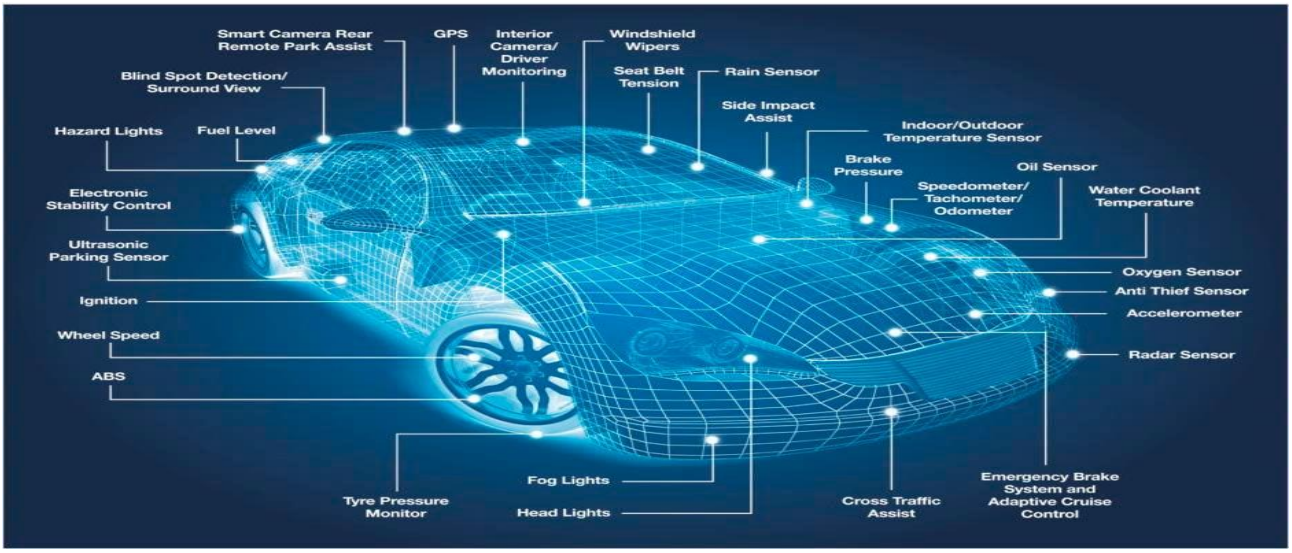
A PROJECT REPORT ON

IMPACT OF CAR FEATURES ANALYSIS



Submitted By: -

Pritesh Kumar Bag

Trainity

C-97, C-97, Ahinsa Cir, Panch Batti, C Scheme,

Ashok Nagar, Jaipur, Rajasthan 302001

AUGUST 29,2023

PROJECT DESCRIPTION

Overview

In recent years, there has been a growing trend towards electric and hybrid vehicles and increased interest in alternative fuel sources such as hydrogen and natural gas. At the same time, traditional gasoline-powered cars remain dominant in the market, with varying fuel types and grades available to consumers. By dissecting vehicle specifications, it unveils insights crucial for manufacturers and consumers alike. It empowers manufacturers to tailor offerings to customer demands, enhancing competitiveness. Car features data analysis offers numerous advantages in understanding market trends, consumer preferences, and technological advancements. For consumers, it facilitates informed decisions based on detailed feature comparisons. This data-driven approach fosters innovation, steering the automotive industry towards safer, more efficient, and technologically advanced vehicles. Ultimately, car features data analysis serves as the compass guiding both manufacturers and consumers towards improved automotive experiences.

This dataset could be useful to derive various data's that will support business decisions, such as:

- Exploring trends in car features and pricing over time
- Comparing the fuel efficiency of different types of cars
- Investigating the relationship between a car's features and its popularity
- Predicting the price of a car based on its features and market category
-

Description of the data sources used in the project

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)
- **Solved Dataset with Dashboard –** https://docs.google.com/spreadsheets/d/1k_Uhp3fm_ZCib2BAv882yntmIRCefOes/edit?usp=drive_link&ouid=109940108743911589080&rtpof=true&sd=true

