

IMPACT OF CAR FEATURES

Analyzing the Impact of Car Features
on Price and Profitability

BY - SUKHMANI ARORA



Problem Description

How can a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer demand?

The task is to analyze the relationship between a car's features, market category, and pricing, and identify which features and categories are most popular among consumers and most profitable for the manufacturer.

The dataset contains information on various car models and their specifications, and is titled "Car Features and MSRP".

Approach

Understanding of Data

Firstly, I observed the data. It contains 11915 total records and has columns representing the car's make, model, year, fuel type, engine power, etc.

Data Cleaning

In this step, 715 duplicate rows in the dataset were removed and the records with missing values were also removed as there were very few records.

11097 records remained after data cleaning.

Analysis

For analysis, I created different types of charts and dashboards to find insights, trends, patterns, and relationships within the data.

Workbook Link :

<https://drive.google.com/file/d/1JohDTBIFeGoW5u6YWUhG7rohts9qV39f/view?usp=sharing>

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

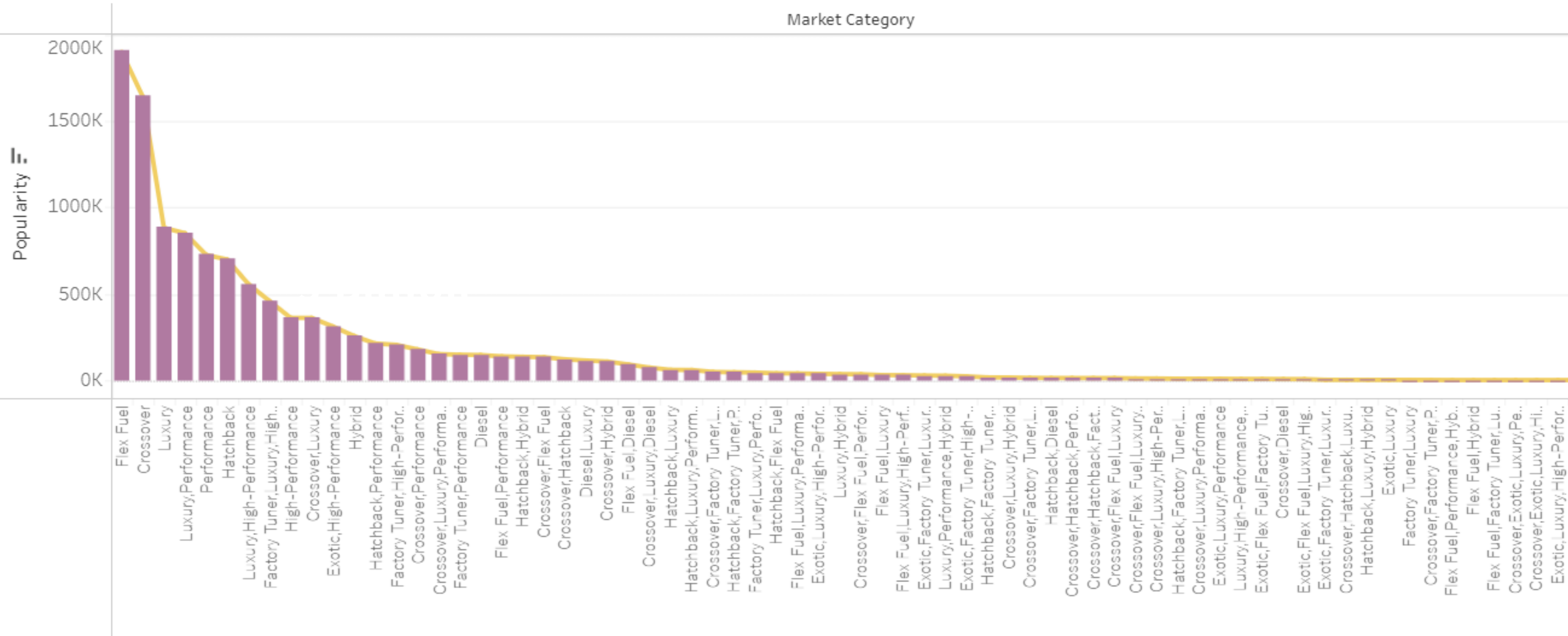
- **Task 1A:** Create a pivot table that shows the number of car models in each market category and their corresponding popularity scores.
- **Task 1B:** Create a combo chart that visualizes the relationship between market category and popularity.

Task 1A Result :

- This pivot table is showing the number of models in each category and their popularity in that category.
- The highest popular market categories are :
 - Hatchback, Flex fuel
 - Flex fuel, Diesel
 - Crossover, Flex Fuel, Performance
- The maximum number of models in the Crossover category.

Market Category	Number of Models	Avg. Popularity
Hatchback, Flex Fuel	7	5,657
Flex Fuel, Diesel	16	5,657
Crossover, Flex Fuel, Perfo..	6	5,657
Crossover, Luxury, Perfor..	2	3,916
Crossover, Factory Tuner, ..	5	2,607
Crossover, Performance	69	2,586
Crossover, Hybrid	42	2,563
Diesel, Luxury	47	2,416
Luxury, Performance, Hybr..	11	2,333
Flex Fuel	855	2,226
Crossover, Luxury, Diesel	33	2,196
Hatchback, Factory Tuner, ..	21	2,174
Factory Tuner, Luxury, Hig..	215	2,133
Hybrid	121	2,117
Hatchback, Hybrid	64	2,111
Crossover, Flex Fuel	64	2,074
Crossover, Hatchback, Perf..	6	2,009
Crossover, Hatchback, Fac..	6	2,009
Factory Tuner, High-Perfo..	104	1,966
Crossover, Factory Tuner, ..	26	1,823
High-Performance	198	1,823
Factory Tuner, Performan..	81	1,818
Diesel	84	1,731
Flex Fuel, Performance	81	1,702
Crossover, Hatchback	72	1,676
Luxury, High-Performance	334	1,668
N/A	3,362	1,659
Hatchback, Luxury, Perfor..	36	1,632
Crossover, Flex Fuel, Luxur..	6	1,624
Crossover	1,068	1,539
Performance	503	1,443
Factorv Tuner. Luxurv. Perf..	31	1.413

