Analyzing the Impact of Car Features on Price and Profitability

Project Description: 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.

For the given dataset, as a Data Analyst, the client has asked How can a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer demand?

It was required to build an interactive dashboard in Excel from the tasks given below:

1. How does the popularity of a car model vary across different market categories?

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

3. Which car features are most important in determining a car's price?

4. How does the average price of a car vary across different manufacturers?

5. What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

Now for the Next portion of the Project, it was required to create the Interactive Dashboard. The client has requested these questions given below:

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

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

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

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

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

Approach: First I gone through dataset to know all the columns present in the table. Then I saw all the questions and thought of functions which could be used to answer each question. After that I applied those functions and found the answer to each question and plotted the graph wherever was required. I started with cleaning the data set using following steps:

- First, I found all the rows with blank cells. For this first, I selected whole table. After that, I selected go to special option inside find & select menu. Inside the menu I selected blanks option and all the blank cells were selected. Then I searched for the missing values on internet and filled those values accordingly.
- After that, I deleted rows in which value of missing data was not available on internet.
- Lastly, I removed all the duplicate rows.

Link for working excel file:-

https://docs.google.com/spreadsheets/d/1pPKrBcpqESWNod4xyNoGvjoDHRkvEFKr/edit?usp=sharing&ouid=106942457558004201317&rtpof=true&sd=true

Tech-Stack Used: The software used for the project is Microsoft Excel 365. It is used to run the functions and get answers of each question. It is also used to plot the graphs.

Insights:

Analysis

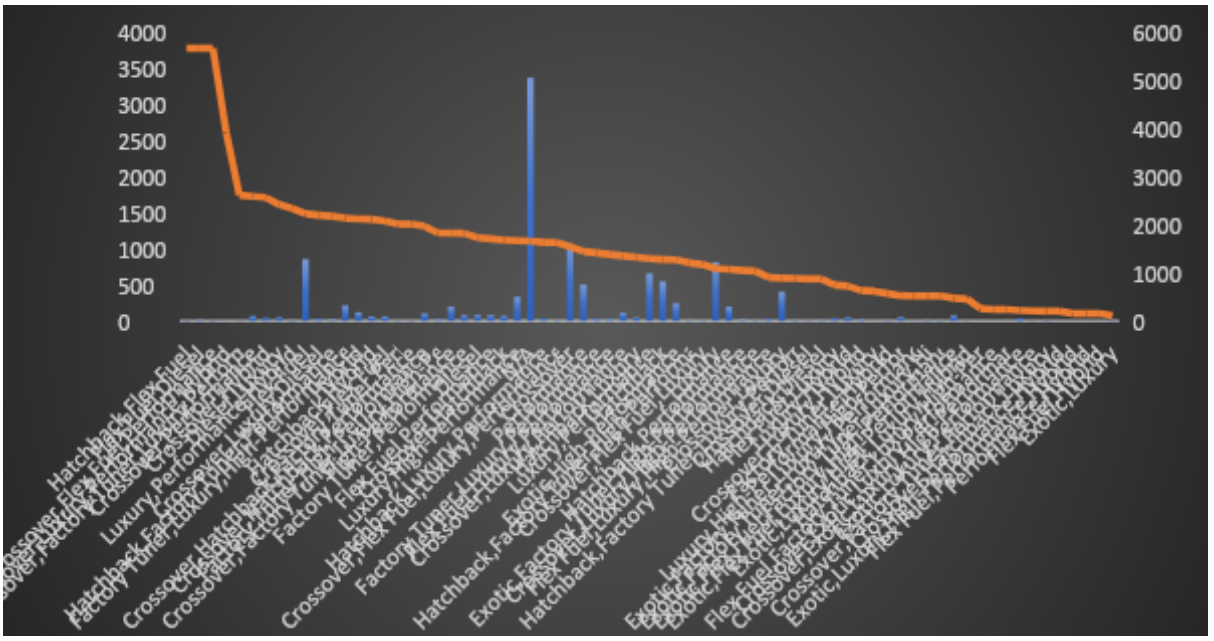
1. How does the popularity of a car model vary across different market categories?

