

A man in a dark blue suit and light blue shirt is leaning into the open driver's side door of a car. He is smiling and gesturing with his hands while talking to a woman sitting in the driver's seat. The woman has curly hair and is wearing a light-colored patterned blouse. She is also smiling and looking towards the man. The car's interior, including the steering wheel and dashboard, is visible. The background is slightly blurred, showing what appears to be a car dealership or showroom with other vehicles and bright lighting.

Analyzing the Impact of Car Features on Price and Profitability

Name: Praharsh Kumar
courses: Data Analytics

Email: Praharshkumar@protonmail.com



Project Description:

☐ Overview:

This project aims to analyze the impact of car features on price and profitability in the automotive industry. It involves exploring a dataset titled "Car Features and MSRP" to provide valuable insights to a car manufacturer on optimizing pricing and product development decisions.

☐ Business Problem:

The primary business problem is to help the car manufacturer maximize profitability while meeting consumer demand. To achieve this, we need to understand the factors driving consumer demand, pricing strategies, and how car features influence pricing.

☐ Data Sources:

The dataset used for analysis is sourced from Kaggle and contains information on over 11,000 car models. It includes variables such as make, model, year, engine specifications, fuel type, transmission type, market category, size, style, fuel efficiency, popularity, and manufacturer's suggested retail price (MSRP).




Assumptions:

- The dataset represents a comprehensive sample of car models available in the market.
- The relationships identified in the analysis are general trends and may not account for all individual cases.
- The dataset, last updated in 2017, provides historical insights and may not reflect the current market.

Approach:

- Descriptive statistics: Used to understand the basic characteristics of the data.
- Visualization: Employed various charts and graphs to visualize trends and relationships in the data.
- Regression analysis: Utilized to identify variables influencing car prices.
- Pivot tables: Created to summarize and analyze data based on different criteria.
- Correlation analysis: Used to quantify relationships between variables.
- Interactive dashboards: Developed using Excel to make data exploration user-friendly and interactive.

- 
- Data Cleaning and Preprocessing:
 - Before conducting the analysis, the following data cleaning and preprocessing steps were performed:
 - Handling missing values: Checked and addressed any missing data in the dataset.
 - Data type conversion: Ensured appropriate data types for each variable.
 - Duplicate removal: Removed duplicate entries if any.
 - Outlier handling: Identified and dealt with outliers.
 - Categorization: Created categories where needed (e.g., market categories).
 - Data normalization: Ensured consistency and accuracy in the dataset.

Tech-Stack Used:

- Microsoft Excel was chosen as the primary tool for data analysis and dashboard creation due to its versatility and compatibility with business needs.
- Excel functions, pivot tables, and charts were used extensively.
- No additional libraries or packages were required.



Power BI



Dataset Description:

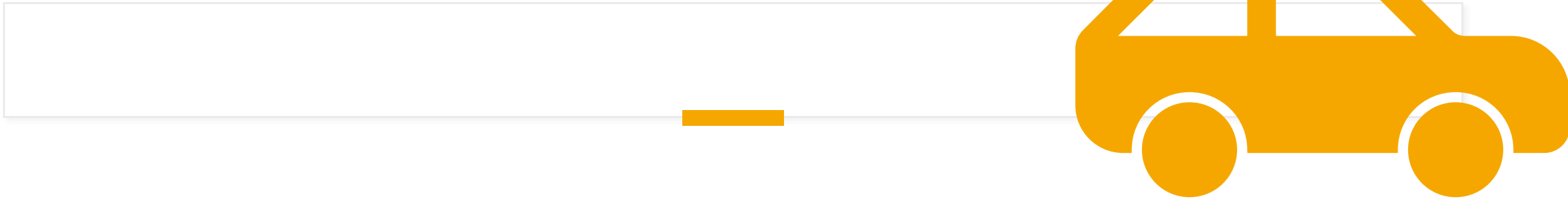
The dataset contains information on various car models and their specifications, and is titled "Car Features and MSRP". It was collected and made available on Kaggle by Cooper Union, a private college located in New York City.

Here is a brief overview of the dataset:

- Number of observations: 11,159
- Number of variables: 16
- File type: CSV (Comma Separated Values)

The variables in the dataset are:

- Make: the make or brand of the car
- Model: the specific model of the car
- Year: the year the car was released
- Engine Fuel Type: the type of fuel used by the car (gasoline, diesel, etc.)
- Engine HP: the horsepower of the car's engine
- Engine Cylinders: the number of cylinders in the car's engine
- Transmission Type: the type of transmission (automatic or manual)
- Driven_Wheels: the type of wheels driven by the car (front, rear, all)
- Number of Doors: the number of doors the car has
- Market Category: the market category the car belongs to (Luxury, Performance, etc.)
- Vehicle Size: the size of the car
- Vehicle Style: the style of the car (Sedan, Coupe, etc.)
- Highway MPG: the estimated miles per gallon the car gets on the highway
- City MPG: the estimated miles per gallon the car gets in the city
- Popularity: a ranking of the popularity of the car (based on the number of times it has been viewed on Edmunds.com)
- MSRP: the manufacturer's suggested retail price of the car

