

Impact of Car Features on Price and Profitability

Project Description

In this project, we will be analysing the impact of car features on price and profitability for various manufacturers. The automotive industry has been rapidly evolving over the past few decades, with a growing focus on fuel efficiency, environmental sustainability, and technological innovation. With increasing competition among manufacturers and a changing consumer landscape, it has become more important than ever to understand the factors that drive consumer demand for cars. We will be doing EDA to find out the relationships between various factors and features that affect the price of a car and help manufacturers make data driven decisions.

Approach

First, we clean the data using appropriate methods to identify missing values and remove/replace them. Then, we will be doing various analyses and making graphs and plots to visualize the insights gained. We will also be making interactive dashboards so that data can be well read and understood by the clients.

Tech-Stack Used

For this project, we used Microsoft Excel 2021 to both analyse and make interactive dashboards.

Insights

The analysis of the dataset gave us various outcomes and insights as follows:

1. Cleaning the data:

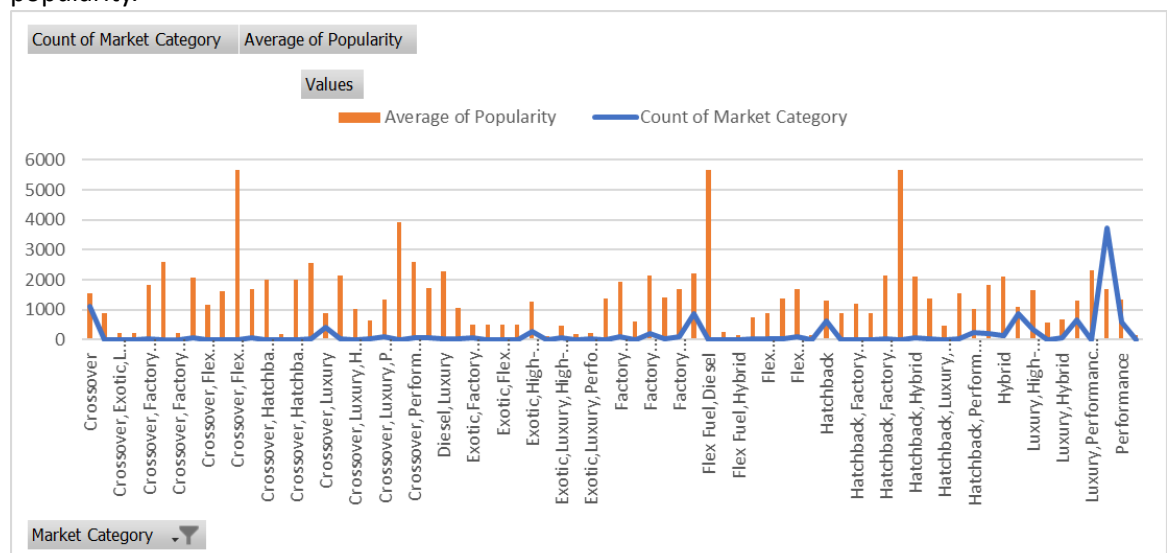
We checked the dataset for missing values and found that there were 69 and 30 values missing from the 'Engine HP' and 'Engine Cylinders' columns respectively.

We replaced the missing values with the mean of the columns i.e., 249 HP and 6 cylinders respectively.

2. 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.