   Function:-

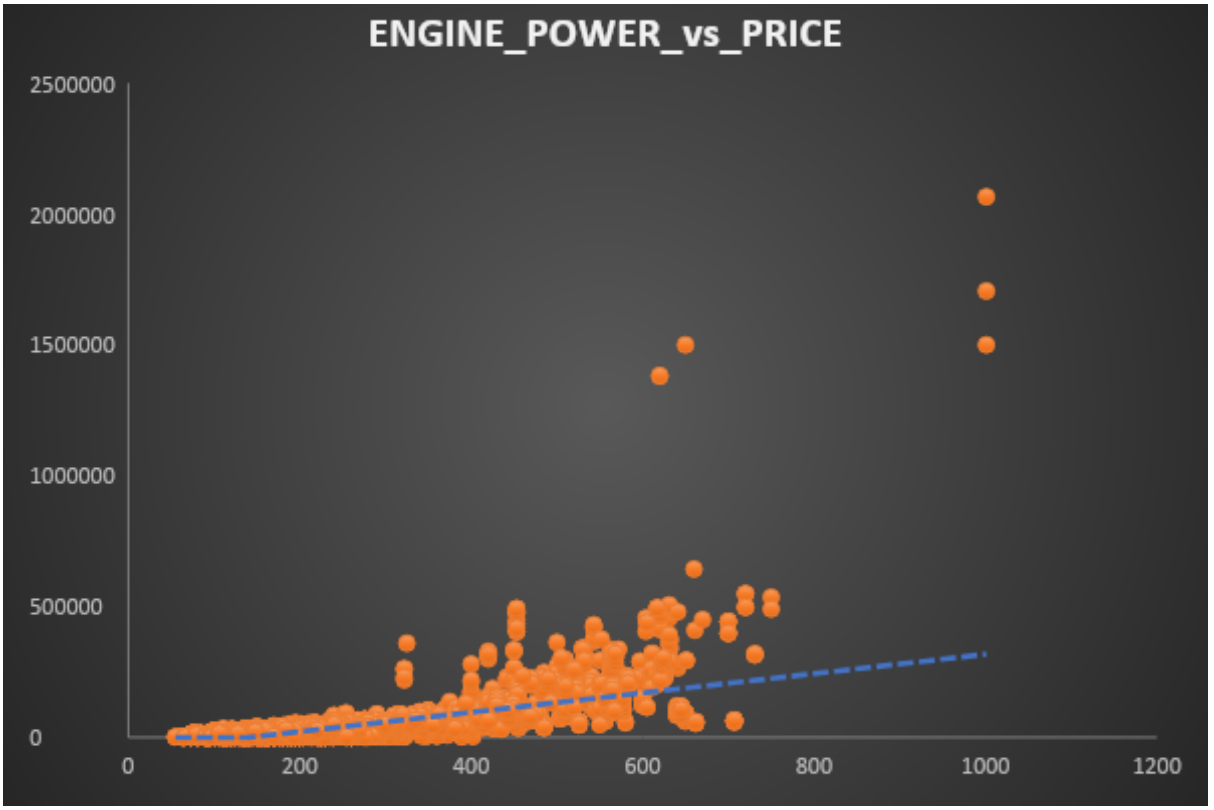
   I created the pivot table with market category in column and count of model and average of popularity in value.