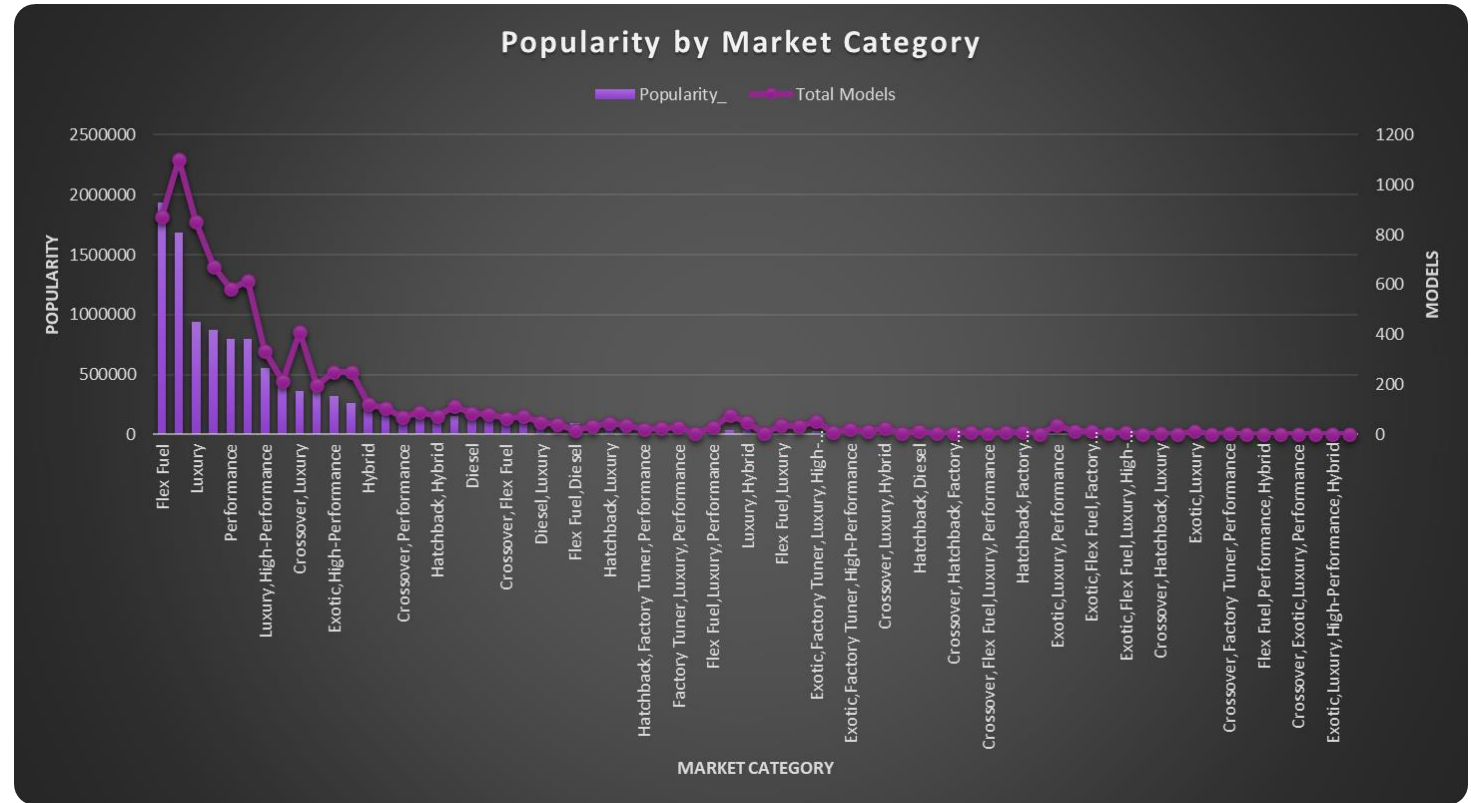


Tasks: Analysis

Insight Required: How does the popularity of a car model vary across different market categories?

Task 1.A: Create a pivot table that shows the number of car models in each market category and their corresponding popularity scores.

Task 1.B: Create a combo chart that visualizes the relationship between market category and popularity.



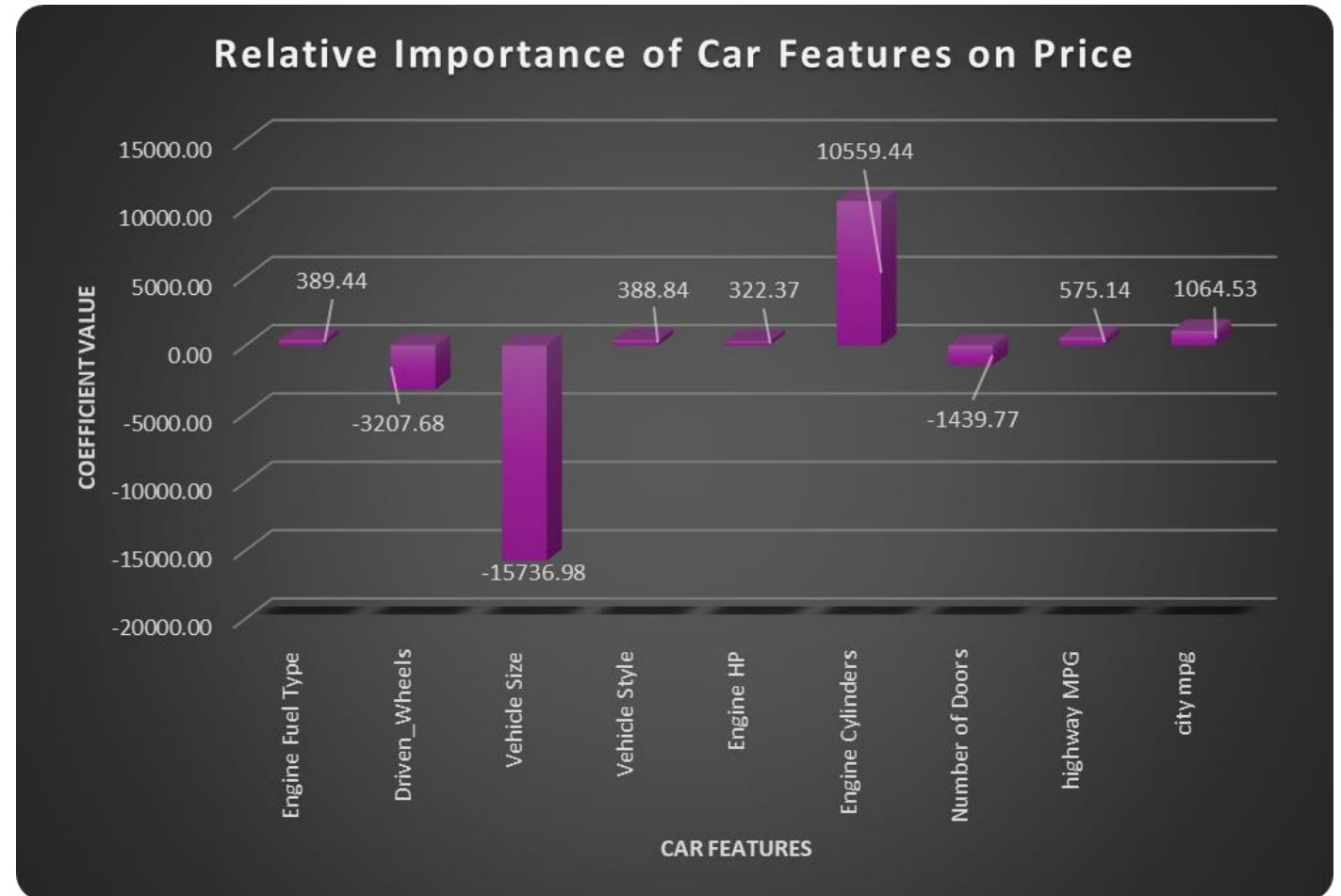
Insight Required: What is the relationship between a car's engine power and its price?

