ANALYZING IMPACT OF CAR FEATURES ON PRICE AND PROFITABILITY

Project Description:

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.

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.

ROLE: 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?

Things we have to find out through the case study:

- How the popularity of a car model vary across different market categories
- the relationship between a car's engine power and its price
- car features are most important in determining a car's price
- the average price of a car vary across different manufacturers
- the relationship between fuel efficiency and the number of cylinders in a car's engine

Approach:

Problem Statement:

How can a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer demand? Identification of such relationships **using REGRESSION ANALYSIS** is the aim of this case study.

Analysis Approach:

- Downloading the dataset (Car_data)
- Identifying the missing values and dealing with it.
- Analyzing the relationship between a car's features, market category, and pricing.
- Identifying which features and categories are most popular among consumers.
- Finally, Data Visualization with the help of Charts, Graphs.

Tech Stack Used:

I used

- **Microsoft® Excel® 2019 MSO (Version 2304 Build 16.0.16327.20200) 64-bit** which enables us to Clean, Format, Organize and Calculate the data in a spreadsheet.
- **Ms Word 2019** for the preparation of the document to be presented.

Insights:

Based on the achieved results I can conclude the following things:

- Crossover, Flex Fuel, Performance & Flex Fuel, Diesel & Hatchback, Flex Fuel are the Market categories with Highest average Popularities.
- Crossover is the market category with highest number of car models.
- There is a linear relationship between Car's Engine power and it's Price.
- Engine HP is strongly correlated with the Car's Price.
- Among all the Manufacturers, Bugatti has the Highest Average Price.

Results: The detailed answers to the questions are below:

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.

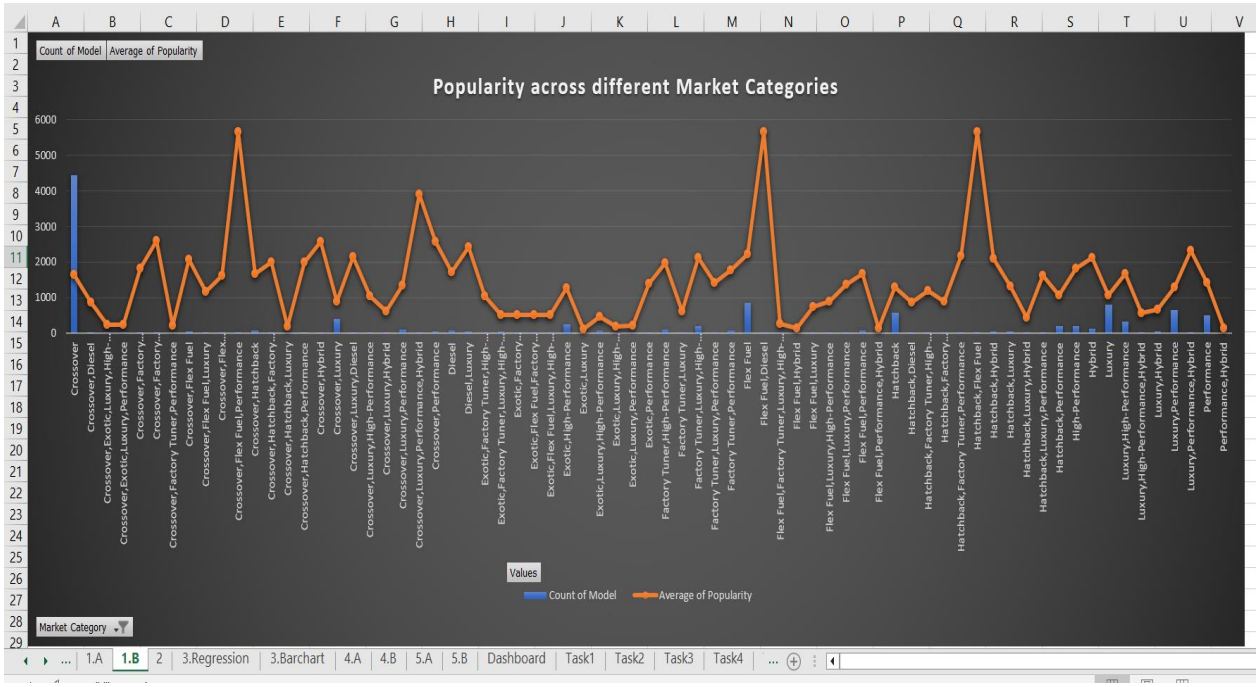
  I created a pivot table from the Model, Market Category, Popularity columns data and taken the Model into values section, Market category into rows section and took the average of popularity into values section which shows the number of car models in each market category and their popularity scores.