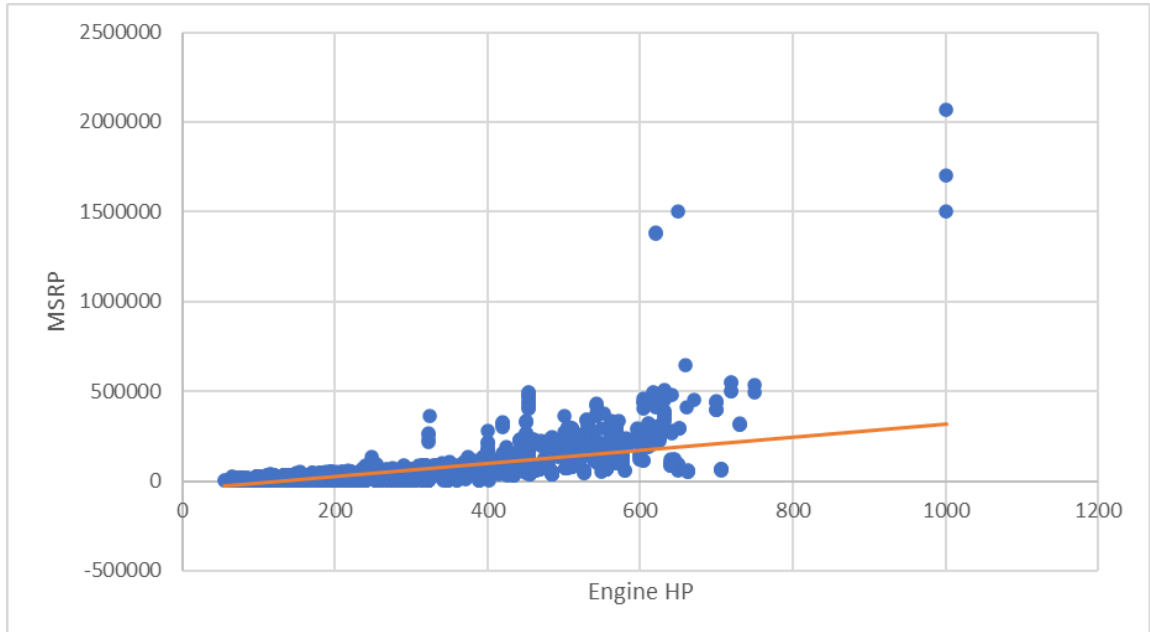
Task 1B: Create a combo chart that visualizes the relationship between market category and popularity.



As we can observe from the graph, the average popularity score of 'Hatchback, Flex Fuel', 'Flex Fuel, Diesel', and 'Crossover, Flex Fuel, Performance' is the highest with 5657 score. Whereas the highest count of market category of car is 'N/A' with 3742 but apart from this 'Crossover' category has the highest count with 1110 number of cars.

3. What is the relationship between a car's engine power and its price?

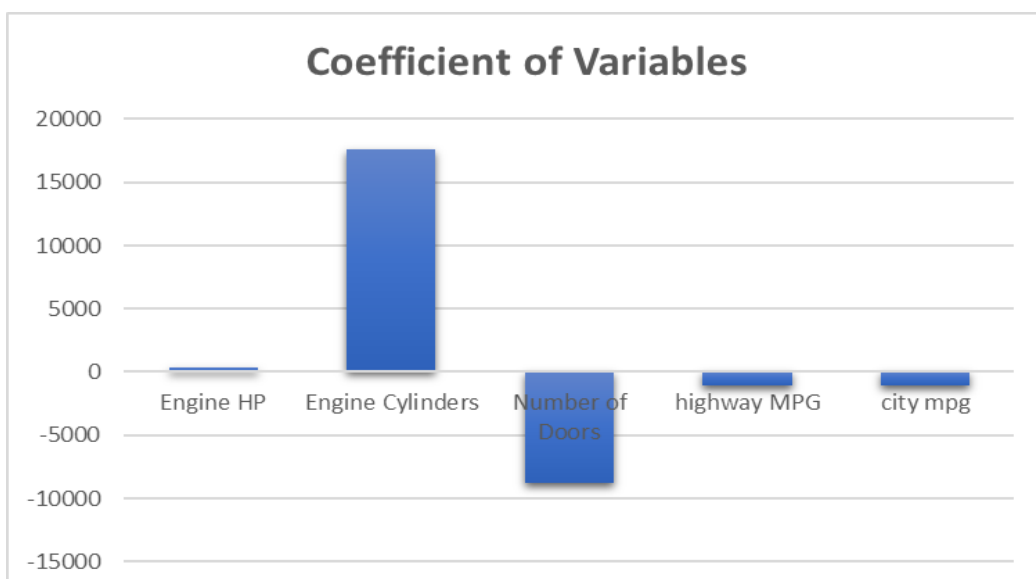
Task 2: Create a scatter chart that plots engine power on the x-axis and price on the y-axis. Add a trendline to the chart to visualise the relationship between these variables.