Task 2: Create a scatter chart that plots engine power on the x-axis and price on the y-axis. Add a trendline to the chart to visualize the relationship between these variables



Insight Required: Which car features are most important in determining a car's price?

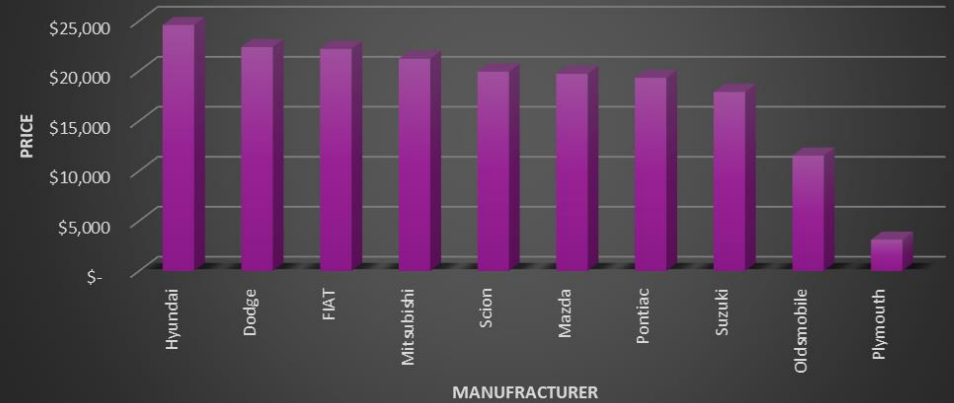
Task 3: Use regression analysis to identify the variables that have the strongest relationship with a car's price. Then create a bar chart that shows the coefficient values for each variable to visualize their relative importance.



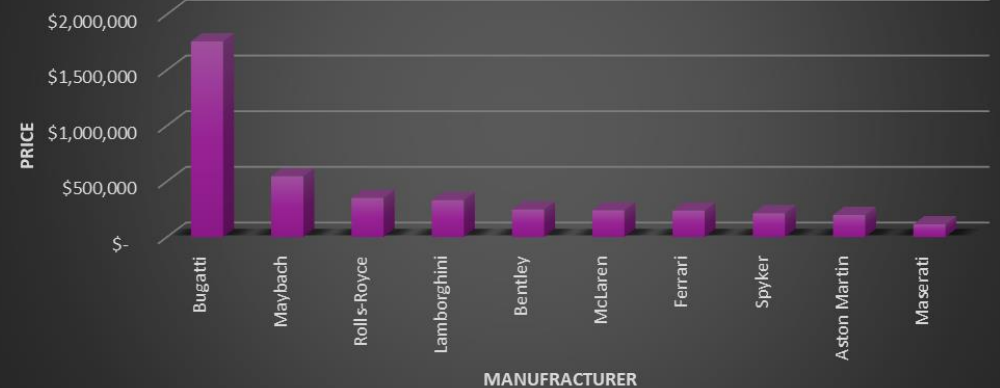
Insight Required: How does the average price of a car vary across different manufacturers?

- **Task 4.A:** Create a pivot table that shows the average price of cars for each manufacturer.
- **Task 4.B:** Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.