Market Category	Count of Model	Average of Popularity
Crossover	4451	1638.588407
Crossover,Diesel	7	873
Crossover,Exotic,Luxury,High-Performance	1	238
Crossover,Exotic,Luxury,Performance	1	238
Crossover,Factory Tuner,Luxury,High-Performance	26	1823.461538
Crossover,Factory Tuner,Luxury,Performance	5	2607.4
Crossover,Factory Tuner,Performance	4	210
Crossover,Flex Fuel	64	2073.75
Crossover,Flex Fuel,Luxury	10	1173.2
Crossover,Flex Fuel,Luxury,Performance	6	1624
Crossover,Flex Fuel,Performance	6	5657
Crossover,Hatchback	72	1675.694444
Crossover,Hatchback,Factory Tuner,Performance	6	2009
Crossover,Hatchback,Luxury	7	204
Crossover,Hatchback,Performance	6	2009
Crossover,Hybrid	42	2563.380952
Crossover,Luxury	406	889.2142857
Crossover,Luxury,Diesel	34	2149.411765
Crossover,Luxury,High-Performance	9	1037.222222
Crossover,Luxury,Hybrid	24	630.9166667
Crossover,Luxury,Performance	112	1349.089286
Crossover,Luxury,Performance,Hybrid	2	3916
Crossover,Performance	69	2585.956522
Diesel	84	1730.904762
Diesel,Luxury	47	2416.106383
Exotic,Factory Tuner,High-Performance	21	1046.380952
Exotic,Factory Tuner,Luxury,High-Performance	51	523.0196078
Exotic,Factory Tuner,Luxury,Performance	3	520

5.A | 5.B | Dashboard | Task1 | Task2 | Task3 | Task4 | ...

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

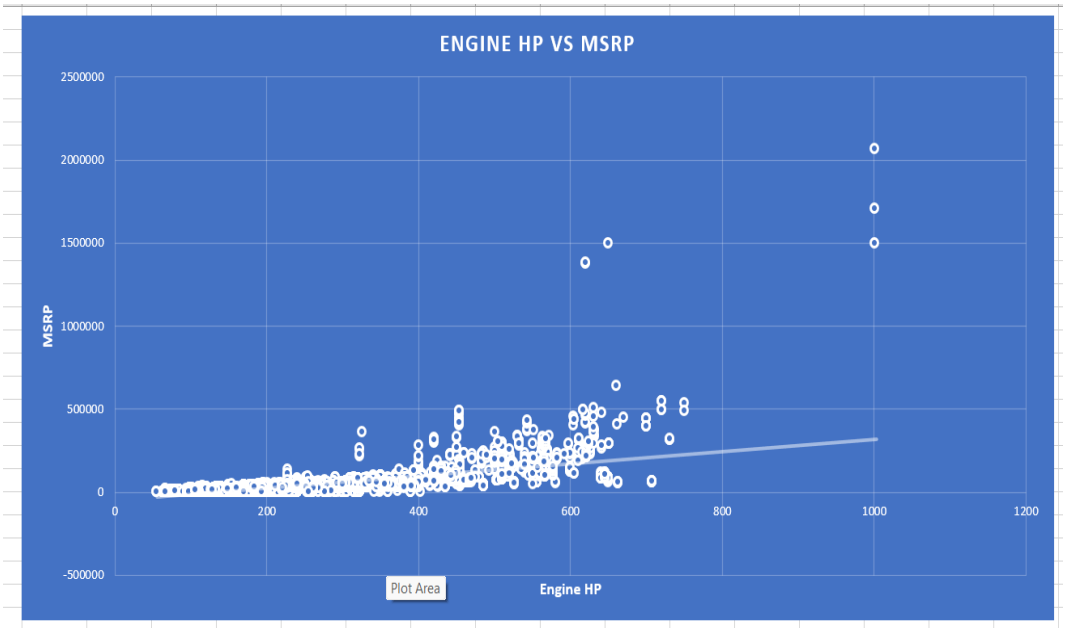
  From the Pivot table created in 1.A, I inserted a combo chart showing the market category on X axis and Count of model on Y axis with a line chart in combination which resembles the popularity.