APPROACH

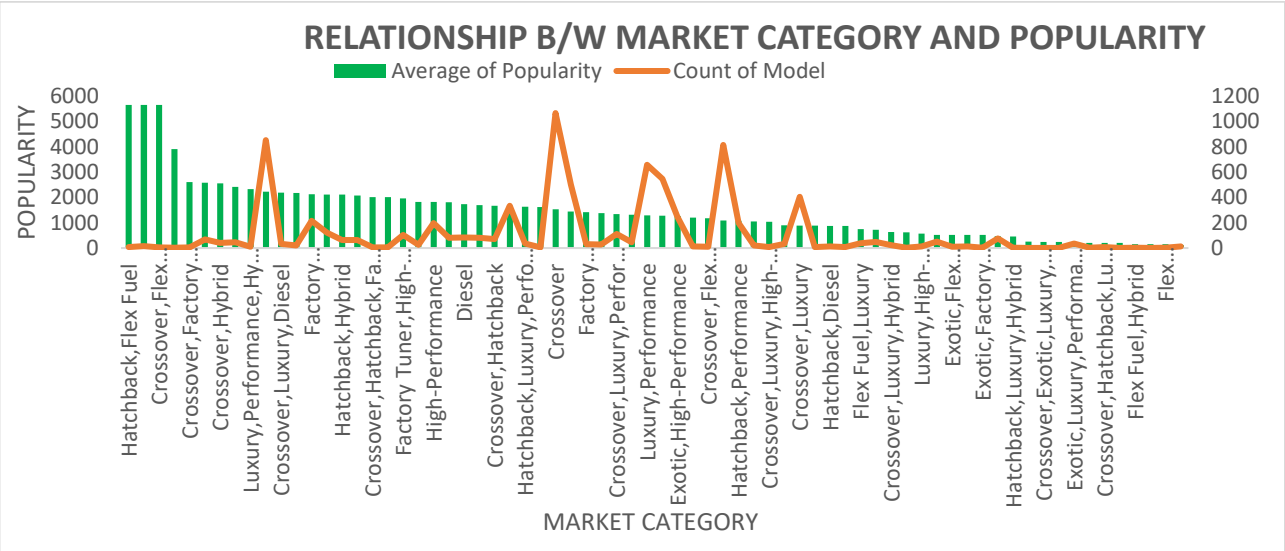
- Downloaded the given datafile
- Understanding the datafile
- Checking the blanks and outliers
- Removing unwanted data
- Drawing summary from the data
- Used formulas, pivot table, charts to build interactive dashboard
-