Price Comparison across Manufacturers: Bottom 10



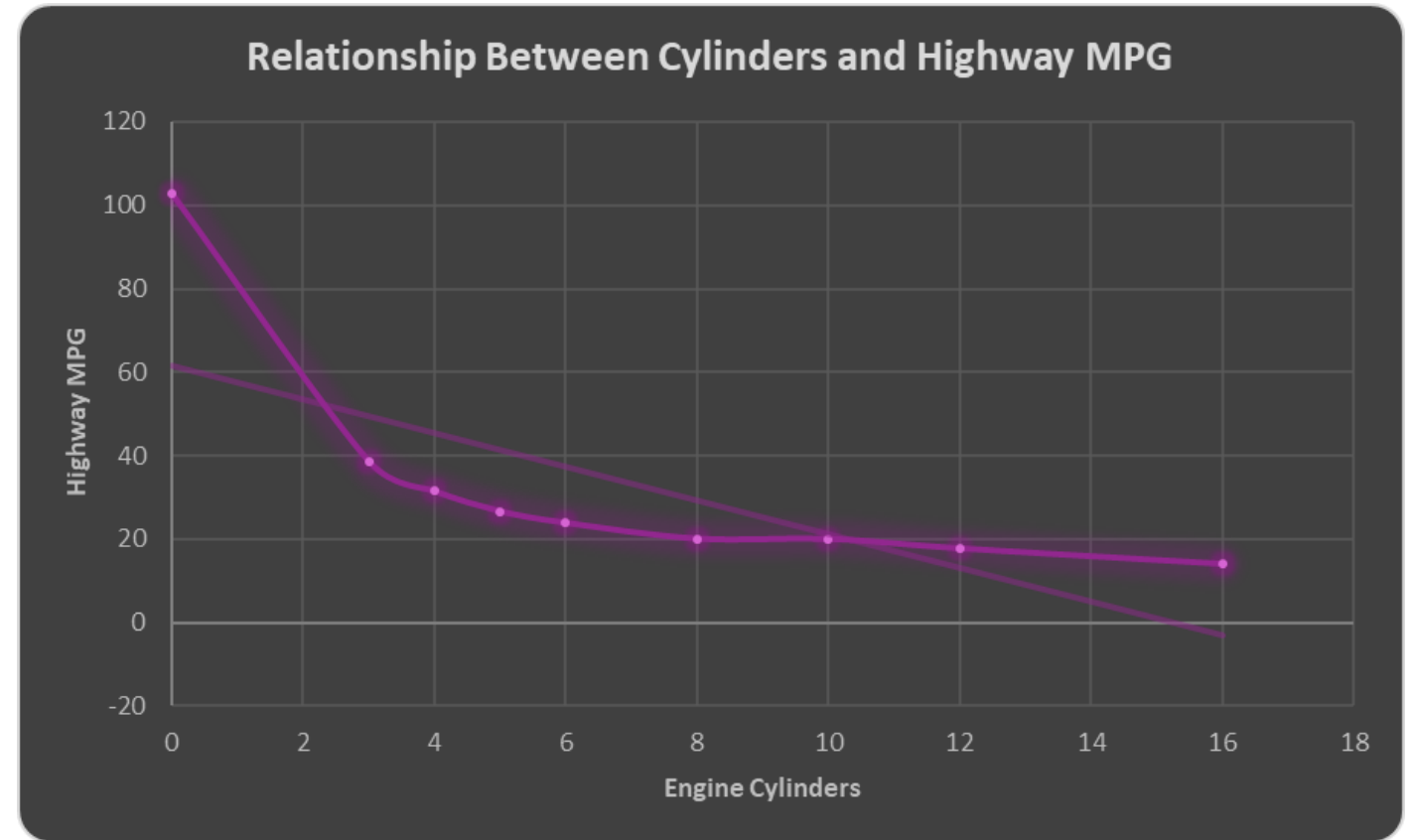
Price Comparison across Manufacturers:
Top 10



Insight Required: What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

Task 5.A: Create a scatter plot with the number of cylinders on the x-axis and highway MPG on the y-axis. Then create a trendline on the scatter plot to visually estimate the slope of the relationship and assess its significance.

Task 5.B: Calculate the correlation coefficient between the number of cylinders and highway MPG to quantify the strength and direction of the relationship.



BODY STYLE

Select all

2dr
Hatchback

2dr SUV

4dr
Hatchback

4dr SUV

Cargo
Minivan

Refresh

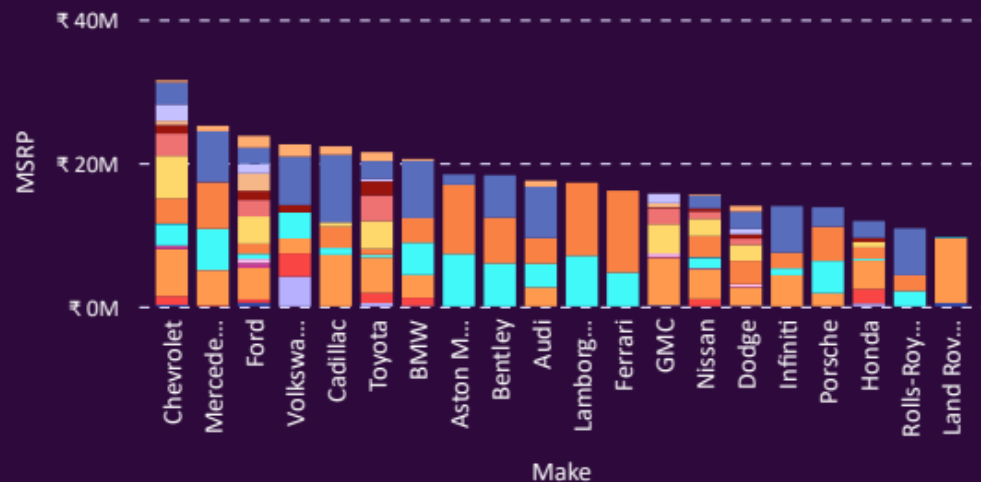


Back



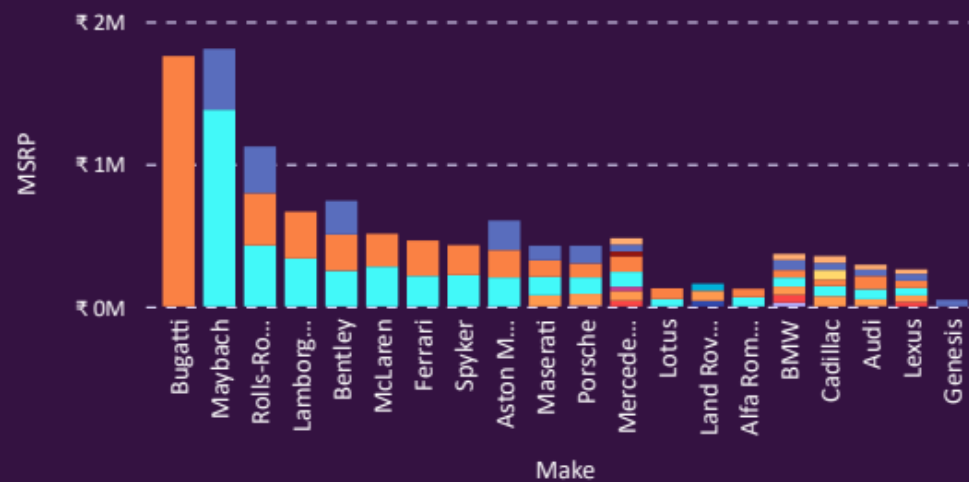
CAR PRICE DISTRIBUTION BY BRAND AND BODY STYLE

Body Style ● 2dr Hatchb... ● 2dr SUV ● 4dr Hatch... ● 4dr SUV ● Cargo Mini...



