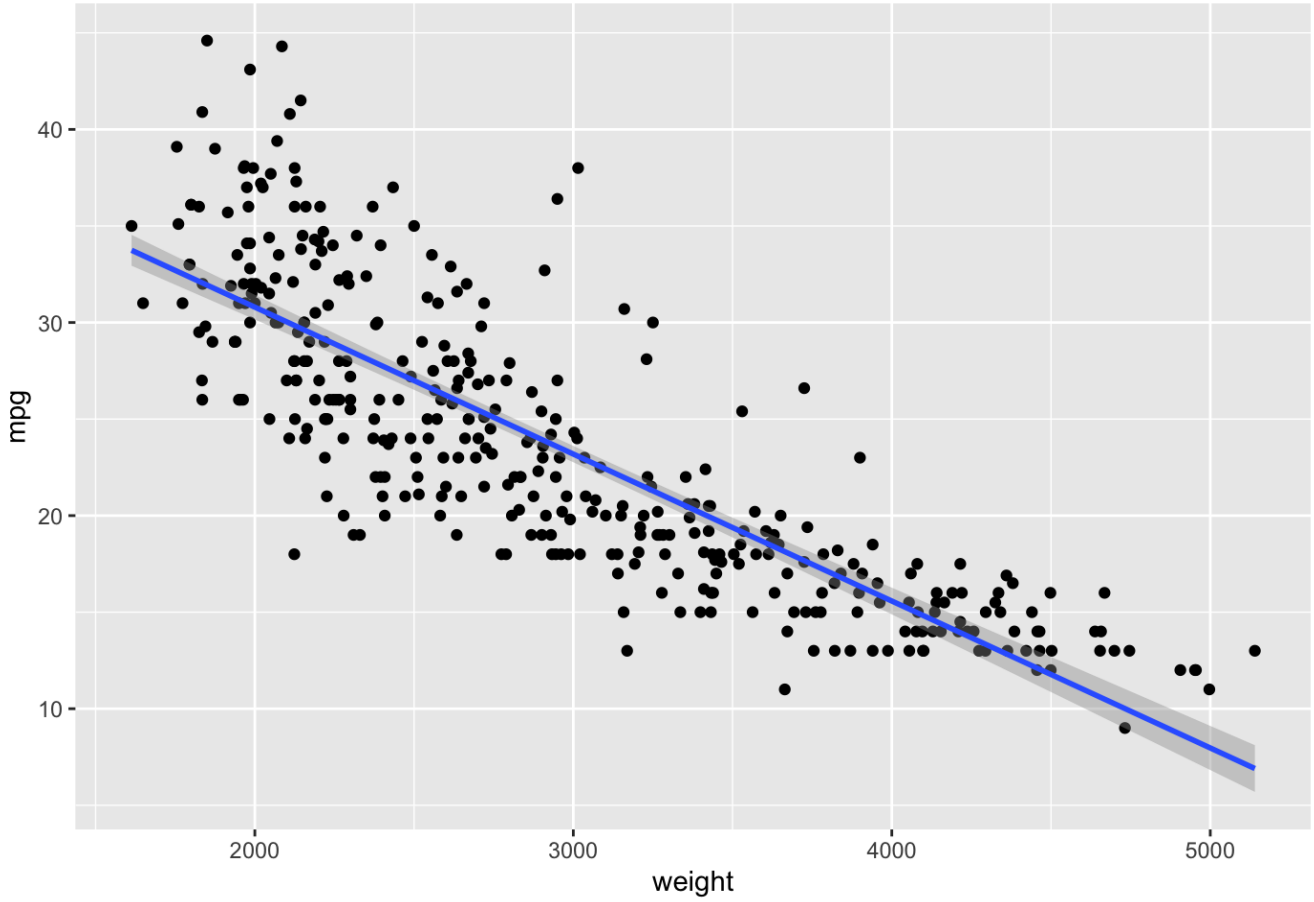

Figure6: Exploring the Relationship Between Weight and MPG Across Car Origins.