TECH STACK USED

- Microsoft Excel 2016

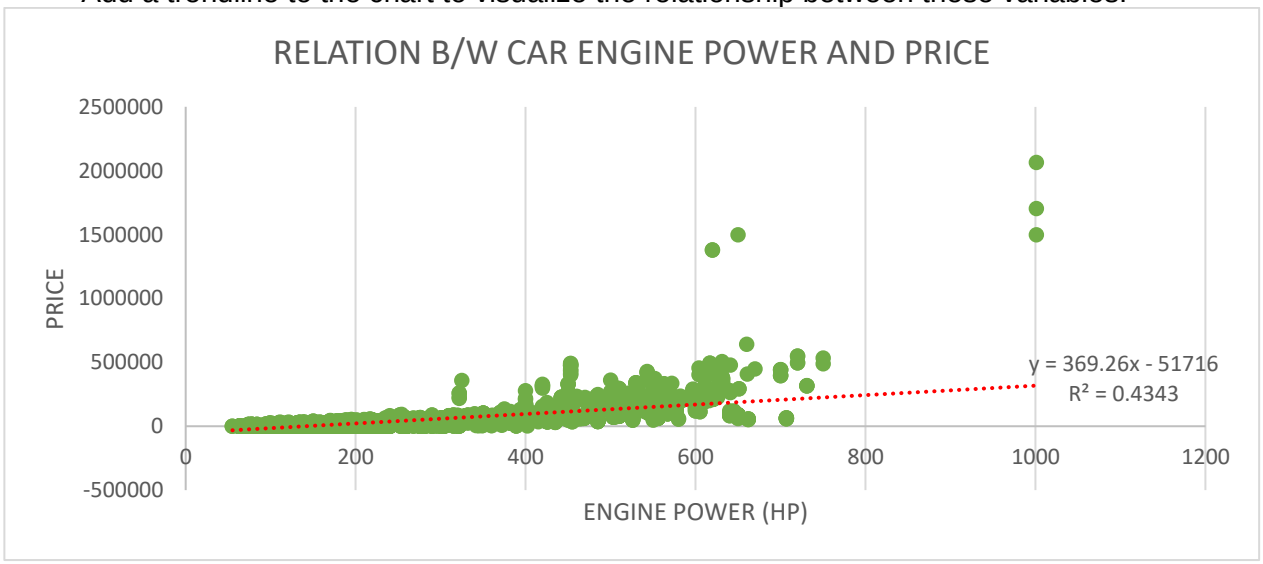
PART-1-INSIGHTS

1. 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.



Interpretation- The crossover market category has highest popularity among customers which includes crossover flex, factory and hybrid.

2. 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.



Interpretation- There lies positive correlation between engine power and price. Equation y=369.26x-51716 shows a positive relation and 43 percent of variability in price can be explained by engine horsepower.

3. 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.

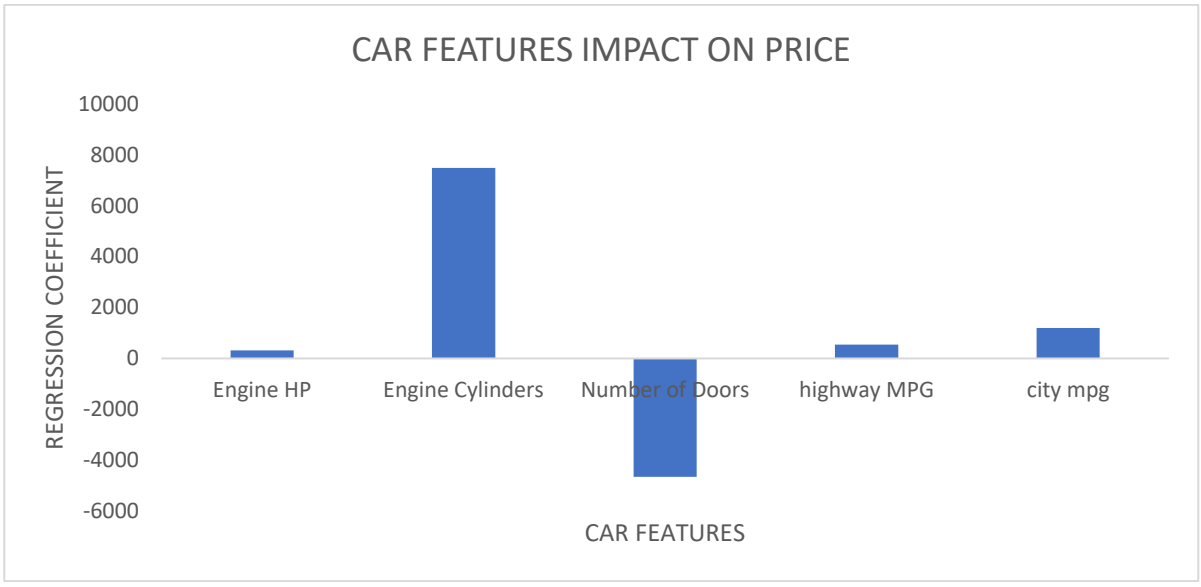
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.678347
R Square	0.460155
Adjusted R Square	0.459912
Standard Error	45366.26
Observations	11097