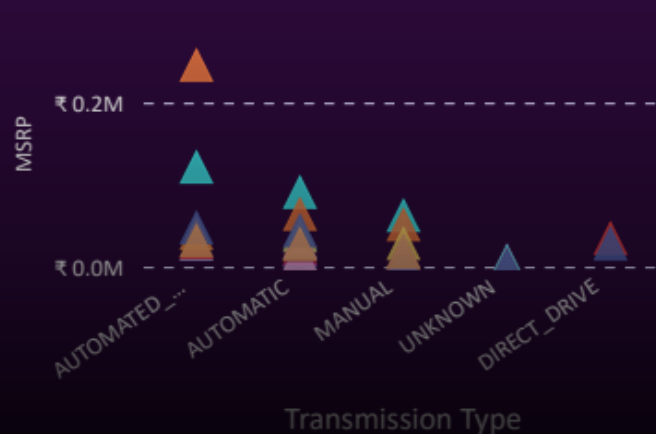
AVERAGE MSRPs BY CAR BRAND AND BODY STYLE

Body Style ● 2dr Hatchb... ● 2dr SUV ● 4dr Hatch... ● 4dr SUV ● Cargo Mini...



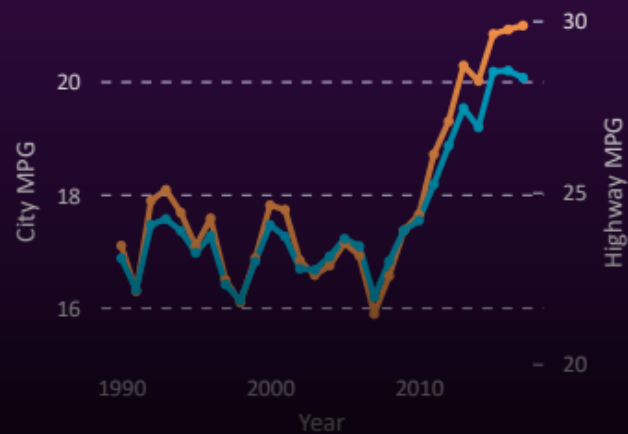
IMPACT OF TRANSMISSION TYPE ON MSRP BY BODY STYLE

Body Style ▲ 2dr Hatch... ▲ 2dr SUV ▲ 4dr Hatch...



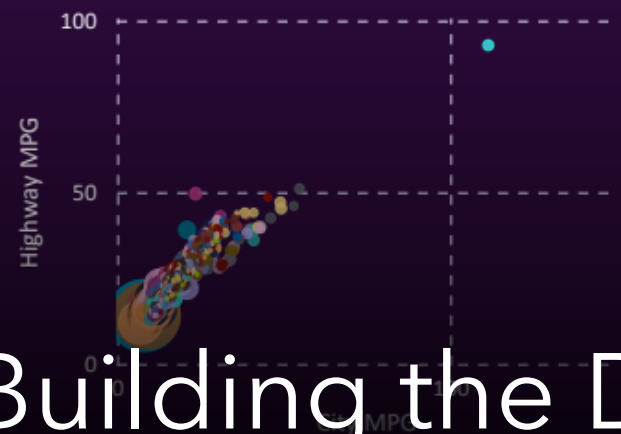
FUEL EFFICIENCY ACROSS BODY STYLES AND MODEL YEARS