   After that I created the graph for that pivot table

   Output:-



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



3. Which car features are most important in determining a car's price?

   Output:-

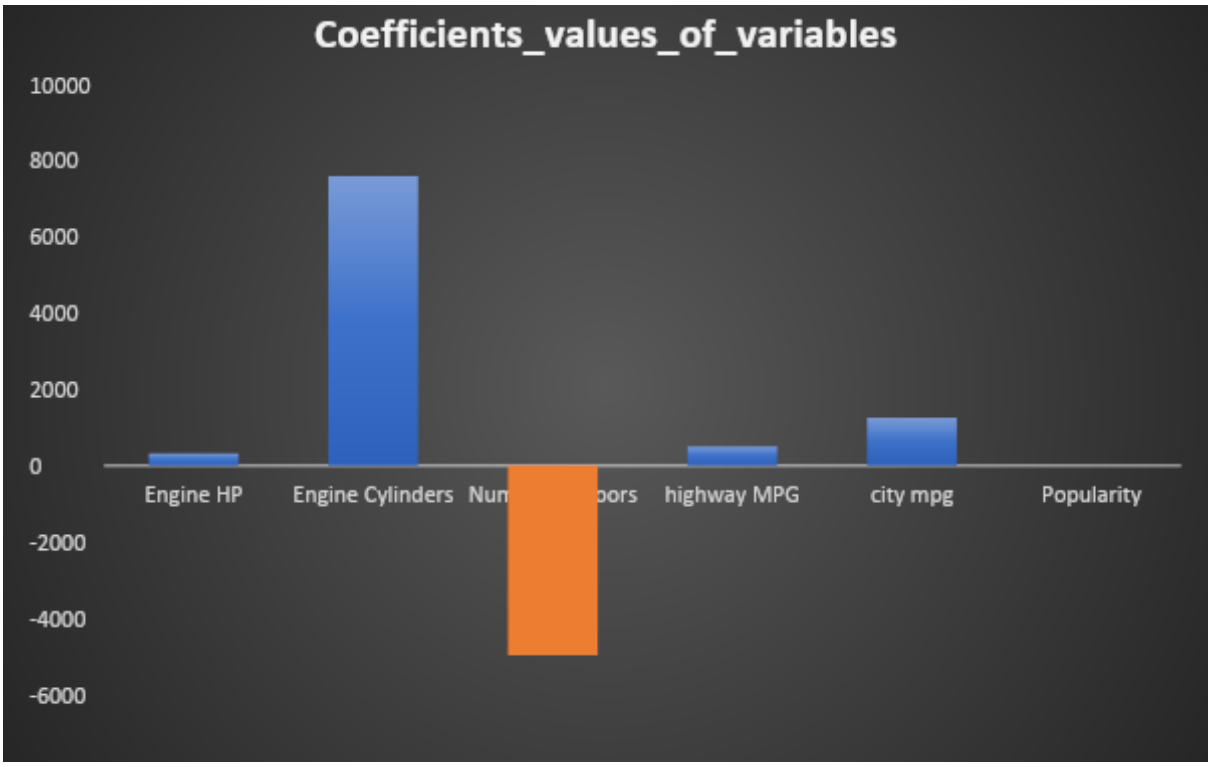
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.683398734
R Square	0.46703383
Adjusted R Square	0.466745558
Standard Error	45073.11516
Observations	11100

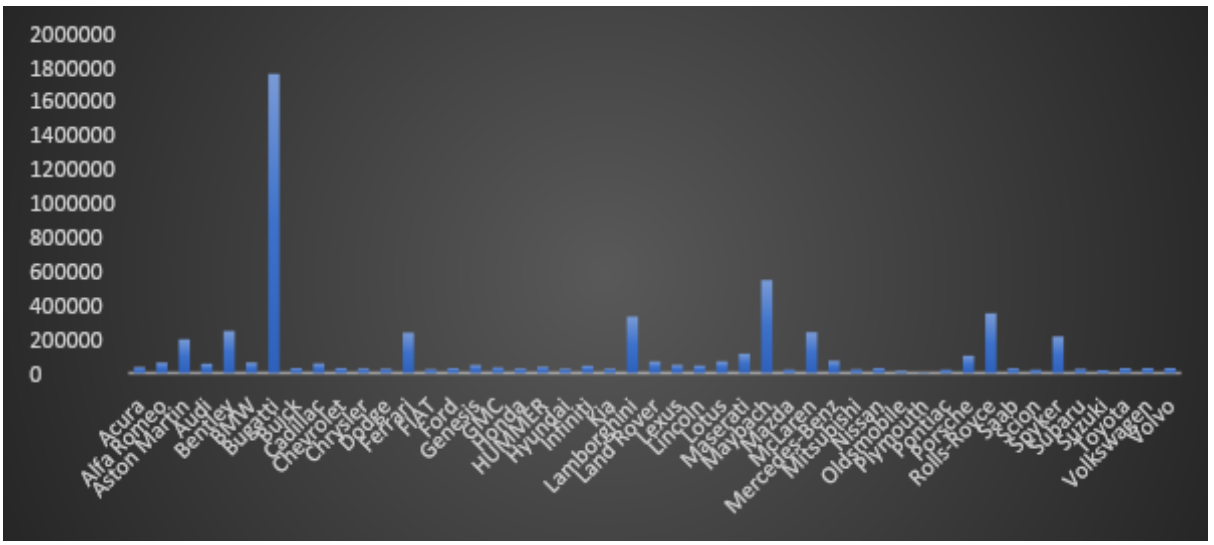
ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	1.97484E+13	3.29141E+12	1620.117298	0
Residual	11093	2.25364E+13	2031585710		
Total	11099	4.22848E+13			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-97172.91379	3897.538609	-24.93186689	1.5283E-133	-104812.7827	-89533.0449	-104812.7827	-89533.0449
Engine HP	320.437386	6.374182606	50.27113369	0	307.9428544	332.9319176	307.9428544	332.9319176
Engine Cylinders	7581.861491	461.1014751	16.4429348	4.82527E-60	6678.020588	8485.702393	6678.020588	8485.702393
Number of Doors	-4978.61342	496.3152295	-10.03115182	1.40155E-23	-5951.479544	-4005.747295	-5951.479544	-4005.747295
highway MPG	503.8302762	109.2603646	4.6112813	4.04669E-06	289.6605284	718.000024	289.6605284	718.000024
city mpg	1253.15324	125.642737	9.973940953	2.48501E-23	1006.871129	1499.435352	1006.871129	1499.435352
Popularity	-3.554458264	0.297295546	-11.95597549	9.59937E-33	-4.137210412	-2.971706117	-4.137210412	-2.971706117