ANOVA

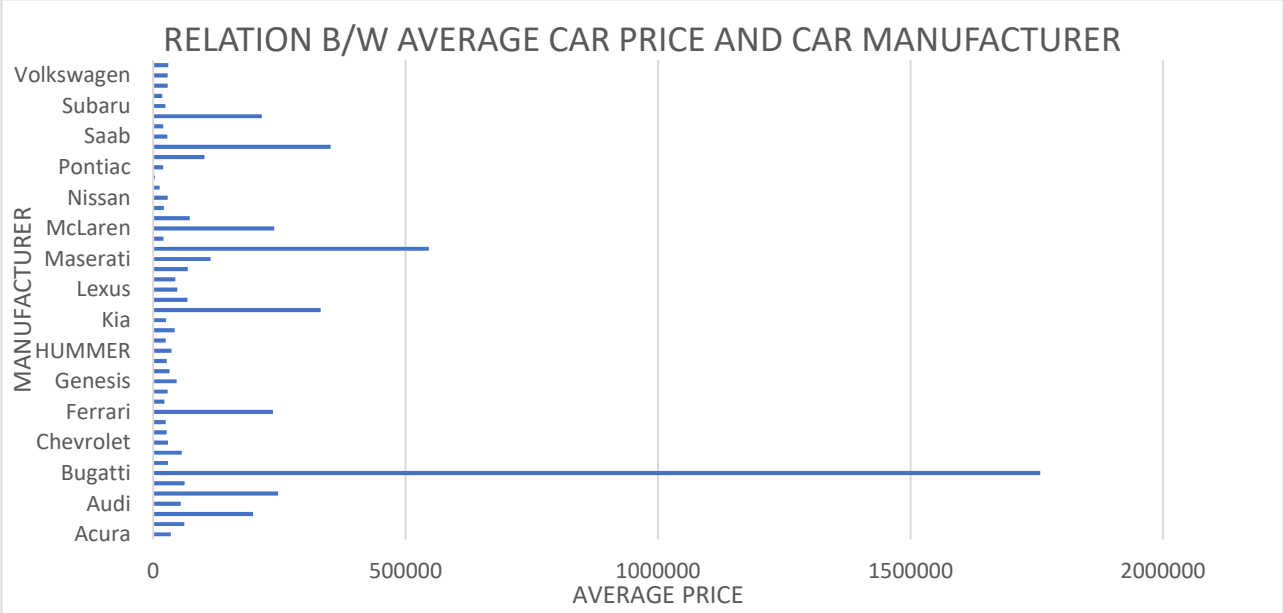
	df	*SS*	*MS*	*F*	*ignificance F*
Regression	5	1.95E+13	3.89E+12	1890.759	0
Residual	11091	2.28E+13	2.06E+09		
Total	11096	4.23E+13			

	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	-102835	3893.782	-26.41	4E-149	-110467	-95202.2	-110467	-95202.2
Engine HP	319.4607	6.417509	49.77955	0	306.8813	332.0402	306.8813	332.0402
Engine Cylinders	7483.326	464.13	16.12334	7.93E-58	6573.548	8393.103	6573.548	8393.103
Number of Doors	-4654.19	498.813	-9.33052	1.25E-20	-5631.95	-3676.42	-5631.95	-3676.42
highway MPG	540.5581	109.9296	4.917312	8.9E-07	325.0765	756.0396	325.0765	756.0396
city mpg	1193.479	126.3628	9.444864	4.27E-21	945.7857	1441.173	945.7857	1441.173



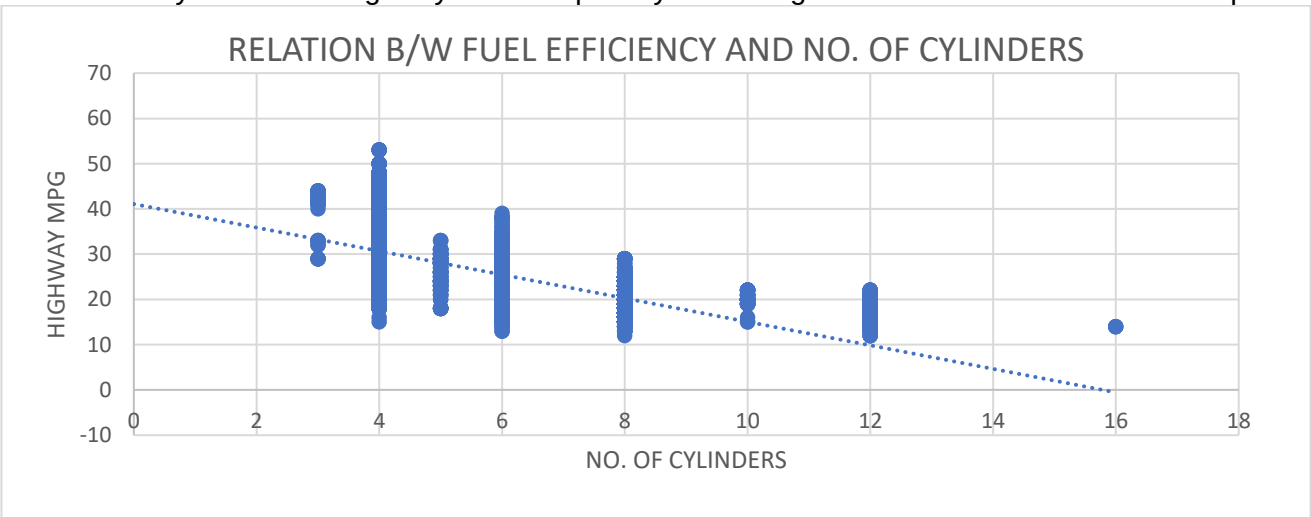
Interpretation- The regression analysis shows that number of doors has negative impact on the price while there is increase in price with features like engine cylinders, engine hp, highway mpg and city mpg.