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.

  I inserted a XY Scatter chart based on the Engine HP and MSRP columns data and added a trendline to show the direction of relation.

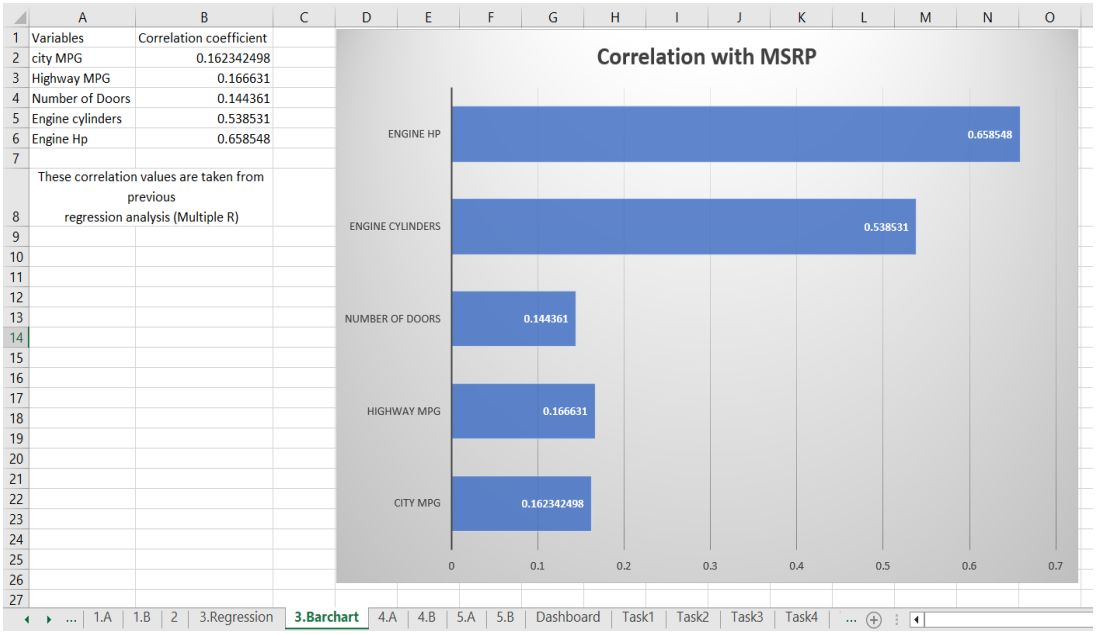


Insight Required: Which car features are most important in determining a car's price?

- **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.

  I have taken multiple columns for the correlation with MSRP(Price) using the Regression analysis from the Toolpak. Then, I got the Summary output of the Regression analysis. From the Summary, Looking at the values of Multiple R we can see the correlation of that data column with the MSRP(Price).

  For all those values, I inserted a bar chart which shows the relationship between Car's Price and different variables data.



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.

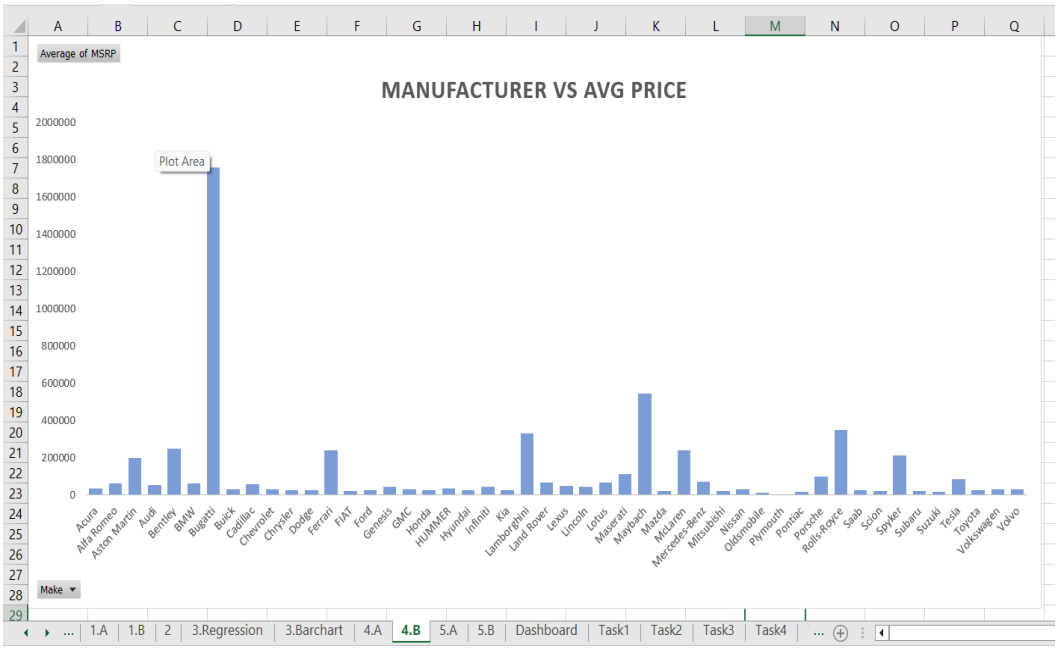
  I have created a pivot table which shows the Manufacturer and average of MSRP from their corresponding columns data.