We can observe that most of the car have a lower HP around 200-600 with most being under 700 HP. We do have exceptions here with Bugatti having 1000HP.

4. 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.

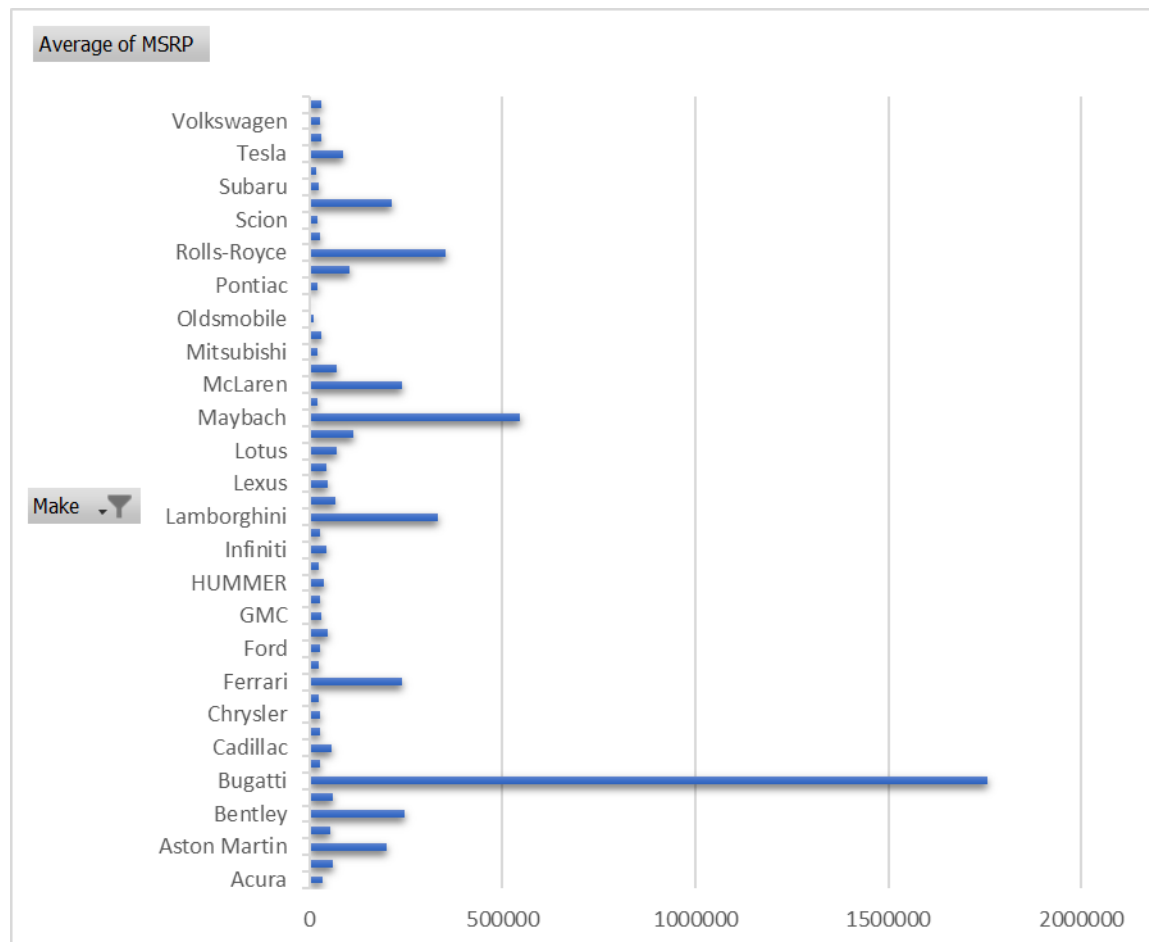


As we can see here, the 'Engine Cylinders' have the highest coefficient of variables in the regression analysis with 17568.96 being the value and 'Number of doors', 'Highway MPG' and 'City MPG' having a negative coefficient.