4. 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



Interpretation- Luxury brands like Buggati, Maybach, Lambhorgini, Mclaren, Ferrari, Aston Martin command high prices as compared to mainstream brands like Honda, Kia and Toyota

5. 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.



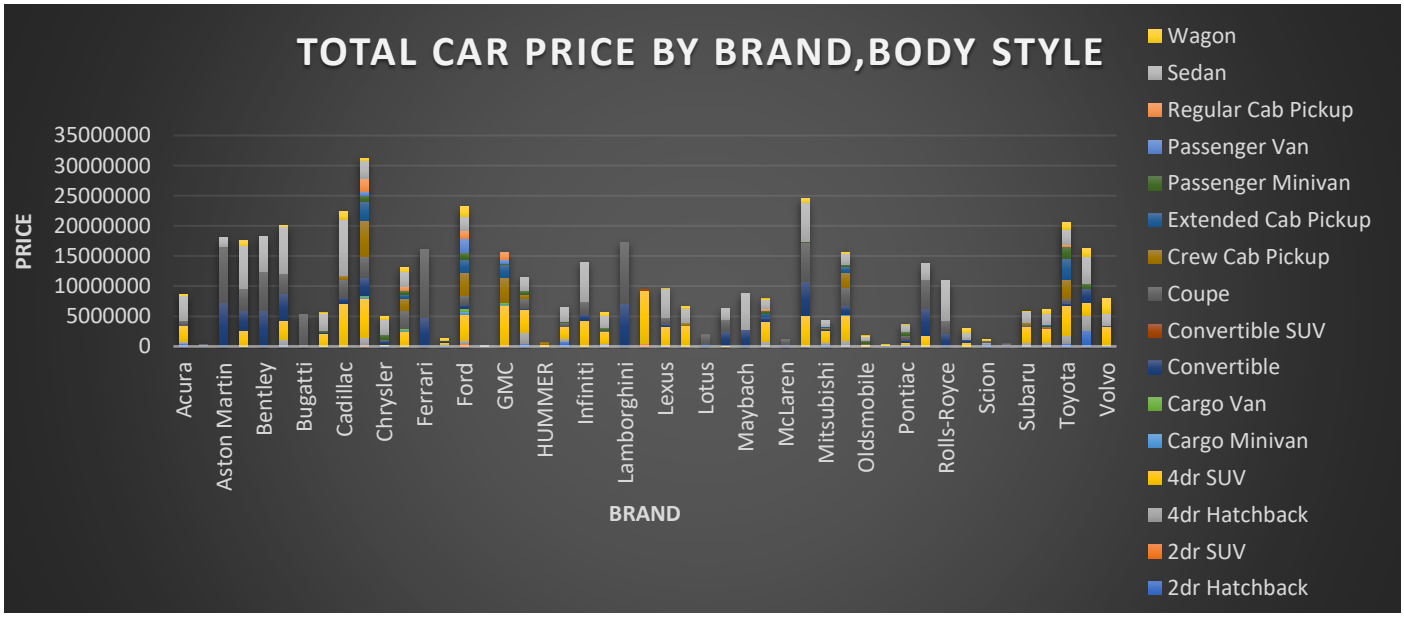
TASKB- CORRELATION COEFFICIENT BETWEEN NO. OF CYLINDERS AND HIGHWAY MPG=CORREL(A:A,B:B)= -0.6147

Interpretation- There is a negative relationship between number of cylinders in the cars engine and its fuel efficiency i.e highway mpg.This can help manufacturers and business to make decisions related to the design of cylinders and increase vehicle efficiency.