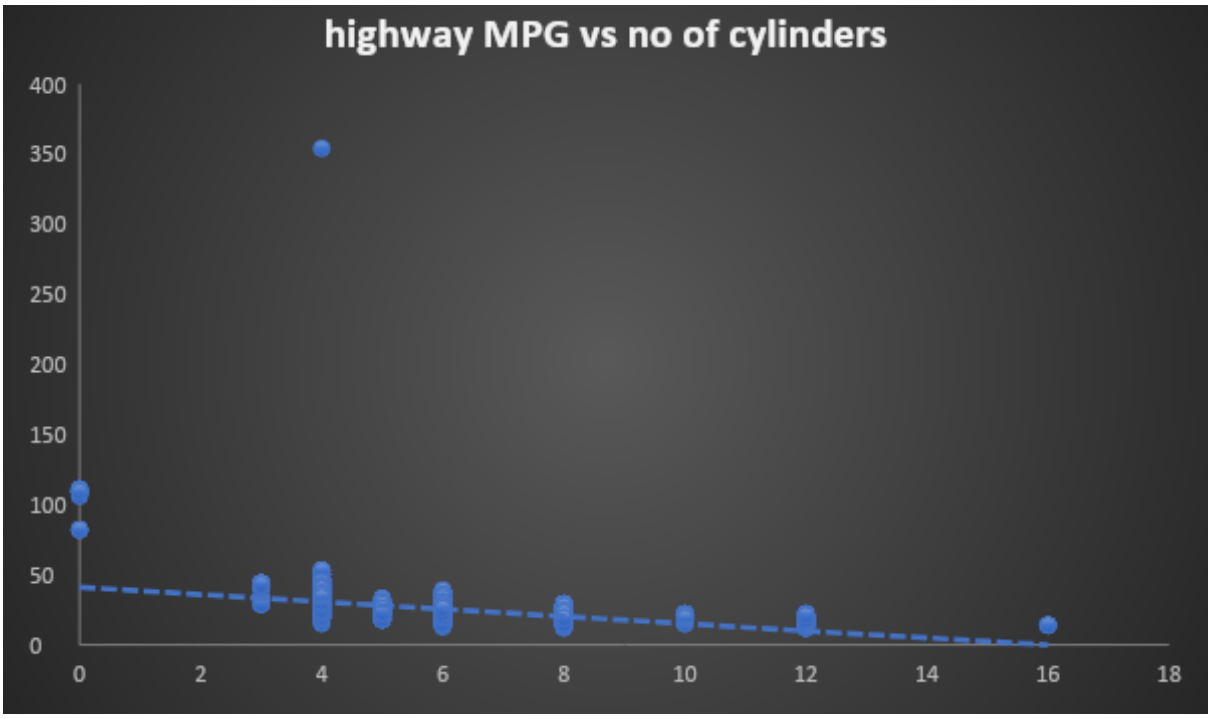
4. How does the average price of a car vary across different manufacturers?