5. 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 of each manufacturer.

Task 4B: Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between manufacturer and average price.

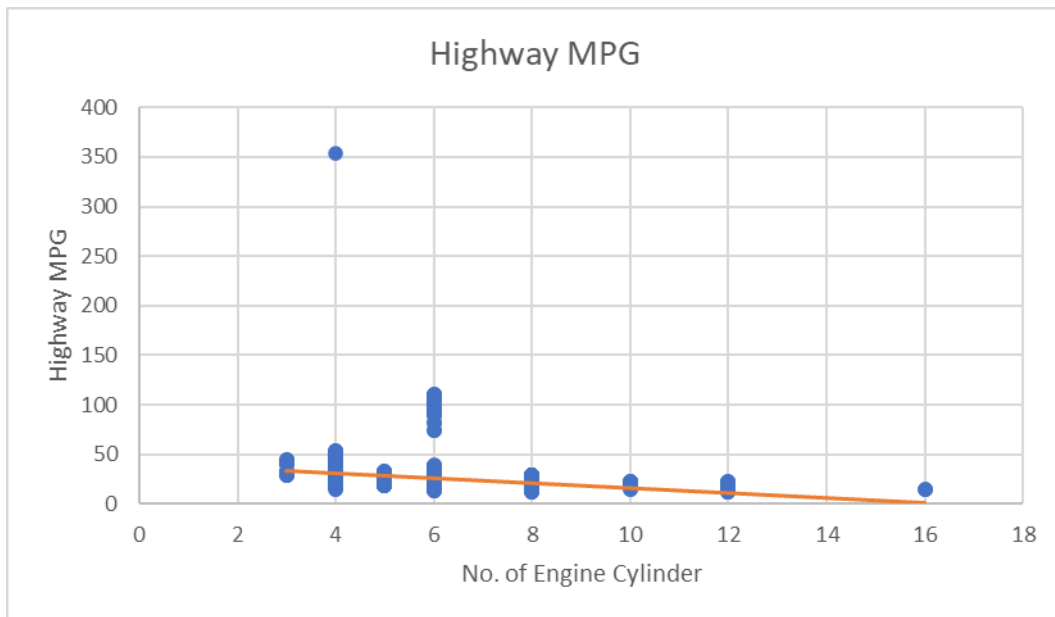


As we can see here, on an average most of the brands have a price range of USD 30,000-50,000. Luxury brands have a price range up to USD 550,000 with the exception of Bugatti which has a huge average price of \$1.7 Million.

6. 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.