PART-2-INSIGHTS- DASHBOARD

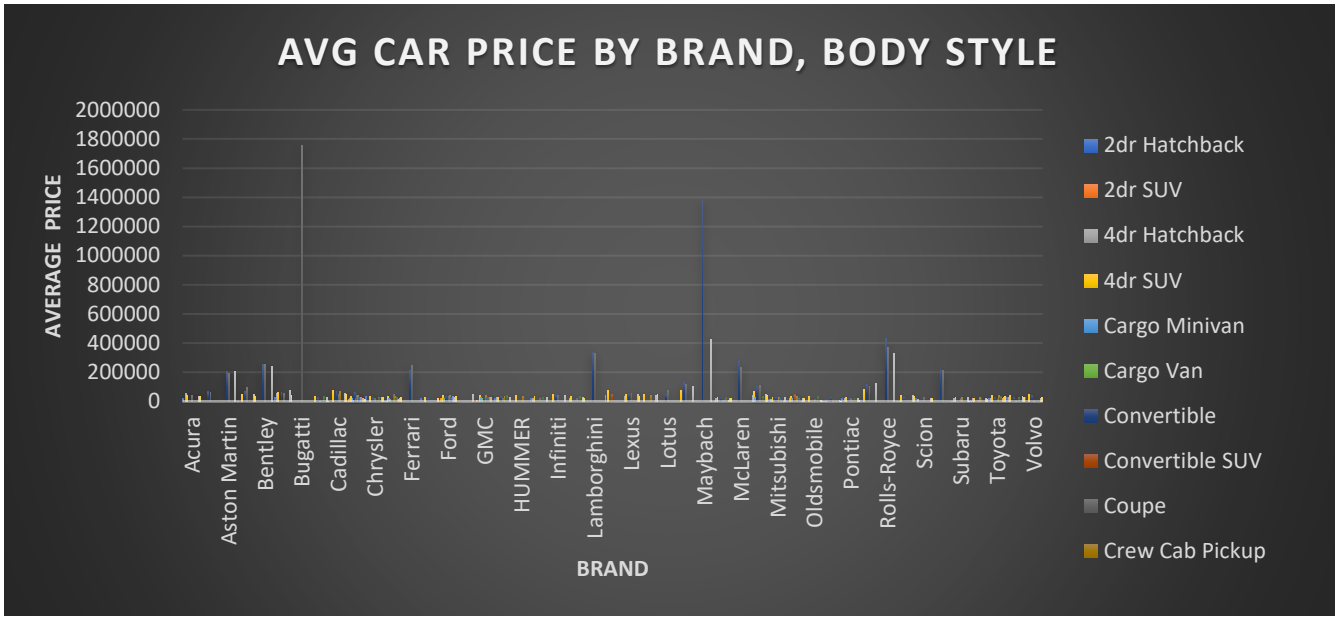
DASHBOARD '!A1

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.