Manufacturer	Average of MSRP
Acura	35087.4878
Alfa Romeo	61600
Aston Martin	198123.4615
Audi	54574.1215
Bentley	247169.3243
BMW	62162.55864
Bugatti	1757223.667
Buick	29034.18947
Cadillac	56368.26515
Chevrolet	29074.72576
Chrysler	26722.96257
Dodge	24857.04537
Ferrari	238218.8406
FIAT	22670.24194
Ford	28511.30788
Genesis	46616.66667
GMC	32444.08506
Honda	26655.14781
HUMMER	36464.41176
Hyundai	24926.26255
Infiniti	42640.27134
Kia	25513.75546
Lamborghini	331567.3077
Land Rover	68067.08633
Lexus	47549.06931
Lincoln	43860.825
Lotus	68377.14286
Maserati	113684.4909

4.B | 5.A | 5.B | Dashboard | T

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

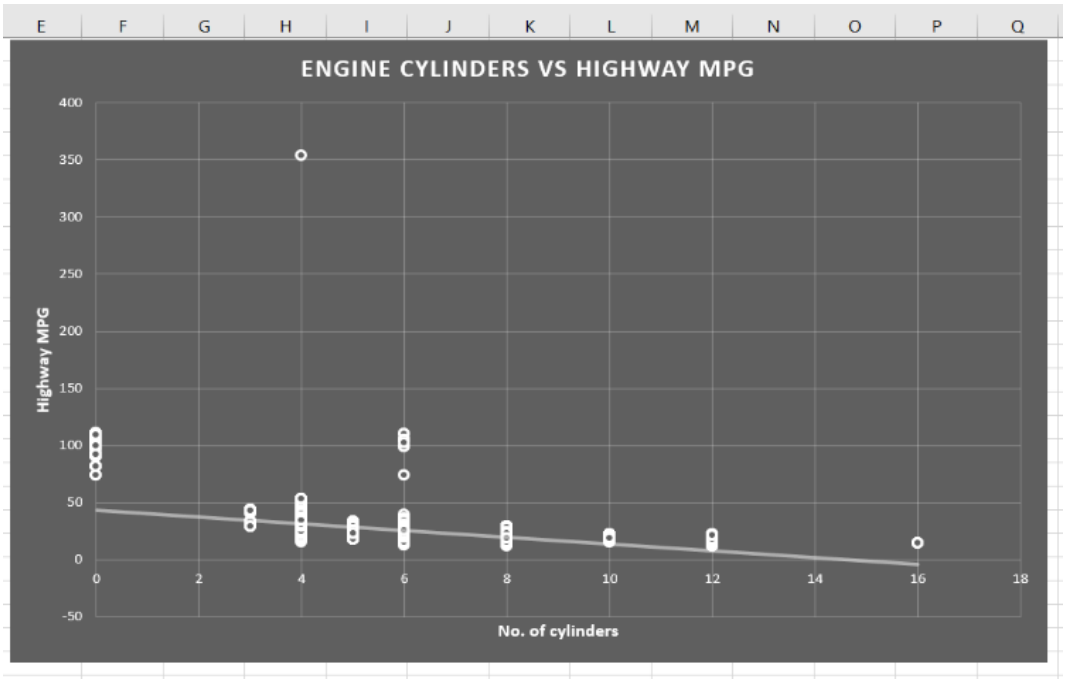
  I inserted a bar chart based on the values provided in the pivot table with variables Manufacturer and Average of car price (MSRP).