● City MPG ● Highway MPG



EXPLORING VARIATION IN HORSEPOWER, MPG AND PRICE ACROSS CAR BRANDS

Make ● Acura ● Alfa Rom... ● Aston M... ● Audi



MAKE

Select all

Dodge

Acura

Ferrari

Alfa Romeo

FIAT

Aston Martin

Ford

Audi

Genesis

Bentley

GMC

BMW

Honda

Bugatti

HUMMER

Buick

Hyundai

Cadillac

Infiniti

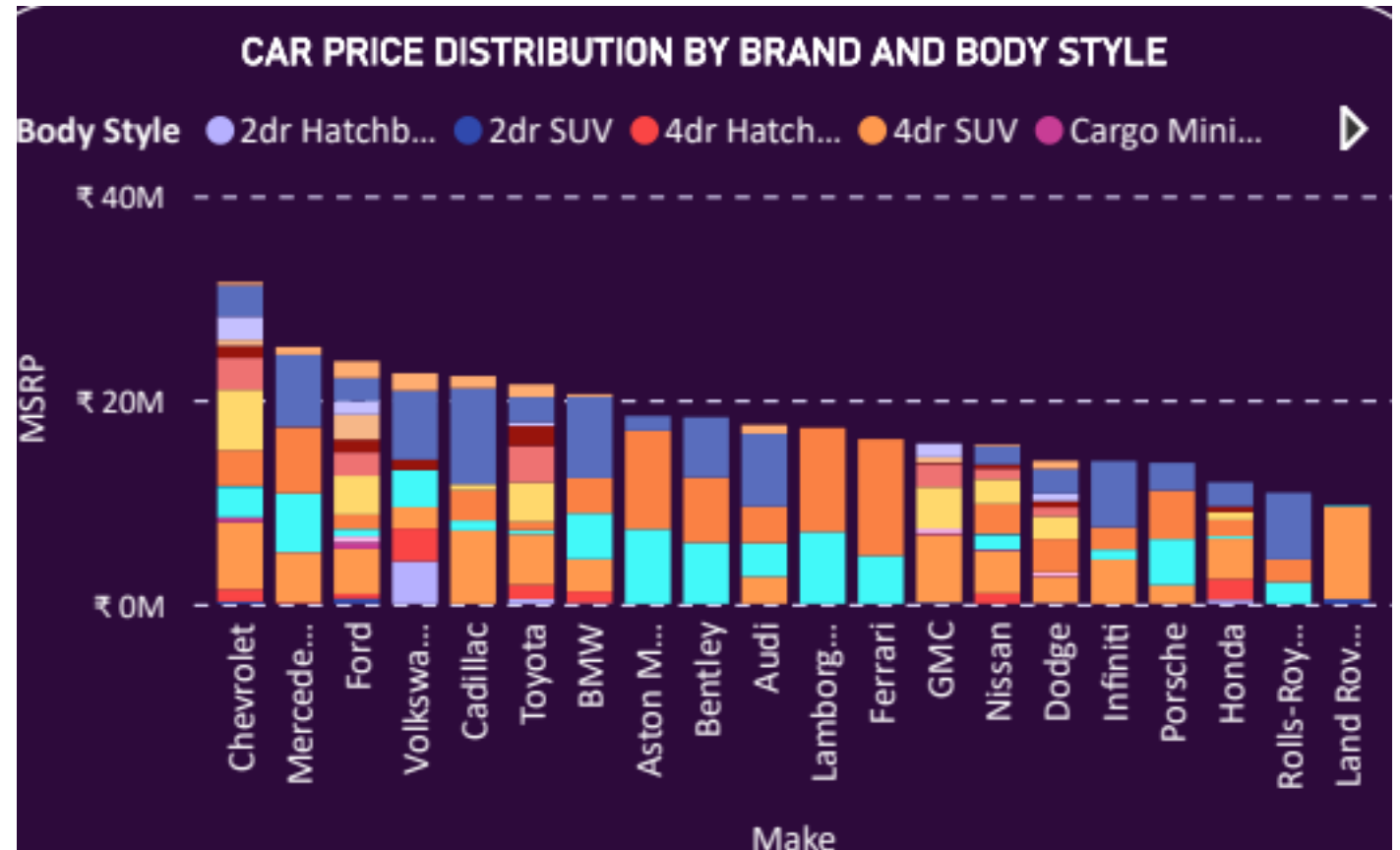
Chevrolet

Kia

Building the Dashboard

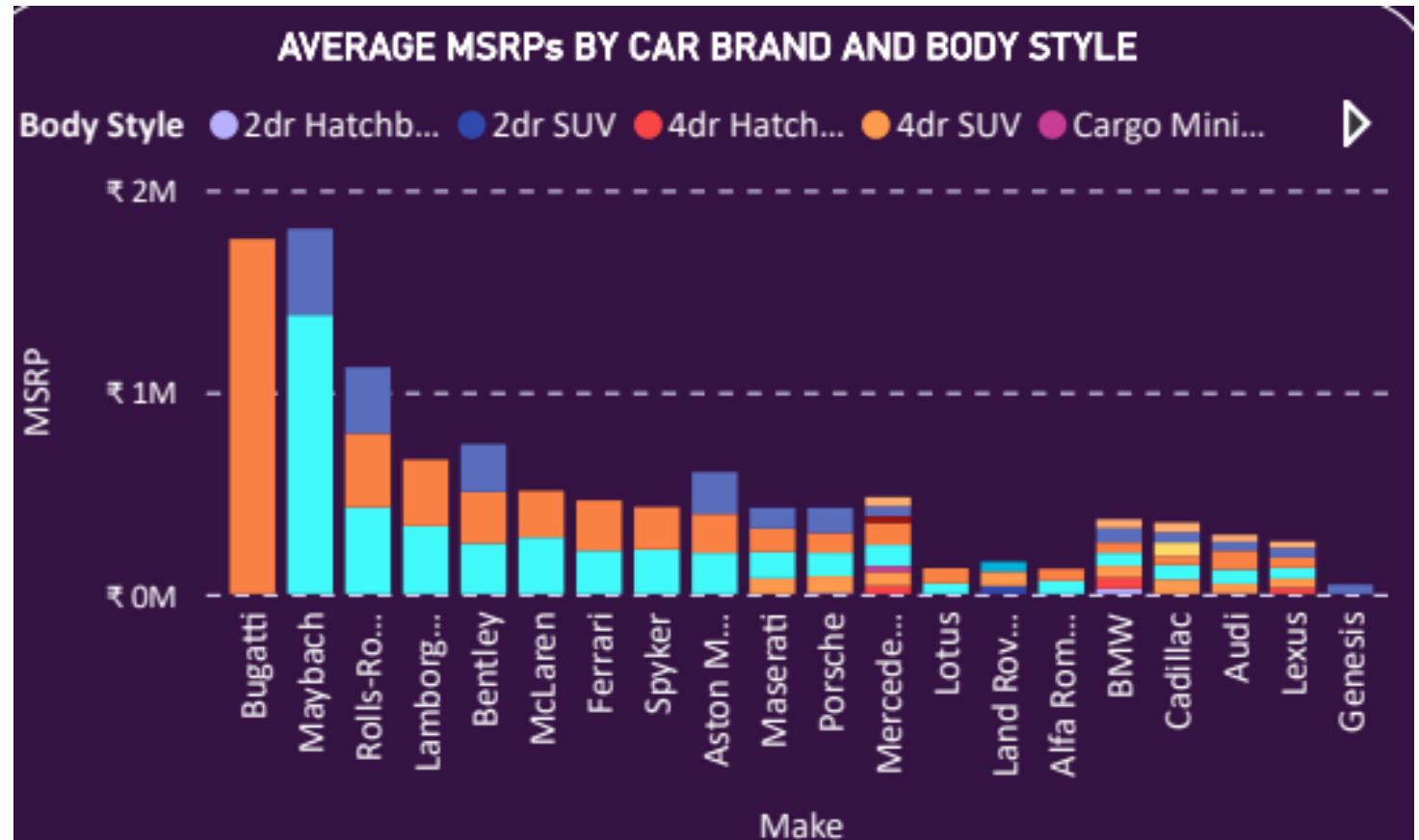
Task 1: How does the distribution of car prices vary by brand and body style?

Hints: Stacked column chart to show the distribution of car prices by brand and body style. Use filters and slicers to make the chart interactive. Calculate the total MSRP for each brand and body style using SUMIF or Pivot Tables.