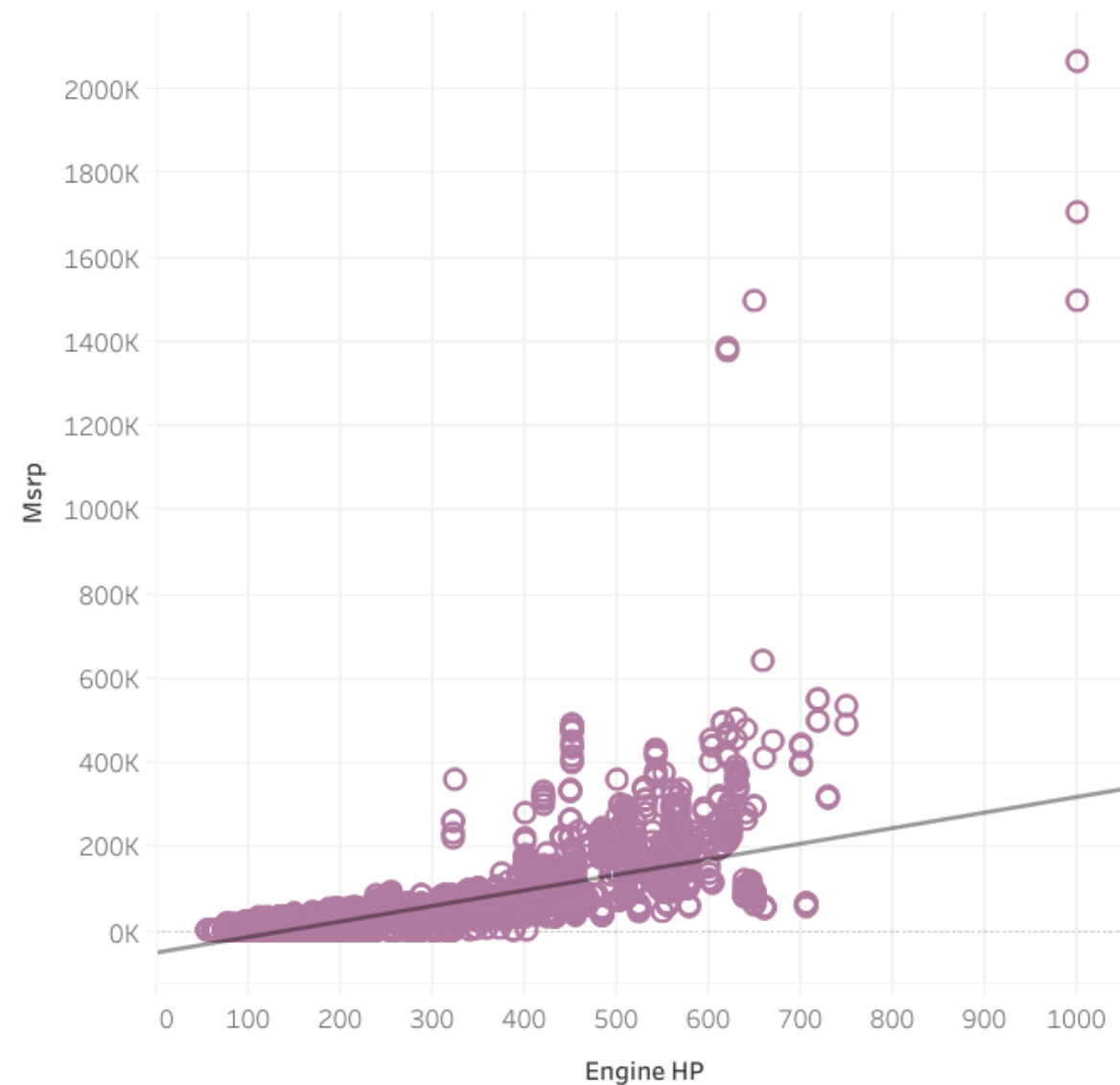
Task 1B Result :



The bars show the total popularity score for each market category and the line chart shows any trend or pattern in popularity score across market categories.

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.



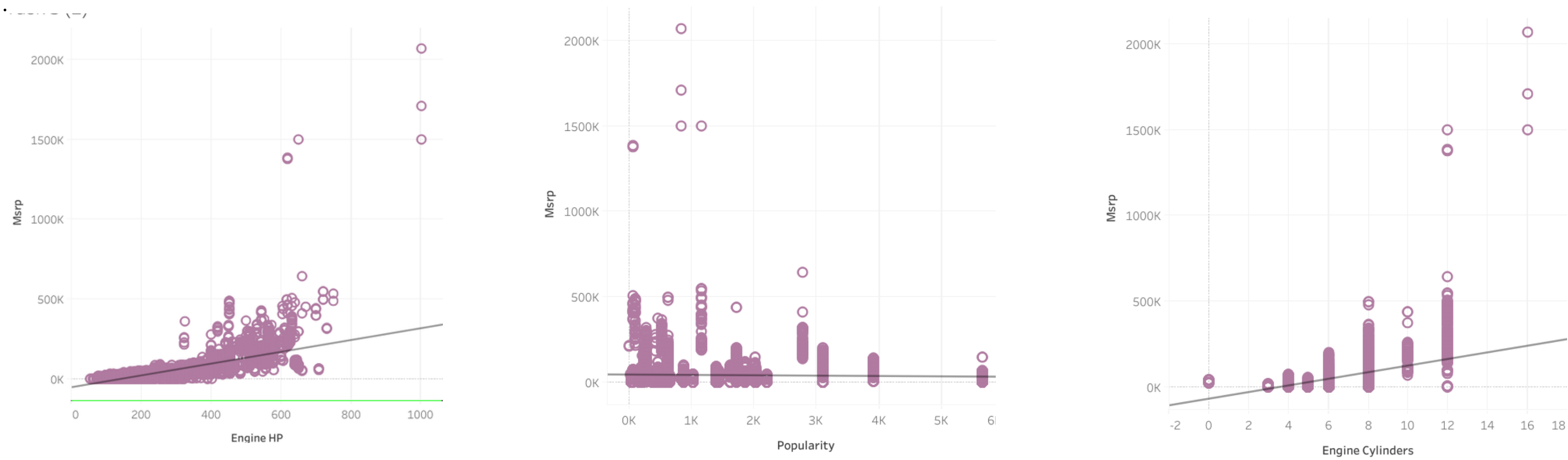
Result :