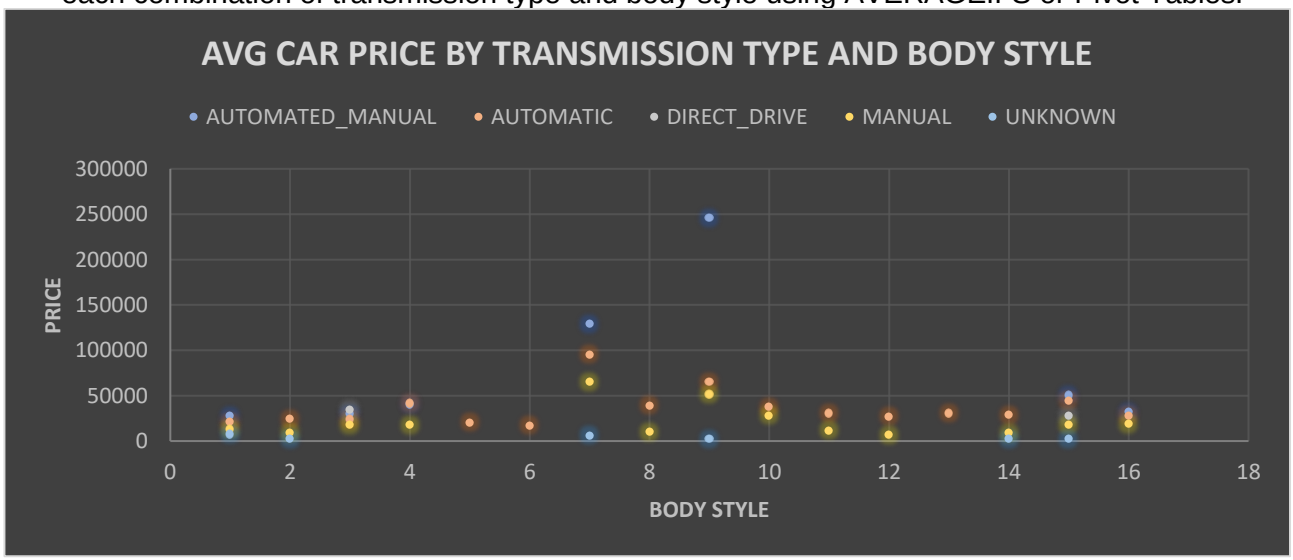
Interpretation- Buggati has the highest car price by brand and body style.

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.



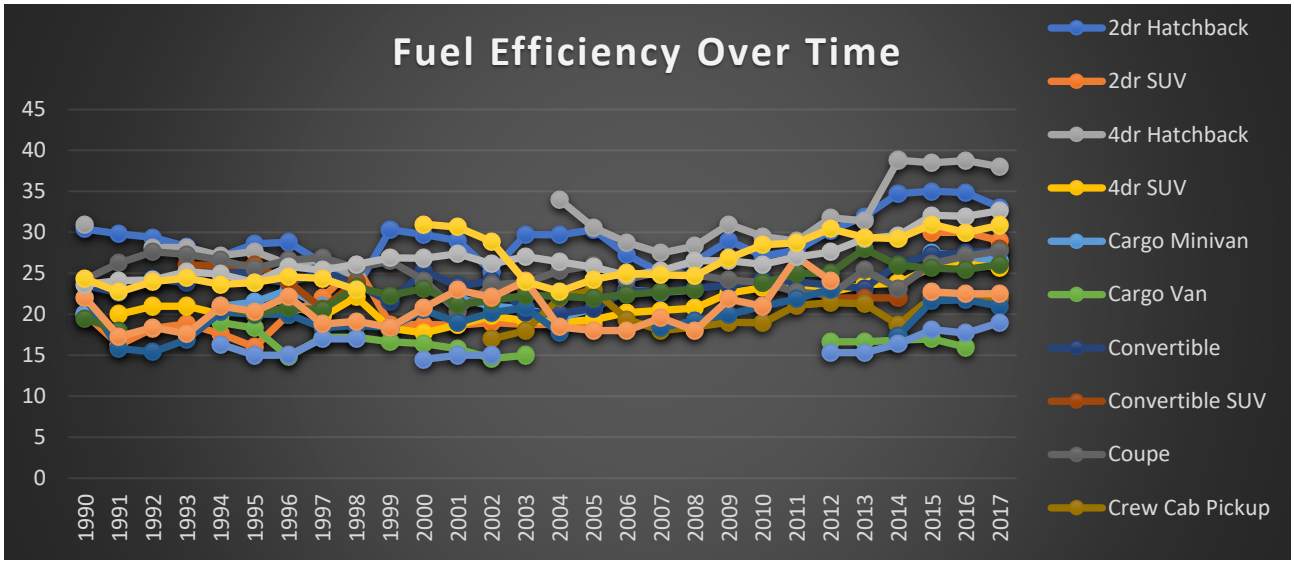
Interpretation- Average price of Buggati and Maybach are high by brand and body style.

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.