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

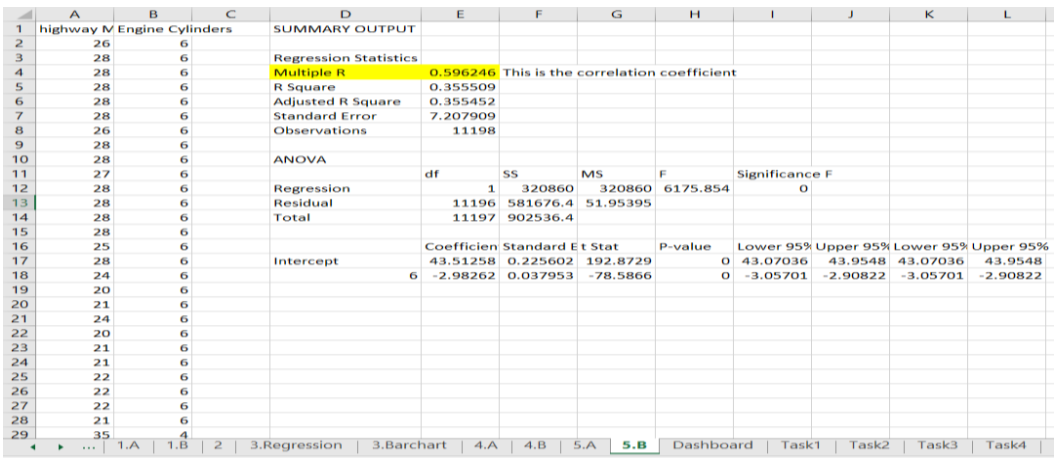
- **Task 5.A:** Create a scatter plot with the number of cylinders on the x-axis and highway MPG on the y-axis. Then create a trendline on the scatter plot to visually estimate the slope of the relationship and assess its significance.

  I created an XY Scatter Plot with the Number of cylinders on x axis and Highway MPG on y axis. Then added a trendline to the direction of their relationship.



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

  From the Highway MPG and Engine Cylinders columns, I did a regression analysis from the Toolpak and got the corresponding values which shows the correlation between them.



	A	B	C	D	E	F	G	H	I	J	K	L
1	highway M	Engine Cylinders		SUMMARY OUTPUT								
2	26	6										
3	28	6		Regression Statistics								
4	28	6		Multiple R	0.596246	This is the correlation coefficient						
5	28	6		R Square	0.355509							
6	28	6		Adjusted R Square	0.355452							
7	28	6		Standard Error	7.207909							
8	26	6		Observations	11198							
9	28	6										
10	28	6		ANOVA								
11	27	6			df	SS	MS	F	Significance F			
12	28	6		Regression	1	320860	320860	6175.854	0			
13	28	6		Residual	11196	581676.4	51.95395					
14	28	6		Total	11197	902536.4						
15	28	6										
16	25	6			Coefficien	Standard E	t Stat	P-value	Lower 95%	Upper 95%	Lower 95%	Upper 95%
17	28	6		Intercept	43.51258	0.225602	192.8729	0	43.07036	43.9548	43.07036	43.9548
18	24	6		6	-2.98262	0.037953	-78.5866	0	-3.05701	-2.90822	-3.05701	-2.90822
19	20	6										
20	21	6										
21	24	6										
22	20	6										
23	21	6										
24	21	6										
25	22	6										
26	22	6										
27	22	6										
28	21	6										
29	35	4										

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.

  From the columns MSRP, Make and Vehicle Style, I created a Pivot table with Make into rows section, Vehicle style into Columns section, MSRP into Values section with value field settings as SUM inorder to calculate the Total MSRP. Based on the pivot table, I inserted a Stacked column chart and used a slicer to make it interactive.

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.

  From the columns Make, Vehicle Style and MSRP, I created pivot table with Make into rows section, vehicle style into columns section, MSRP into values section with value field settings as Average to calculate the Average MSRP for each brand and body style. Based on the pivot , I inserted a clustered column chart which shows the highest and lowest average MSRP's across different body style.

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.

  From the columns Transmission type, MSRP & Vehicle Style, I created a pivot table with Transmission type into rows section, Vehicle style into columns section and MSRP into values section with value field settings as Average to calculate Average of MSRP across each body style. Based on the pivot table, I inserted the XY Scatter Plot which shows how different features affect the MSRP across different body style.