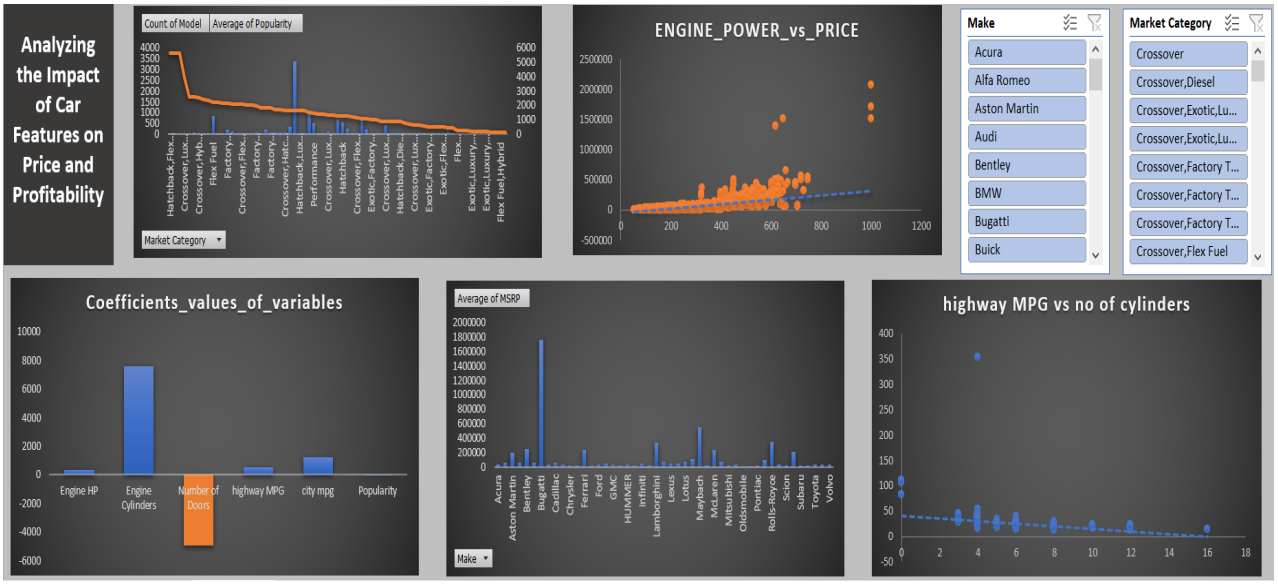
5. What is the relationship between fuel efficiency and the number of cylinders in a car's engine?