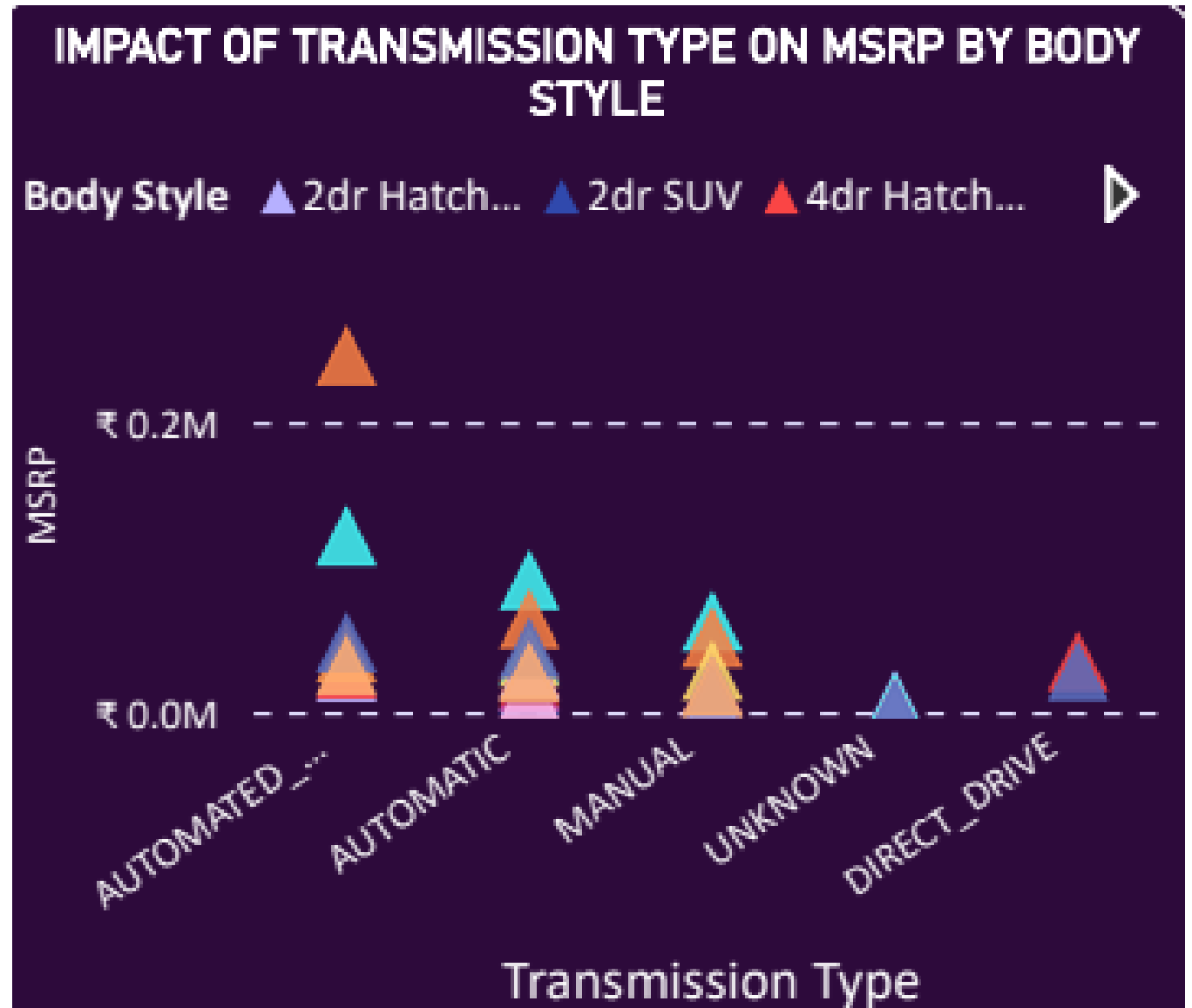
Task 2: Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?

- Hints: Clustered column chart to compare the average MSRPs across different car brands and body styles. Calculate the average MSRP for each brand and body style using AVERAGEIF or Pivot Tables.