- Here, a scatter chart is plotted for engine power and price of cars.
- We can observe from the chart that when engine power increases, the price also increases.
- The trend line shows a positive slope, meaning engine and price are positively correlated.

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.

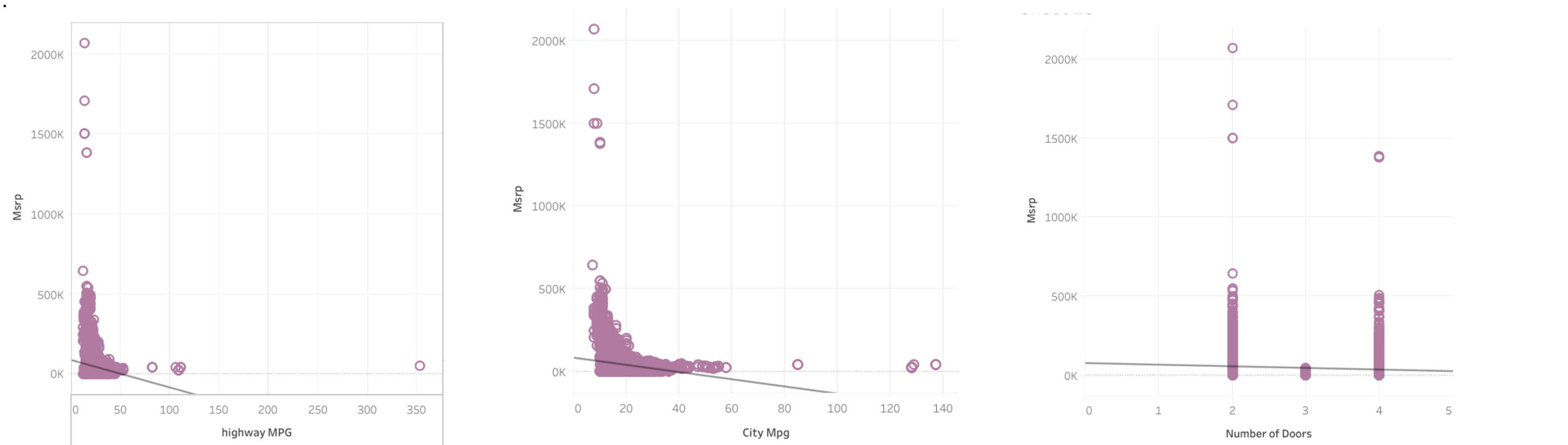
These are the results of regression analysis of all numerical columns to identify the variables which have a strong relationship with the car price.



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.

These are the results of regression analysis of all numerical columns to identify the variables which have a strong relationship with the car price.



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.

Result :

Upon doing regression analysis on the variables with Msrp, a correlation coefficient is calculated for each variable, which represents their relationship with Msrp.

Engine HP: 0.6590

Popularity: -0.04849

Engine Cylinders: 0.5520

highway MPG: -0.2078

city Mpg: -0.2327

Number of Doors: -0.1456

Popularity, highway MPG, city Mpg, and Number of Doors have a negative correlation coefficient which means that with an increase in the value of these variables, price decreases, and Engine HP and Engine Cylinders have a positive correlation coefficient which means that with an increase in the value of these variables, the price also increases.

Engine HP has the strongest relationship with car price.



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

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

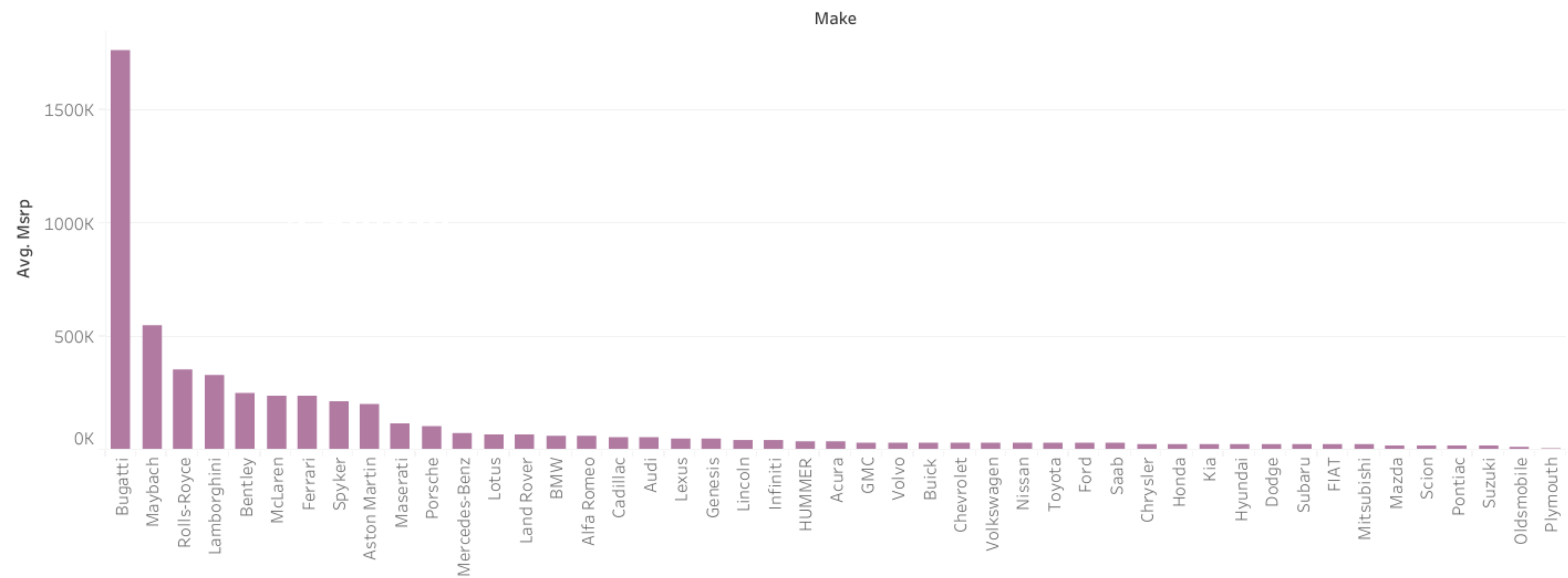
Task 4A Result :