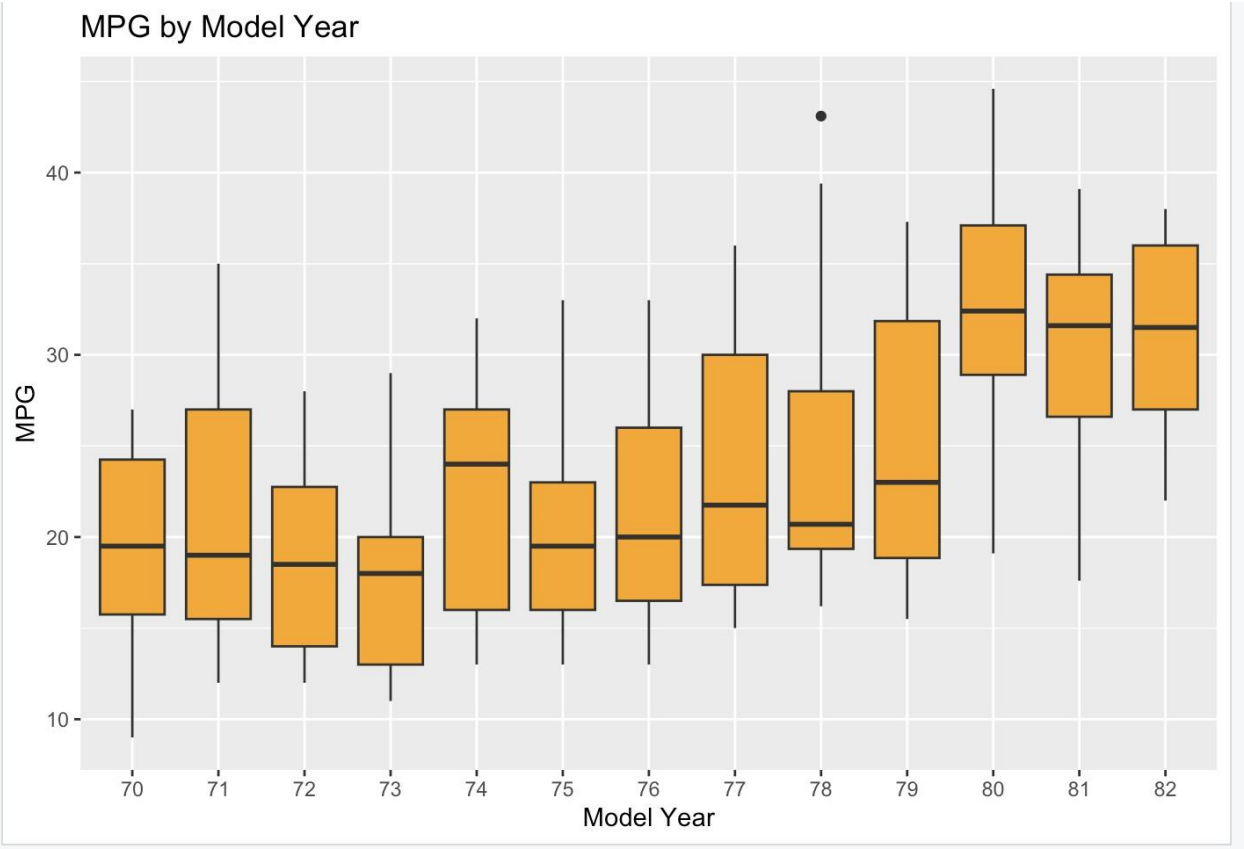

Figure7: Evolution of Average Miles per Gallon (MPG) in U.S. Cars by Model Year.

This box plot graph shows the average miles per gallon (MPG) of cars by model year in the United States. The x-axis shows the model year and the y-axis shows the average MPG.

The graph shows that the average MPG of cars has increased steadily over time. In 1975, the average MPG was 13.5, but by 2022, it had increased to 25.7. This increase in fuel efficiency could be due to a number of factors, including government regulations, technological advances, and consumer demand. The graph also shows that the rate of increase in fuel efficiency has slowed down in recent years. This is likely due to a number of factors, including the increasing popularity of SUVs and trucks, which are typically less fuel-efficient than cars. Overall, it's shown that the average MPG of cars in the United States has increased steadily over time, but the rate of increase has slowed down in recent years.