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.

From the columns Vehicle style, Highway MPG and Model year, I created a pivot table with Vehicle style into rows section and model years into columns section and Highway MPG into values section with value field settings as average to calculate the average MPG across different body styles and model years.Based on the pivot, Inserted a line chart to show the trend of fuel efficiency over time for each body style.

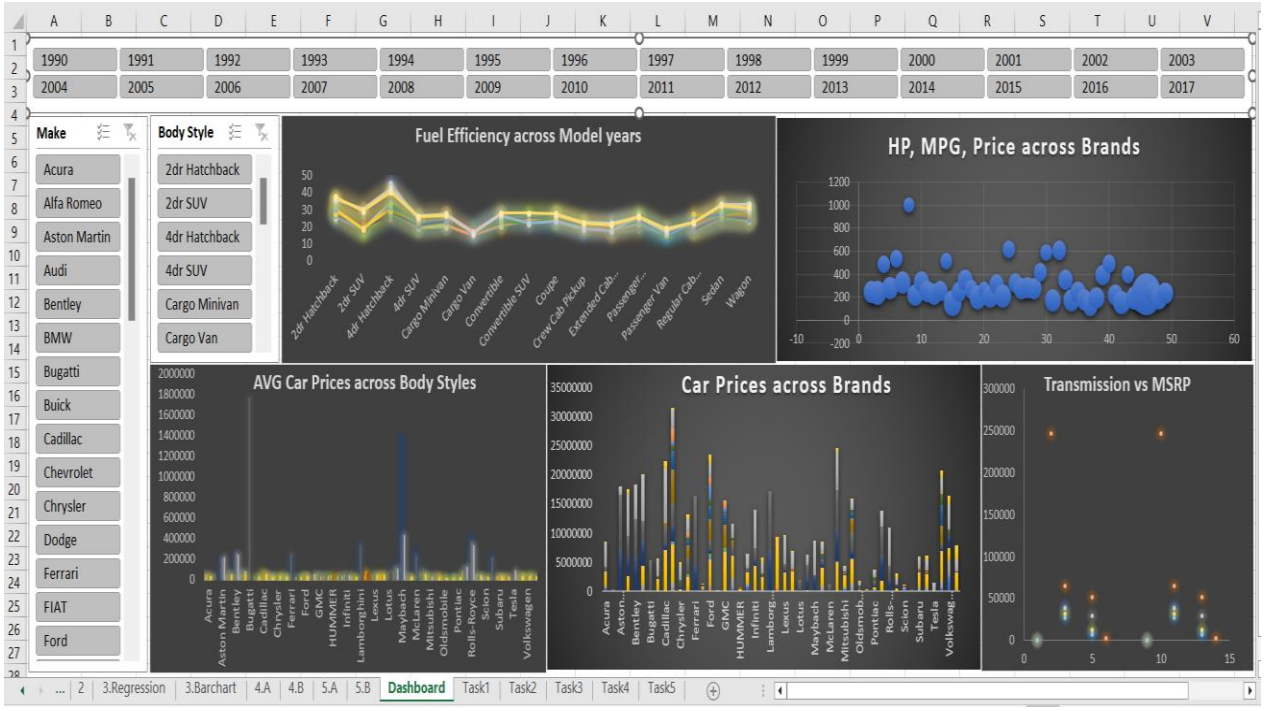
Task 5: How does the car's horsepower, MPG, and price vary across different Brands?

- **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.

From the columns Make, Engine HP, Highway MPG and MSRP , I created a pivot table with Make into rows section, Engine HP, Highway MPG, MSRP into values section with value field settings as average to calculate the average horsepower, MPG and MSRP for each car brand. From the pivot I copied and pasted all the values beside to create a Bubble chart to visualize the relationship between Engine HP, MPG, and Price.

DASHBOARD:

I moved the charts from all the previous tasks and created a dashboard with slicers to make it look interactive.



MY EXCEL FILE:

https://docs.google.com/spreadsheets/d/1pXU3QCNlOV-H2CZTD3I6Ii_7vMNUWW1t/edit?usp=share_link&ouid=104755012826368900391&rtpof=true&sd=true

Note: Download the file from the link and view it in Ms Excel , as Google sheets preview is not showing or responding to some options created in the excel