- This pivot table is showing how the average price of cars varies across different manufacturers.
- The highest price are of brands :
Bugatti
Maybach
- The lowest price are of brands :
Plymouth
Oldsmobile

Make	
Bugatti	1,757,224
Maybach	546,222
Rolls-Royce	351,131
Lamborghini	331,567
Bentley	247,169
McLaren	239,805
Ferrari	237,384
Spyker	214,990
Aston Martin	198,123
Maserati	113,684
Porsche	101,622
Mercedes-Benz	72,135
Lotus	68,377
Land Rover	68,067
BMW	62,163
Alfa Romeo	61,600
Cadillac	56,368
Audi	54,574
Lexus	47,549
Genesis	46,617

Task 4B Result :

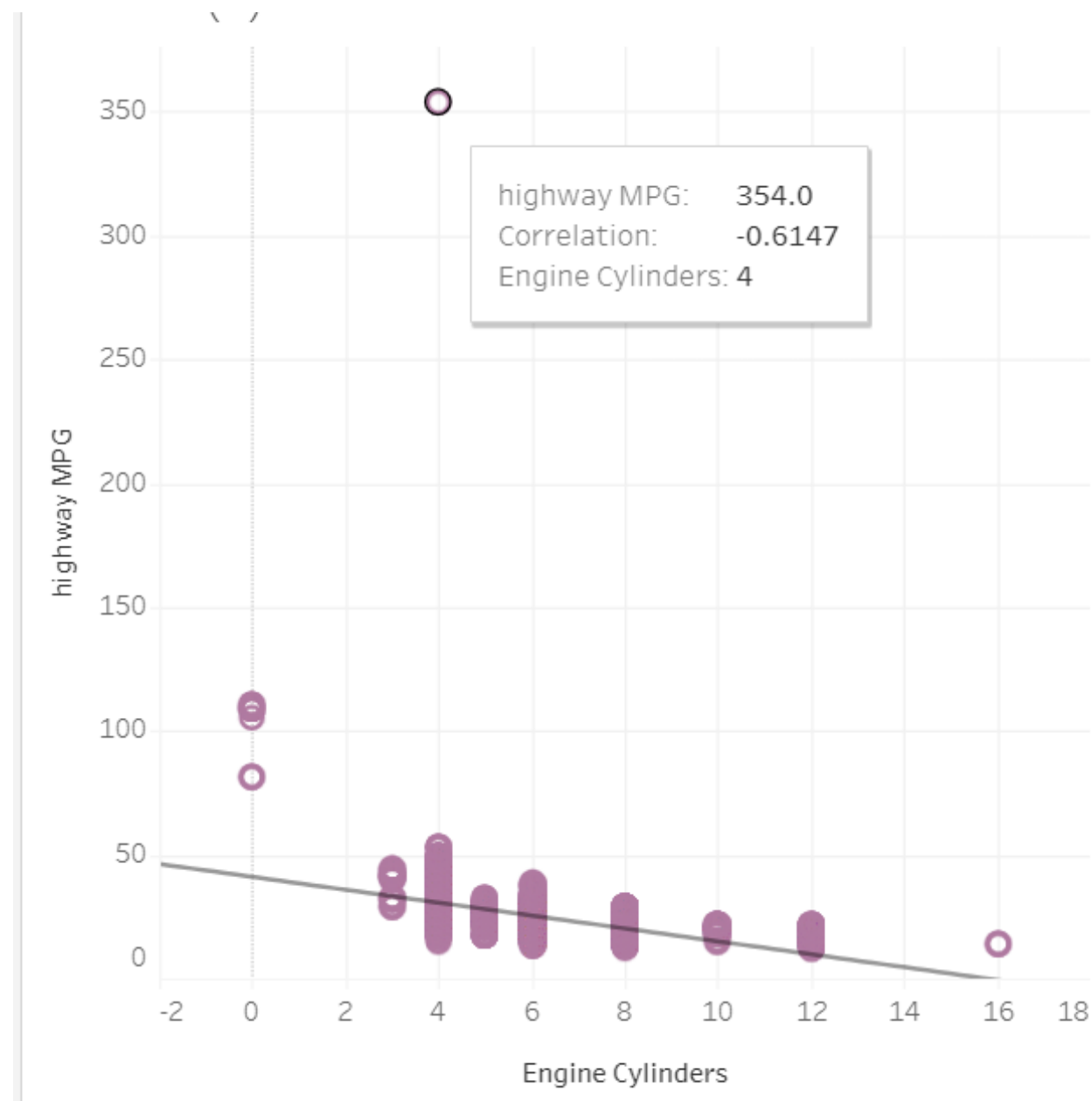
Task 4B



This graph shows the average msrp of all brands, we can easily identify the brands with highest and lowest average msrp which are "Buggati" and "Plymouth".