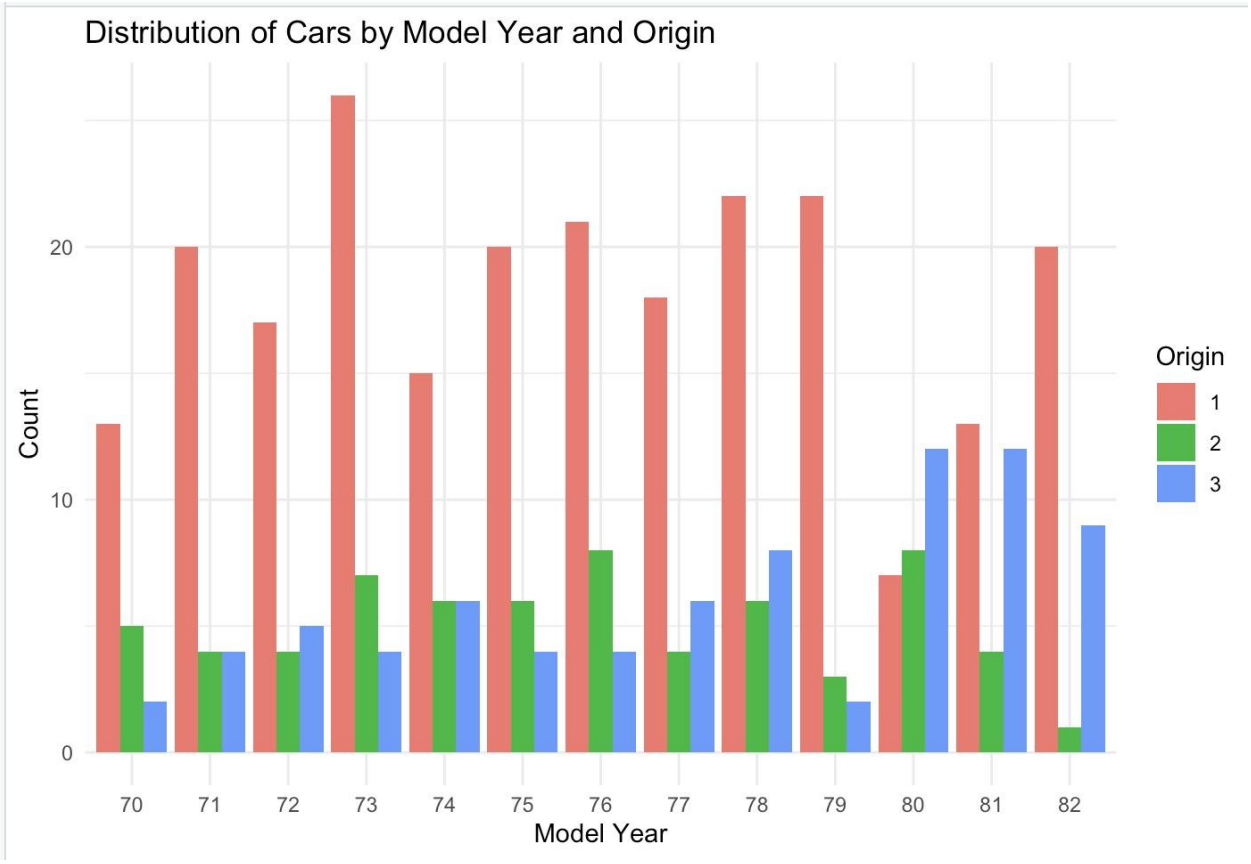

Figure8: Changing Tides: Car Origin Preferences Over Model Years.

The graph displays how cars are spread out across different model years and their countries of origin. Imagine the x-axis as a timeline of car model years, and the y-axis as a list of countries where these cars come from. The length of the bar in the graph shows how many cars there are in each group.

What the graph tells us is that Japanese cars are the most favored, followed by American and German cars. People like Japanese cars because they're reliable and fuel-efficient. American cars, on the other hand, tend to be bigger and more powerful, making them a hit with families and those who need to carry heavy stuff. German cars are famous for their performance and luxury. The graph also reveals that the popularity of cars from these different countries has changed over time. Back in the day, Japanese cars were all the rage in the 1970s and 1980s. Then, in the 1990s and early 2000s, American cars took the lead. Lately, German cars have been gaining more attention. In a nutshell, the graph shows that Japanese cars are number one, followed by American and German cars, and it's interesting to see how people's preferences have shifted over the years.

References:

[1] Auto-mpg dataset. (2017, July 2). Kaggle. https://www.kaggle.com/datasets/uciml/autompg-dataset

[2] Quinlan,R.. (1993). Auto MPG. UCI Machine Learning Repository. (n.d.).https://doi.org/10.24432/C5859H.

[3] RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/

[4] R Core Team (2021). R: The R Project for Statistical Computing. (n.d.-b). https://www.r-project.org/