   Function:-
   =CORREL(A2:A11101,B2:B11101)
   Output:-

   -0.61471

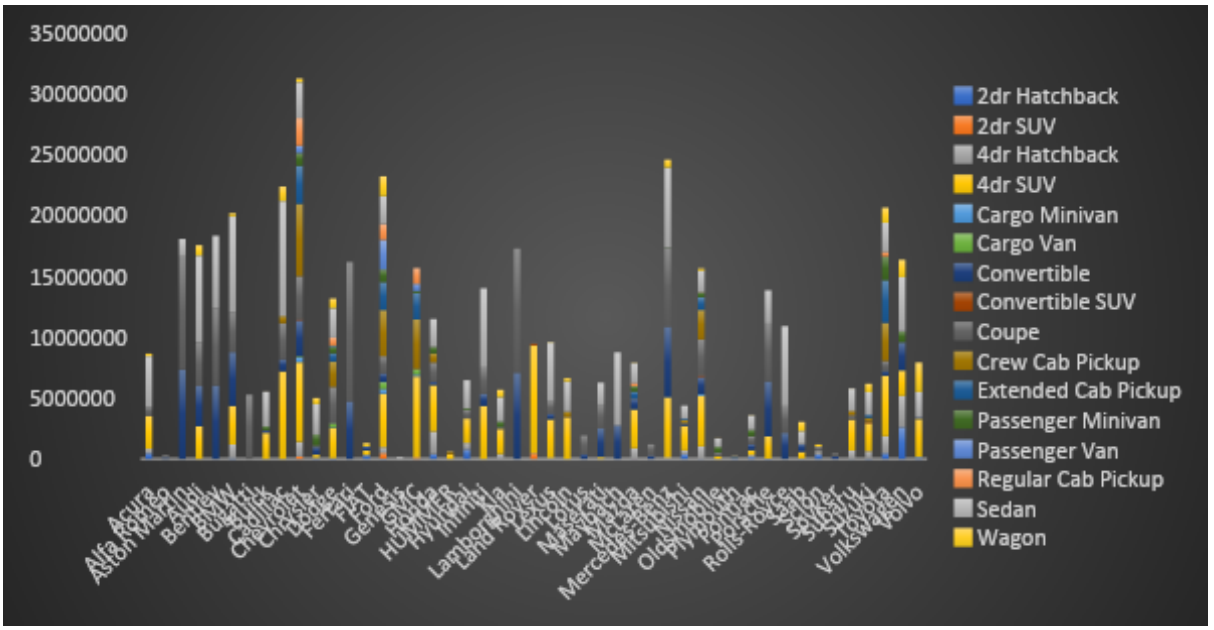


Dashboard

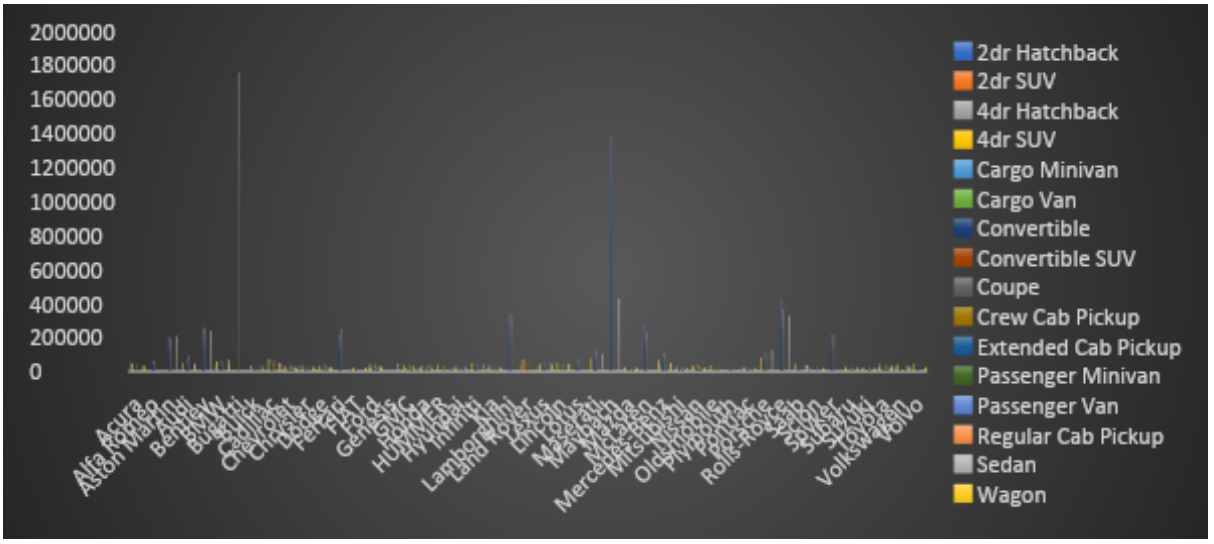


Building the Dashboard:

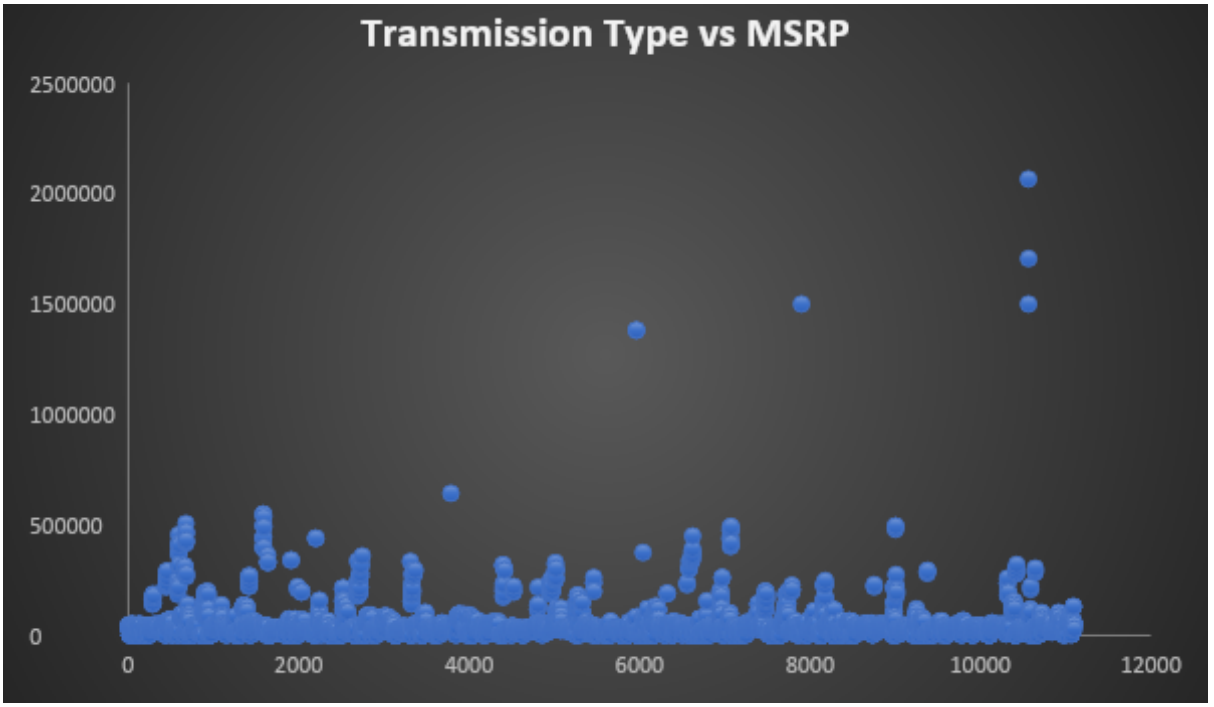
1. How does the distribution of car prices vary by brand and body style?