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

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



Task 5A Result :

- The scatter plot and trendline analysis show a negative slope in the relationship between the number of cylinders in a car's engine and its highway MPG which means that as the number of cylinders increases, the fuel efficiency (highway MPG) tends to decrease.

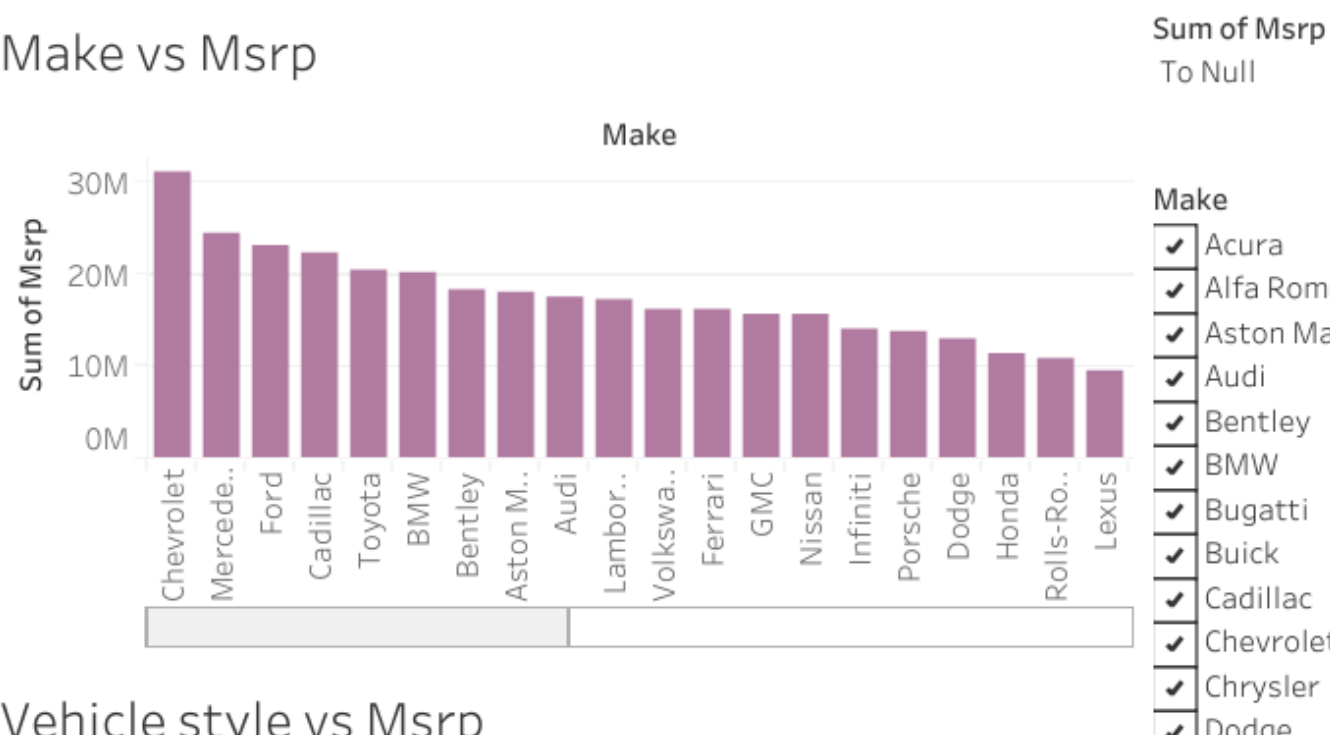
Task 5B Result :

- A calculated field is created for showing the correlation value for highway MPG and Engine Cylinders.
- The correlation value comes out to be -0.6147 which means these two variables are negatively correlated.
- This negative correlation suggests that as the number of engine cylinders increases, the highway MPG tends to decrease.

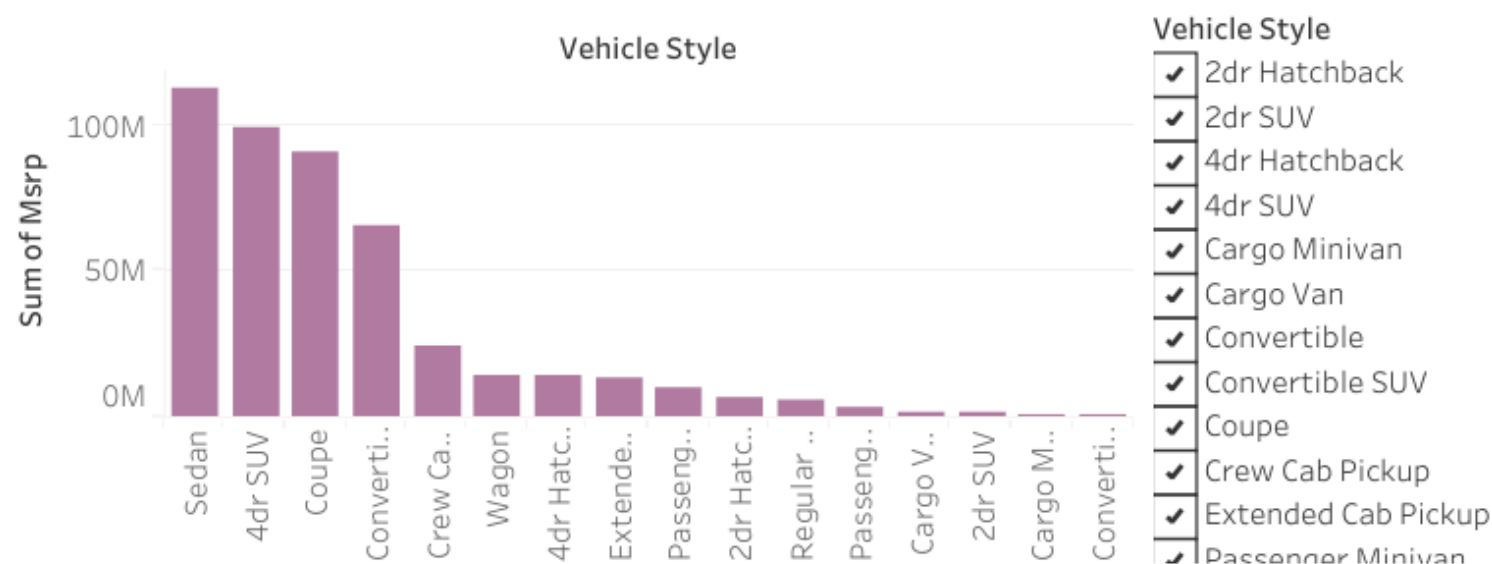
Building dashboard

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

Make vs Msrp



Vehicle style vs Msrp



Results :