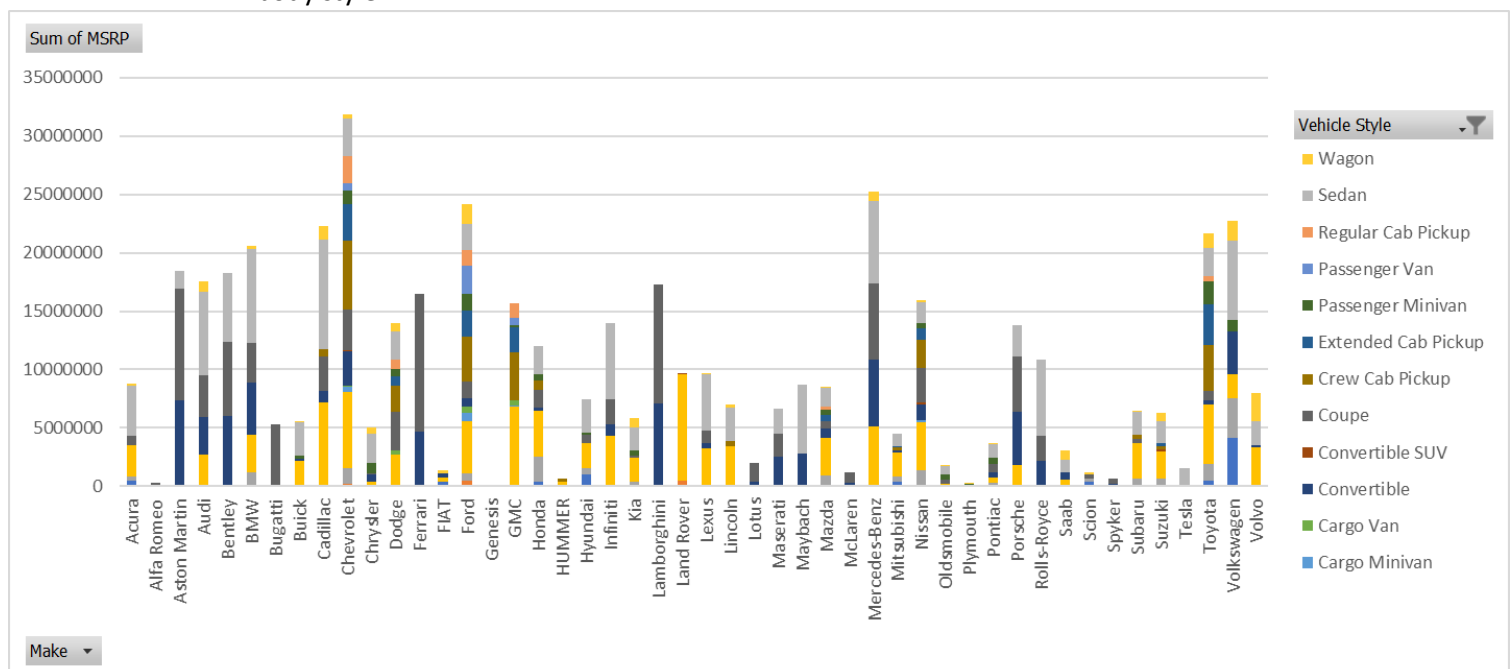


As we can observe, the Highway MPG peaks at 4 cylinders. And as we increase the no. of cylinders it decreases very much.

Building Dashboards

7. How does the distribution of car prices vary by brand and body style?

We used a pivot table to form a table showing the distribution of car prices by brand and body style.



As we can see here that Chevrolet has the highest price distribution and most body type/vehicle styles as well. Whereas, Bugatti has the least number of body types that being one i.e., Coupe. And Genesis has the least price distribution.

Similar to previous dashboard, we made a pivot table to show the distribution of car by brand and body style but with average MSRP this time.

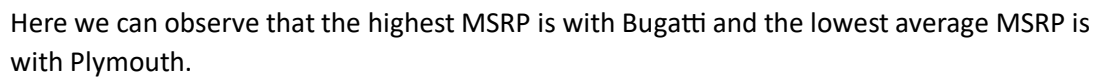


Chart Title

Average MSRP

Legend:

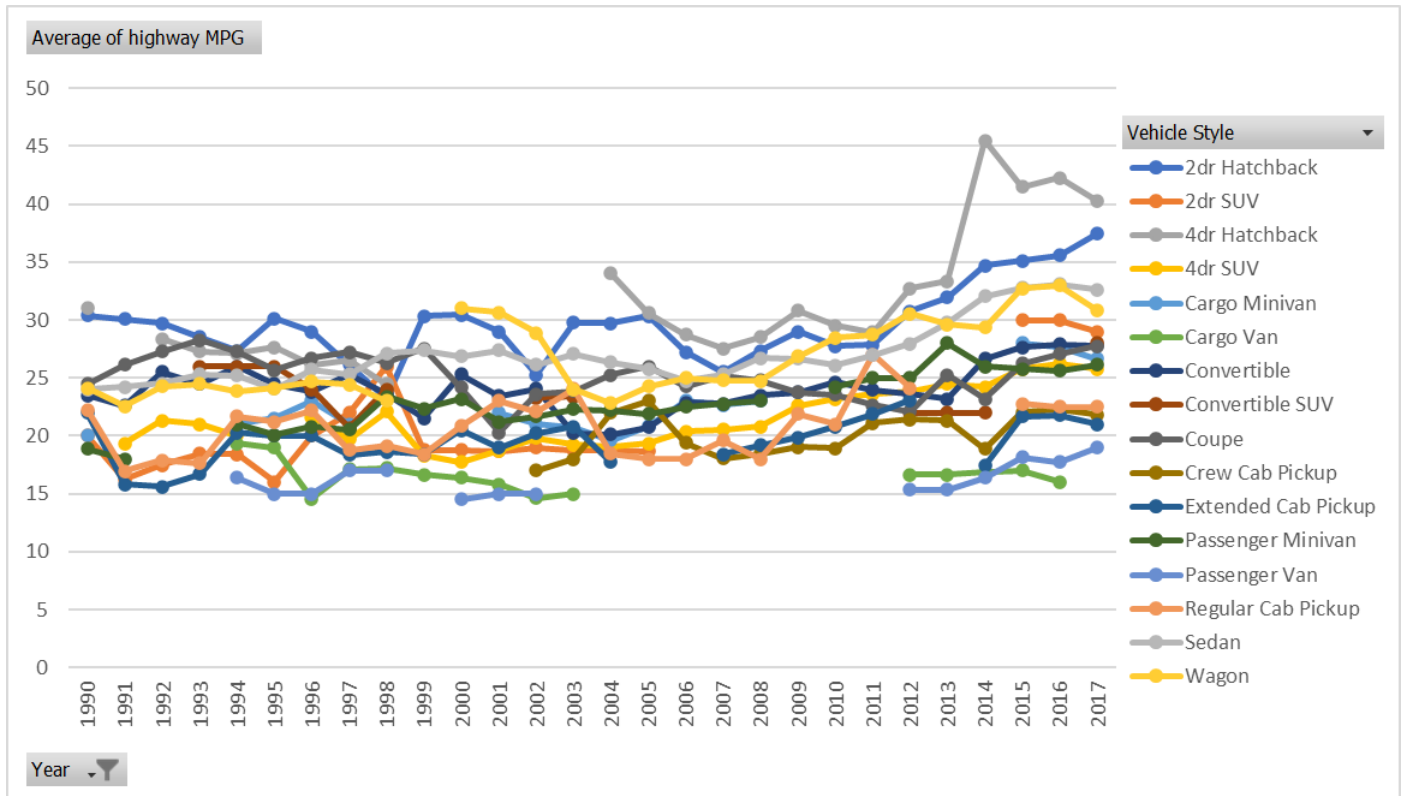
- AUTOMATED_MANUAL
- AUTOMATIC
- DIRECT_DRIVE
- MANUAL
- UNKNOWN

Vehicle Type	AUTOMATED_MANUAL	AUTOMATIC	DIRECT_DRIVE	MANUAL	UNKNOWN
2dr Hatchback	15000	20000	30000	15000	10000
2dr SUV	10000	15000	10000	5000	5000
4dr Hatchback	15000	20000	30000	15000	10000
4dr SUV	15000	20000	30000	15000	10000
Cargo Minivan	15000	20000	30000	15000	10000
Cargo Van	15000	20000	30000	15000	10000
Convertible	15000	20000	30000	15000	10000
Convertible SUV	15000	20000	30000	15000	10000
Coupe	15000	20000	30000	15000	10000
Crew Cab Pickup	15000	20000	30000	15000	10000
Extended Cab Pickup	15000	20000	30000	15000	10000
Passenger Minivan	15000	20000	30000	15000	10000
Passenger Van	15000	20000	30000	15000	10000
Regular Cab Pickup	15000	20000	30000	15000	10000
Sedan	15000	20000	30000	15000	10000
Wagon	15000	20000	30000	15000	10000

As we can observe from the plotted chart above that 'Automated Manual' is the most expensive transmission type and 'Manual' is the least expensive.