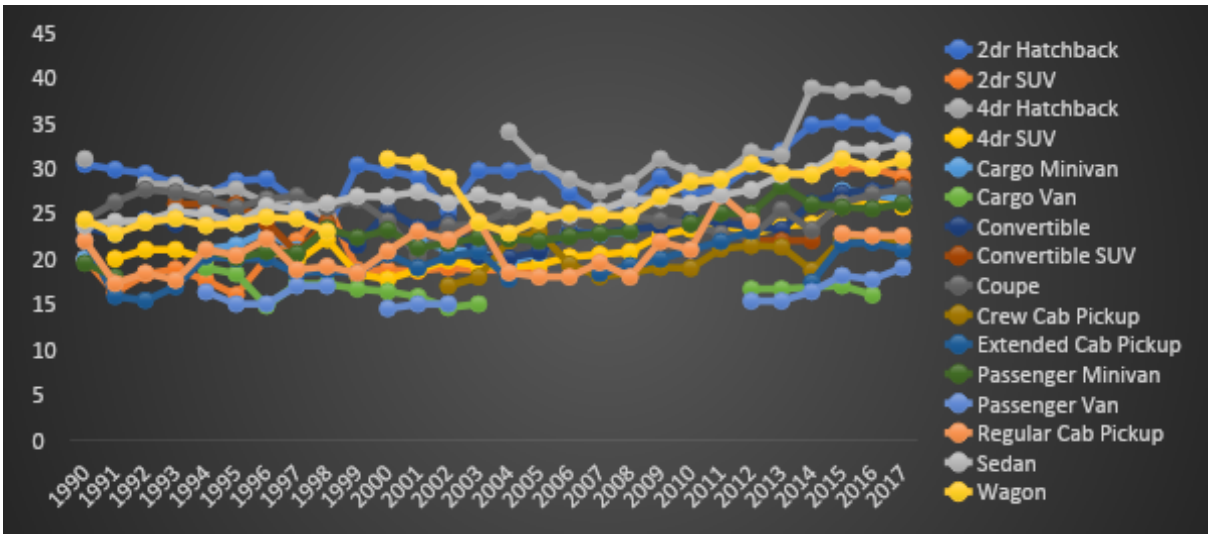
2. Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?