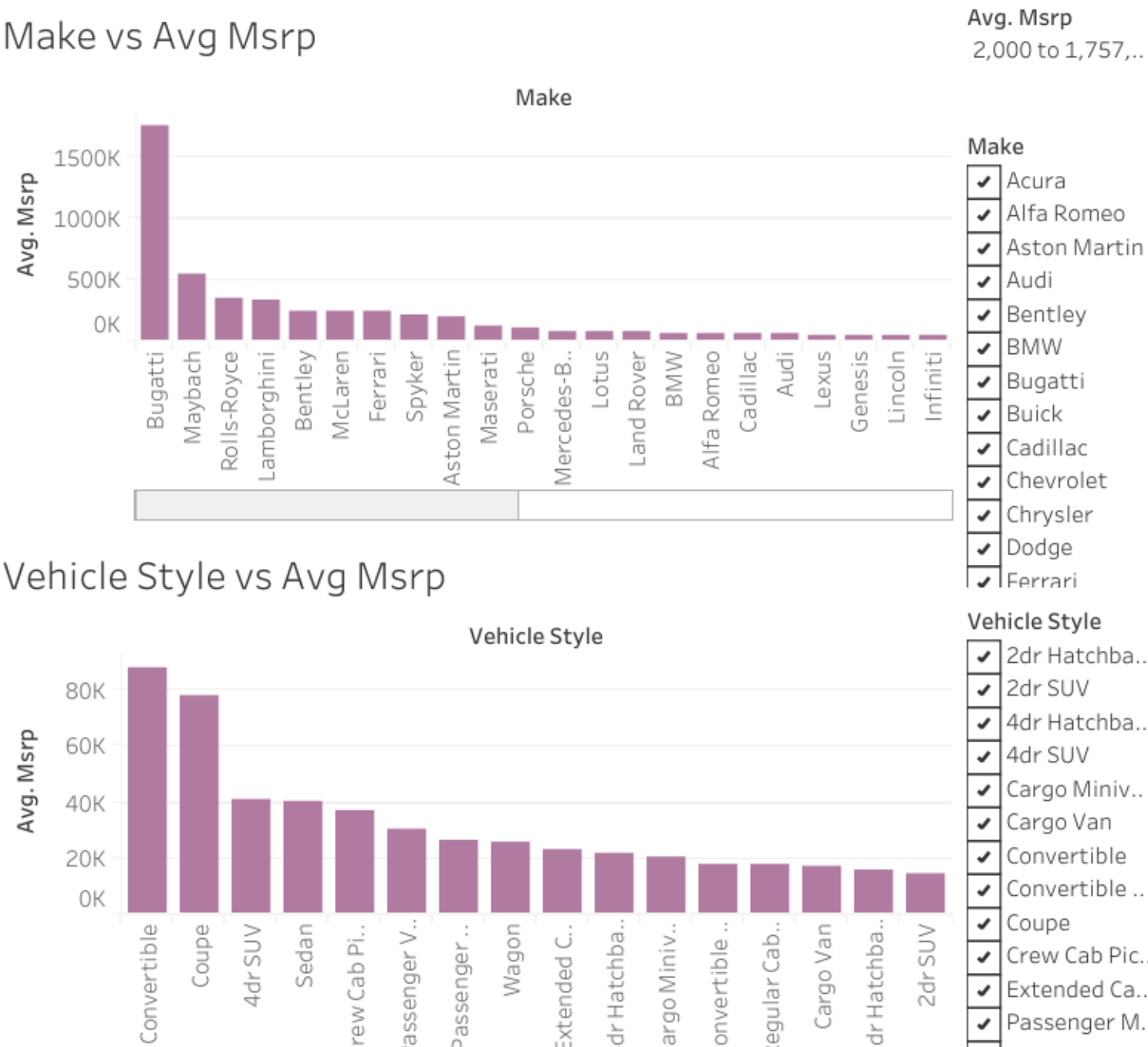
- These graphs show the distribution of car prices by brand and body style.
- Filters are added to see the msrp for any brand and body style separately.
- "Chevrolet" brand has the highest total msrp.
- "Sedan" body style has the highest total msrp.

Building dashboard

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

Results :

- These graphs show the average msrp for each brand and body style.
- Filters are added to see the average msrp for any brand and body style separately.
- "Bugatti" brand has the highest average msrp.
- "Convertible" body style has the highest average msrp.