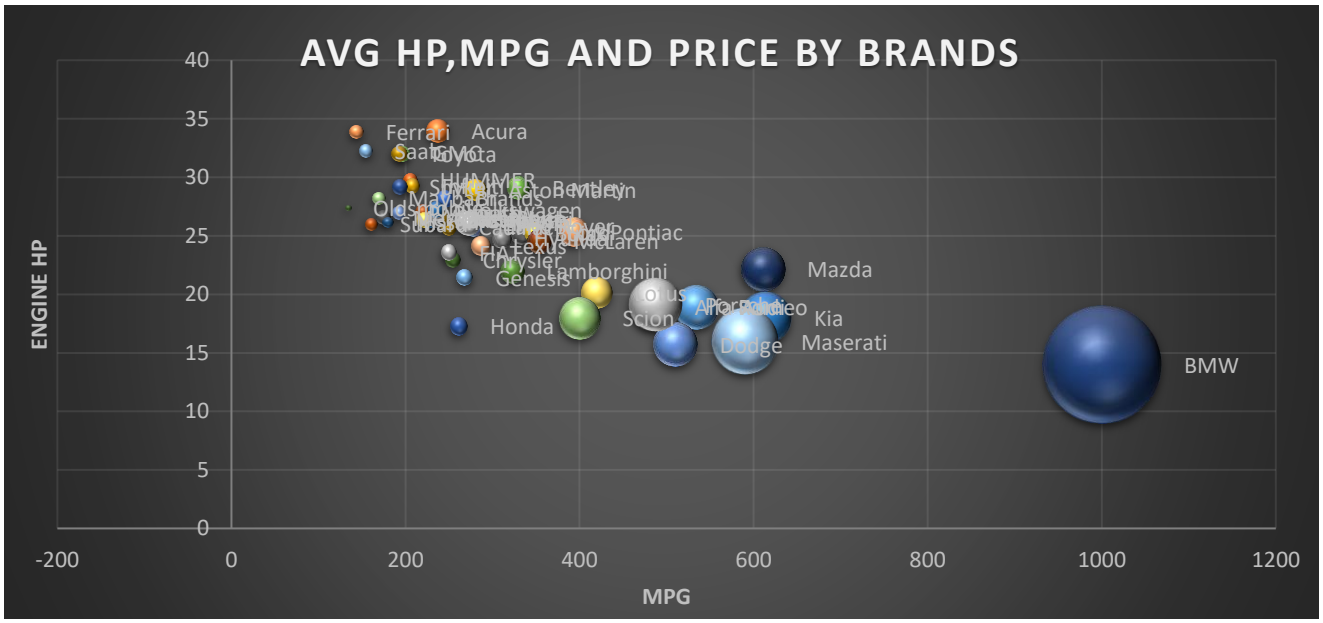
Interpretation- Automated manual with Coupe style is the most expensive category followed by automated manual Convertible body style

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.



Interpretation- The fuel efficiency of most of the categories have increased over time. This can be because of customer demand and governments intervention towards better utilisation of resources and energy.

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 colours 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.



Interpretation- If there is increase in engine price then the highway MPG and engine HP will decrease.

RESULTS

- The crossover market category has highest popularity among customers which includes crossover flex, factory and hybrid.
- There lies positive correlation between engine power and price. Equation y=369.26x-51716 shows a positive relation and 43 percent of variability in price can be explained by engine horsepower.
- The regression analysis shows that number of doors has negative impact on the price while there is increase in price with features like engine cylinders, engine hp, highway mpg and city mpg.
- Luxury brands like Buggati, Maybach, Lambhorgini, Mclaren, Ferrari, Aston Martin command high prices as compared to mainstream brands like Honda, Kia and Toyota
- There is a negative relationship between number of cylinders in the cars engine and its fuel efficiency i.e. highway MPG. This can help manufacturers and business to make decisions related to the design of cylinders and increase vehicle efficiency
- Buggati has the highest car price by brand and body style
- Average price of Buggati and Maybach are high by brand and body style
- The fuel efficiency of most of the categories have increased over time. This can be because of customer demand and governments intervention towards better utilisation of resources and energy.
- Automated manual with Coupe style is the most expensive category followed by automated manual Convertible body style
- If there is increase in engine price then the highway MPG and engine HP will decrease.