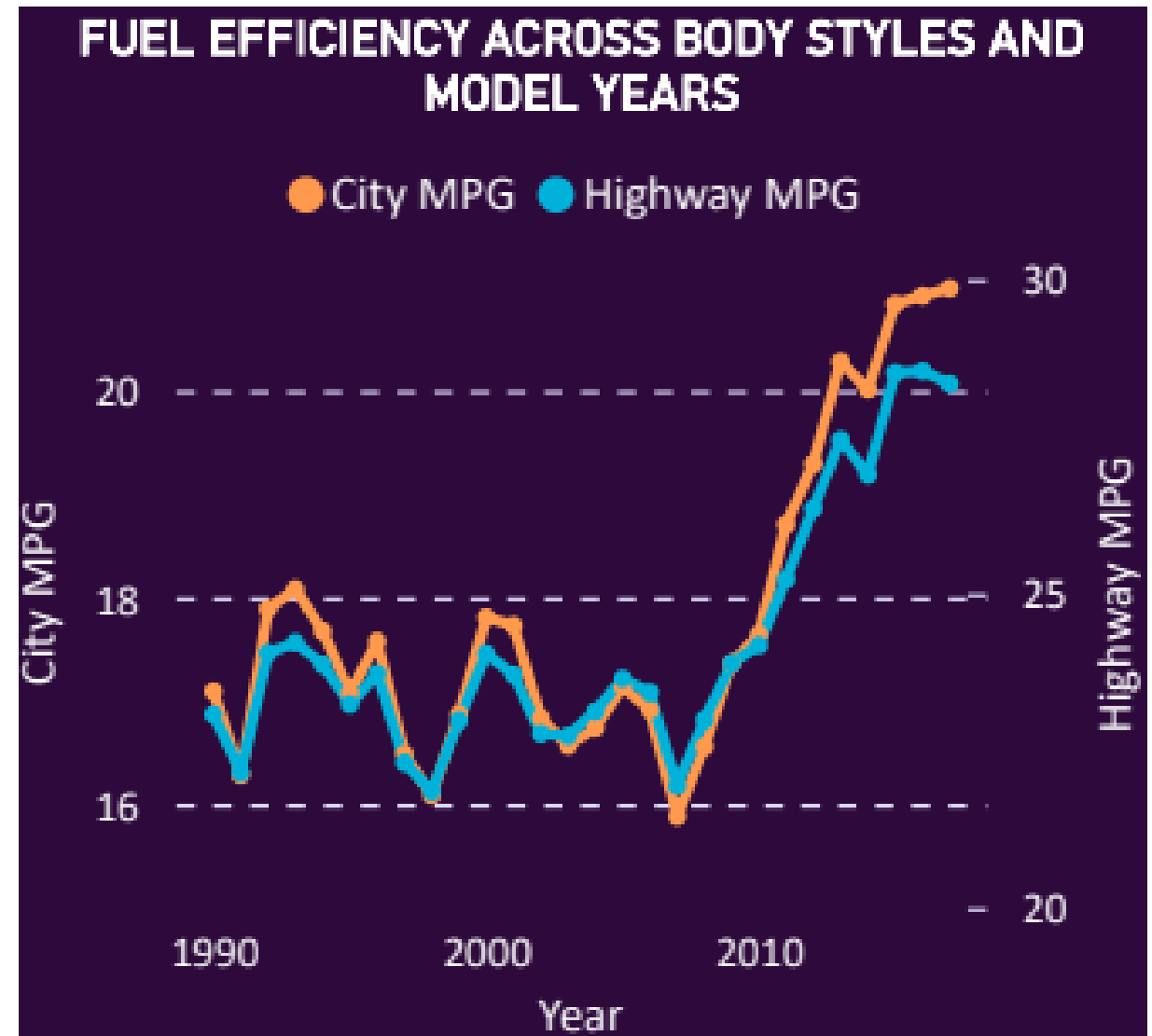
Task 3: How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?

- Hints: Scatter plot chart to visualize the relationship between MSRP and transmission type, with different symbols for each body style. Calculate the average MSRP for each combination of transmission type and body style using AVERAGEIFS or Pivot Tables.



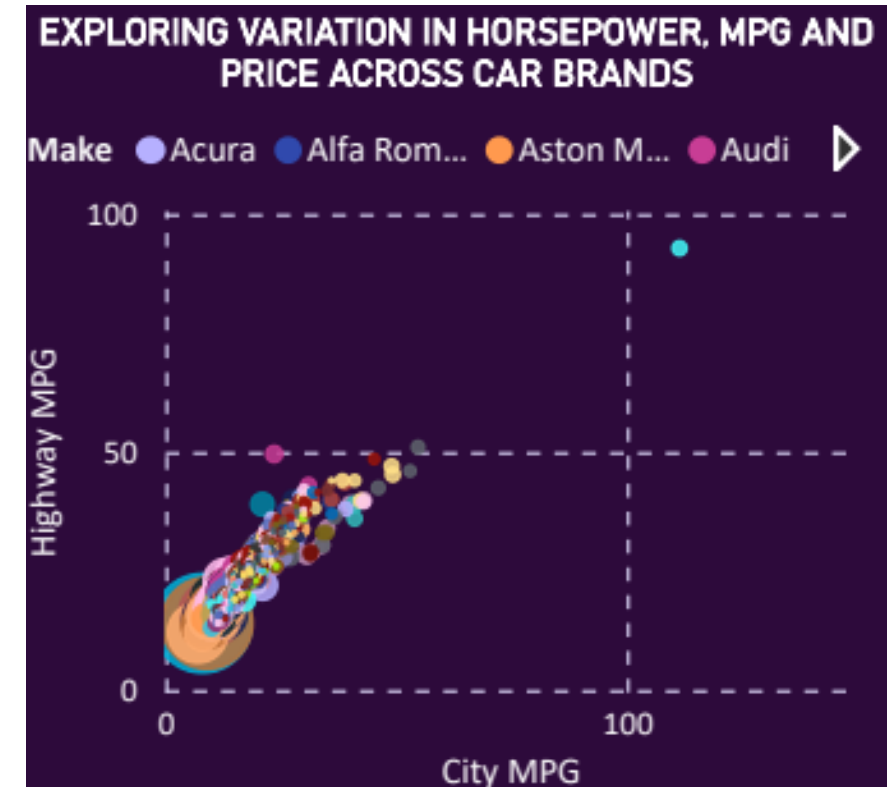
Task 4: How does the fuel efficiency of cars vary across different body styles and model years?

- Hints: Line chart to show the trend of fuel efficiency (MPG) over time for each body style. Calculate the average MPG for each combination of body style and model year using AVERAGEIFS or Pivot Tables.



Task 5: How does the car's horsepower, MPG, and price vary across different Brands?

- Hints: Bubble chart to visualize the relationship between horsepower, MPG, and price across different car brands. Assign different colors to each brand and label the bubbles with the car model name. Calculate the average horsepower, MPG, and MSRP for each car brand using AVERAGEIFS or Pivot Tables.





Insights:



Key Insights:



Market Categories and Popularity: Luxury and Performance market categories tend to have higher popularity scores among car models. This information can guide the manufacturer's focus on these categories.



Engine Power and Price: A positive correlation exists between engine power and price, suggesting that more powerful engines are associated with higher prices.



Influential Car Features: Regression analysis identified specific features that strongly influence car prices, helping prioritize product development efforts.



Manufacturer and Average Price: The analysis revealed variations in average car prices among manufacturers, enabling the identification of pricing strategies.



Fuel Efficiency and Cylinders: A negative correlation between the number of cylinders and highway MPG suggests that fewer cylinders may lead to better fuel efficiency.



Results:

- Interactive dashboards were created in Excel to answer specific questions posed by the client.
- Charts, tables, and graphs were used to visualize trends, correlations, and comparisons.
- Filters and slicers were incorporated to allow users to interact with the data.