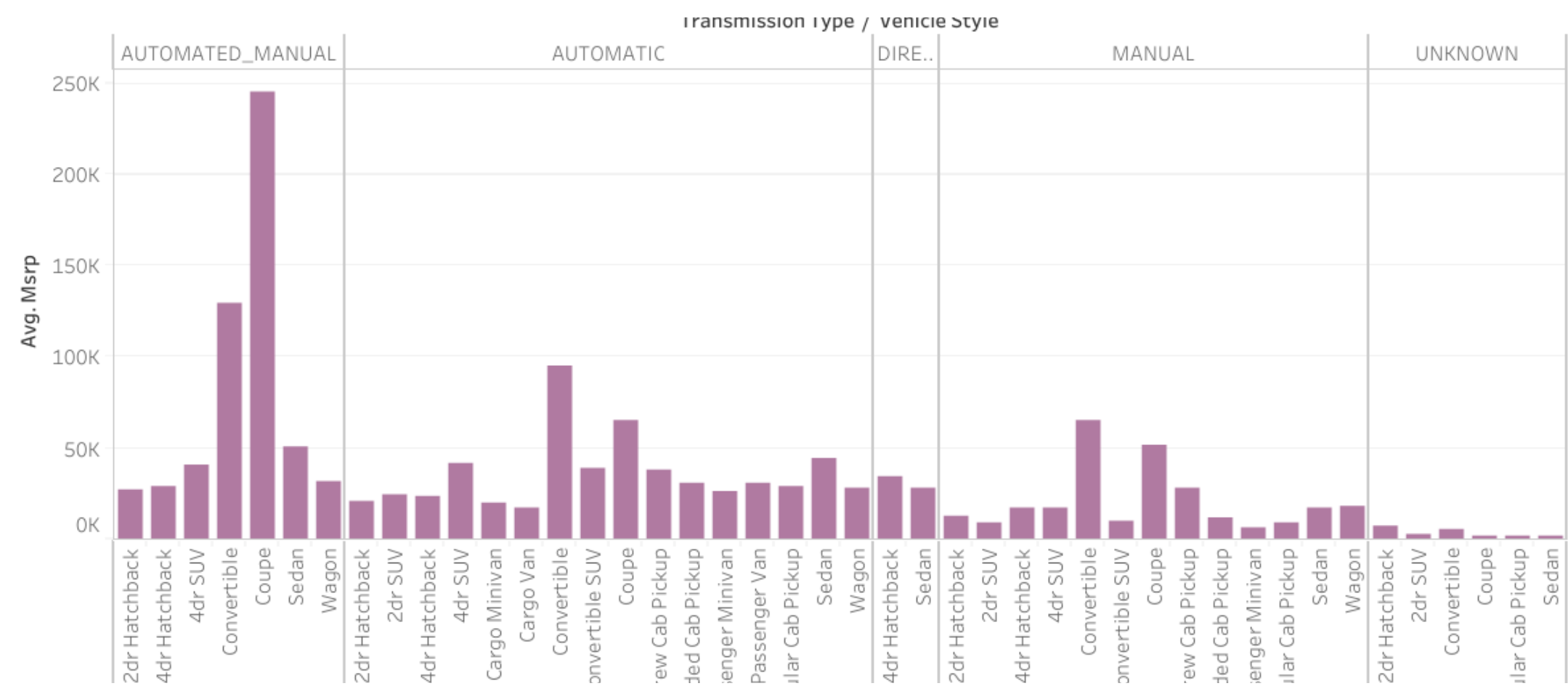


Building dashboard

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

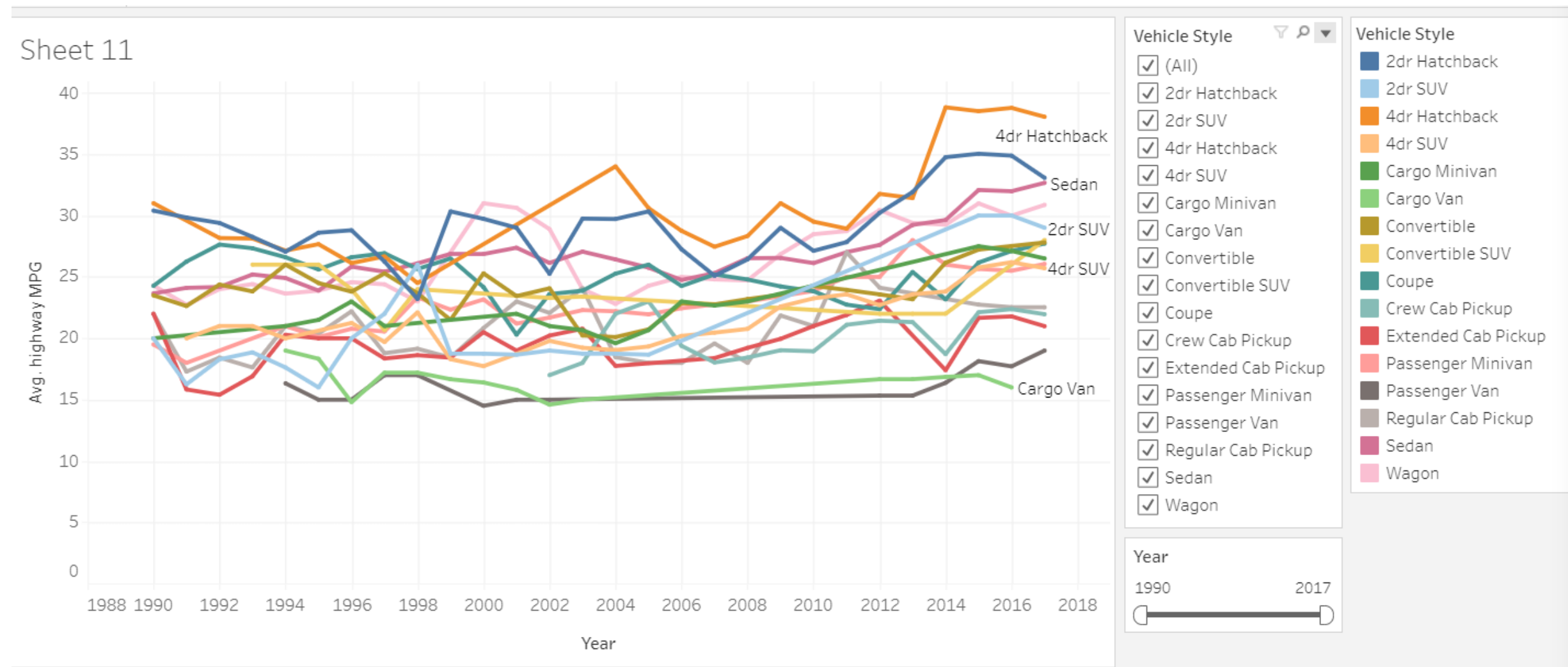
Results :

- This graph shows the average msrp for each transmission type for all body styles.
- Filters are also added to look for the specific value.
- The "Automated/Manual" transmission type with a "Coupe" body style has the highest average Msrp.