10. How does the fuel efficiency of cars vary across different body styles and model years?

We used a pivot table to make a table showing highway MPG across the years for each body style of the cars.

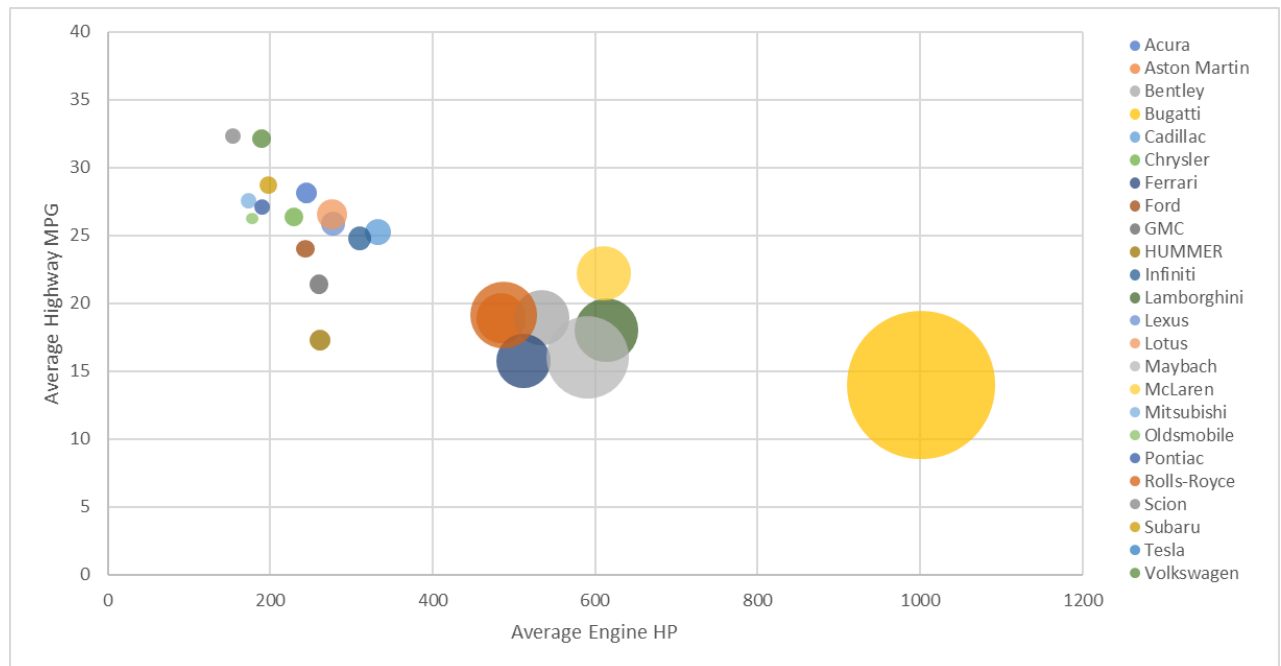


As we can observe here '4dr Hatchback' followed by '2dr hatchback' and 'Sedan' have the highest average highway MPG from 2012 onwards. And 'Cargo van' and 'Cargo Minivan' have the lowest average highway MPG 2012 onwards.

11. How does the car's horsepower, MPG and price vary across different brands?

We used a pivot table to arrange the data by brand for HP, MPG and price and took the data from the pivot table out in a new table to form a bubble chart as shown below.

As we can see below in the plot, the bubbles are plotted with average engine HP on x-axis and average highway MPG on y-axis and the size of the bubble indicates the average price of the brand's cars. Bugatti has the largest bubble size indicating highest average price of the cars. It also has the average highway MPG and highest engine HP. Similarly, Scion has the highest average highway MPG and lowest average engine HP.



Result

From this project we learned a lot of things about the automotive industry and gained various insights revolving various features of the cars. We gained insights on car's features, market category, pricing and identified features and categories most popular among consumers and most profitable for the manufacturer. These insights would help the manufacturers make informed decisions about product development and marketing and help them build cars with features that would be both impactful and profitable.