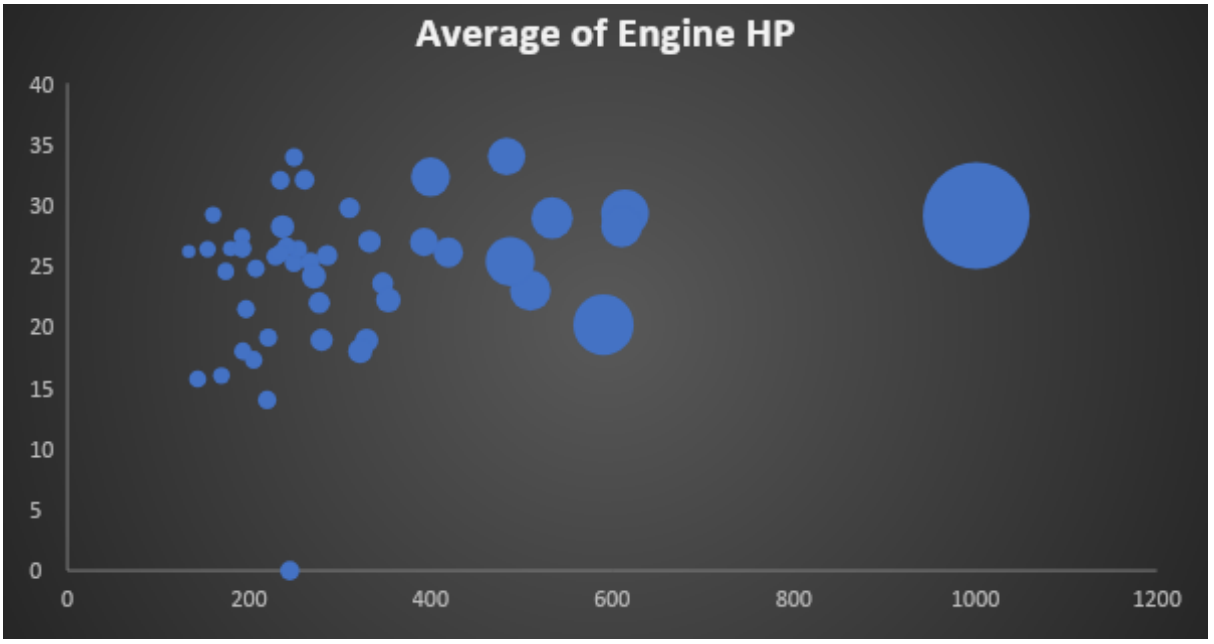
3. How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?



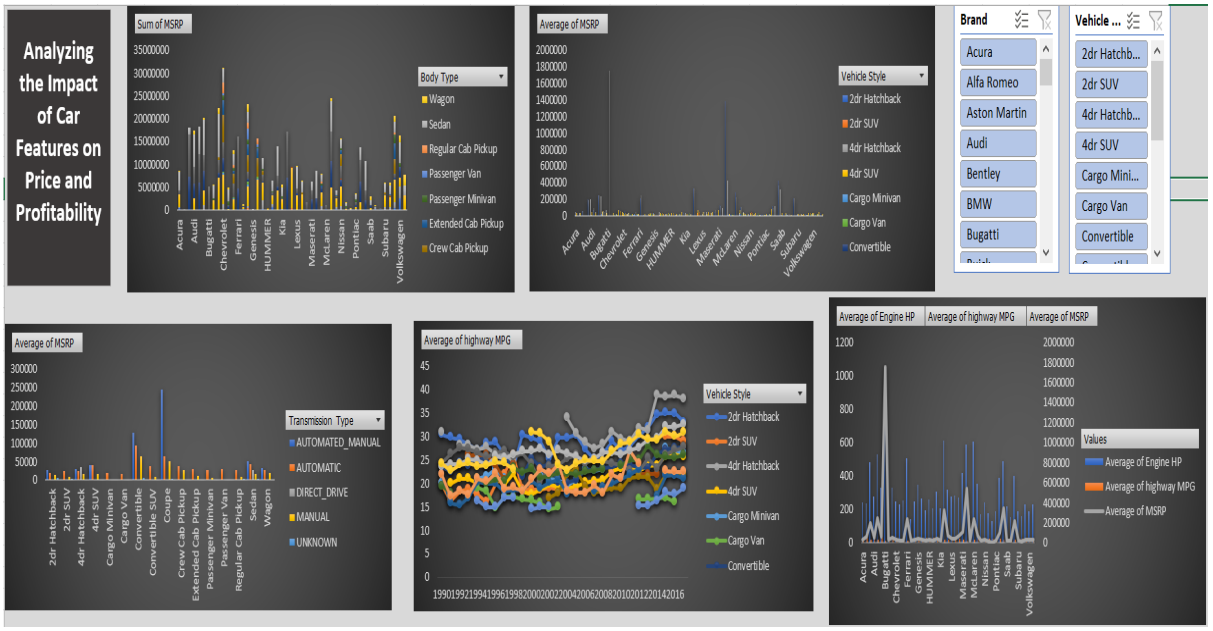
4. How does the fuel efficiency of cars vary across different body styles and model years?



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



Dashboard



Results:

Cleaning the data: Cleaned data consists of 11101 rows including title.

- If engine power increases price will also increase
- Bugatti has the highest average price
- Automatic_manual is most expensive and most popular category
- If engine hp increase price will also increase but highway mpg will decrease.
- Chevrolet has the highest price distribution