Building dashboard

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

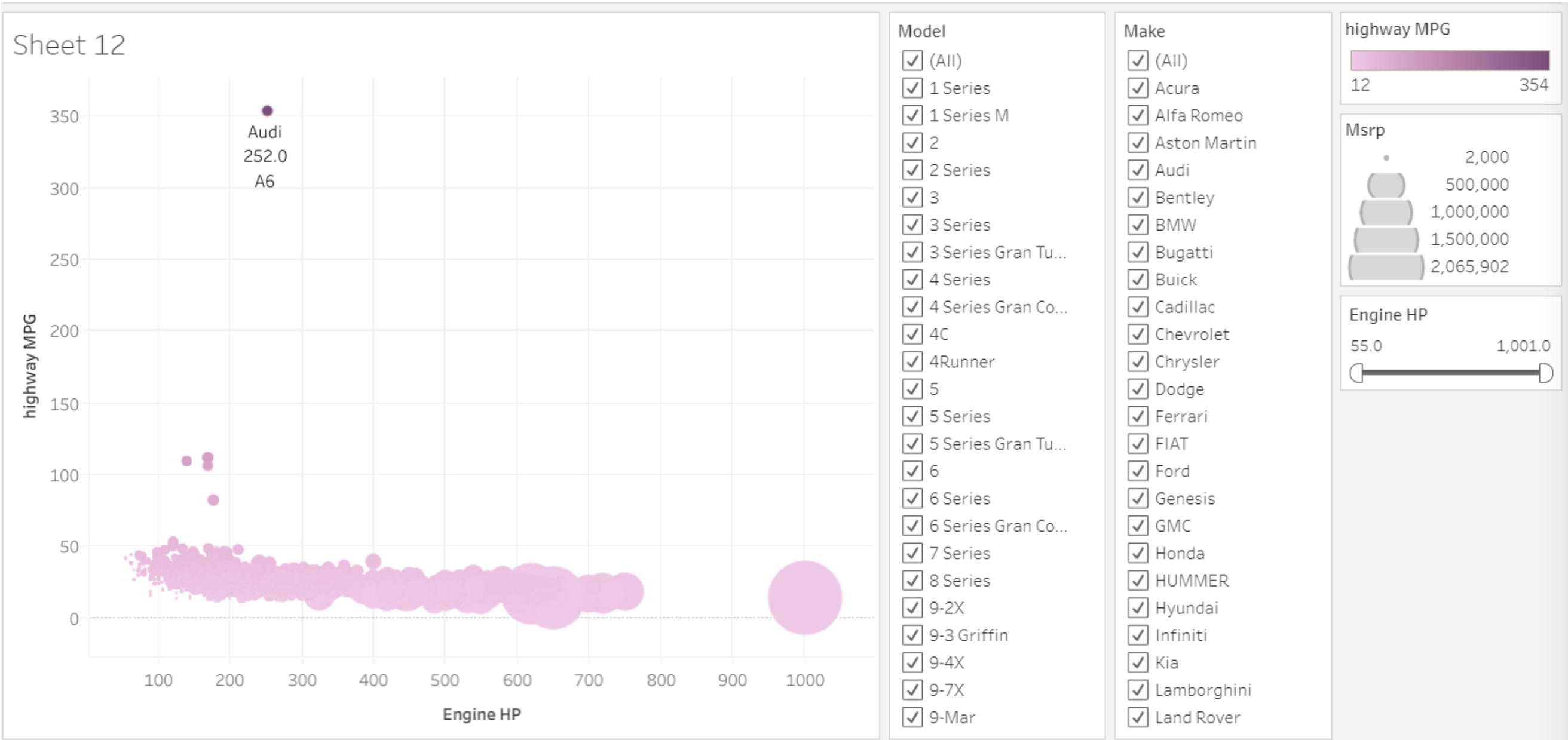


Results :

- This line chart shows the trend of fuel efficiency (MPG) over time for each body style.
- Filters are added for body style and year to observe for any specific body style or year.
- "4dr Hatchback" has the highest fuel efficiency on highways.
- "Cargo Van" has the lowest fuel efficiency on highways.

Building dashboard

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



Results :

- This bubble chart shows how the car's horsepower, MPG (miles per gallon), and price vary across different brands.
- Filters for brands, body style, horsepower, and fuel efficiency(MPG) are added.
- Bubble size represents the price, a bigger bubble means a higher price and a smaller bubble means a lower price.
- The "Bugatti" brand has the highest horsepower value and the highest price as well.
- "Audi" has the highest fuel efficiency on highways.

Tech-Stack Used

Tableau Public

Tableau Public is a free platform to explore, create and publicly share data visualizations online. I have used this as it provides various features to create visually appealing charts.

Insights

Some useful insights were obtained from this analysis for car features and MSRP dataset.

- The highest popular market categories are Hatchback, Flex Fuel; Flex Fuel, Diesel; and Crossover, Flex Fuel, Performance.
- Bugatti and Maybach are the highest-priced brands, while Plymouth and Oldsmobile are the lowest-priced brands.
- Engine HP has the strongest relationship with car price.
- The body style 4dr Hatchback has the highest fuel efficiency on highways.
- Chevrolet brand has the highest total MSRP, and the Sedan body style has the highest total MSRP.
- Vehicles with more number of Engine cylinders typically have lower fuel efficiency on highways.

Results

- ◆ The findings of this analysis highlight the important features which affect the market price of car.
- ◆ This project has helped me understand the power of data visualization, reading plots, and gaining more hands on experience in Tableau.
- ◆ For gaining more insights, statistical tests can also be conducted to measure the significance of the observed relationships.