MODULE 7

AUGUST'2024

IMPACT OF CAR FEATURES

ON PRICE AND PROFITABILITY

REPORT

Gunisha Chopra
DATA ANALYST

MODULE 7

IMPACT OF CAR FEATURES ON PRICE AND PROFITABILITY

FINAL PROJECT-3

PROJECT DESCRIPTION:

The project aims to analyse a dataset containing information on over 11,000 car models and their specifications. The dataset includes details such as the car's make, model, year, fuel type, engine power, transmission, wheels, number of doors, market category, size, style, estimated miles per gallon, popularity, and manufacturer's suggested retail price (MSRP).

BUSINESS PROBLEM:

The automotive industry is constantly evolving, with new car models being introduced regularly and consumer preferences shifting. To stay competitive, manufacturers and dealers need to understand market trends, consumer preferences, and pricing dynamics. However, analysing large datasets containing information on thousands of car models and their specifications can be challenging and time-consuming.

The business problem addressed in this project is to provide actionable insights to stakeholders in the automotive industry based on the analysis of a dataset containing information on over 11,000 car models.

APPROACH:

1.) DATA CLEANING AND PROCESSING: Before starting the analysis, we cleaned and processed the data to ensure its accuracy and reliability. This involved tasks like handling missing values, removing duplicates, standardizing data formats, and addressing any inconsistencies or errors in the dataset.

2.) EXPLORATORY DATA ANALYSIS (EDA)**:** Perform EDA to understand the distribution and relationships between variables in the dataset. This may include creating visualizations such as scatter plots.

3.) STATISTICAL ANALYSIS**:** Conduct statistical analysis to identify trends and patterns in the data.

4.) INSIGHTS GENERATION**:** Generate insights from the analysis, including trends in-car features and prices, fuel efficiency comparisons, and popularity analysis. These insights will help you understand consumer preferences and market dynamics.

5.) REPORTING AND PRESENTATION: Prepare a detailed report documenting the findings of the analysis, including visualisations and key insights. Create a presentation summarising the key findings and insights to communicate them effectively to stakeholders.

6.) ITERATE AND IMPROVE**:** Iterate on the analysis based on feedback and new data, and continually improve the model to make more accurate predictions and generate more valuable insights.

TECH-STACK USED:

MICROSOFT EXCEL

- **Excel Pivot Tables:** Used for summarising and analysing large datasets, allowing for easy filtering, sorting, and grouping of data.

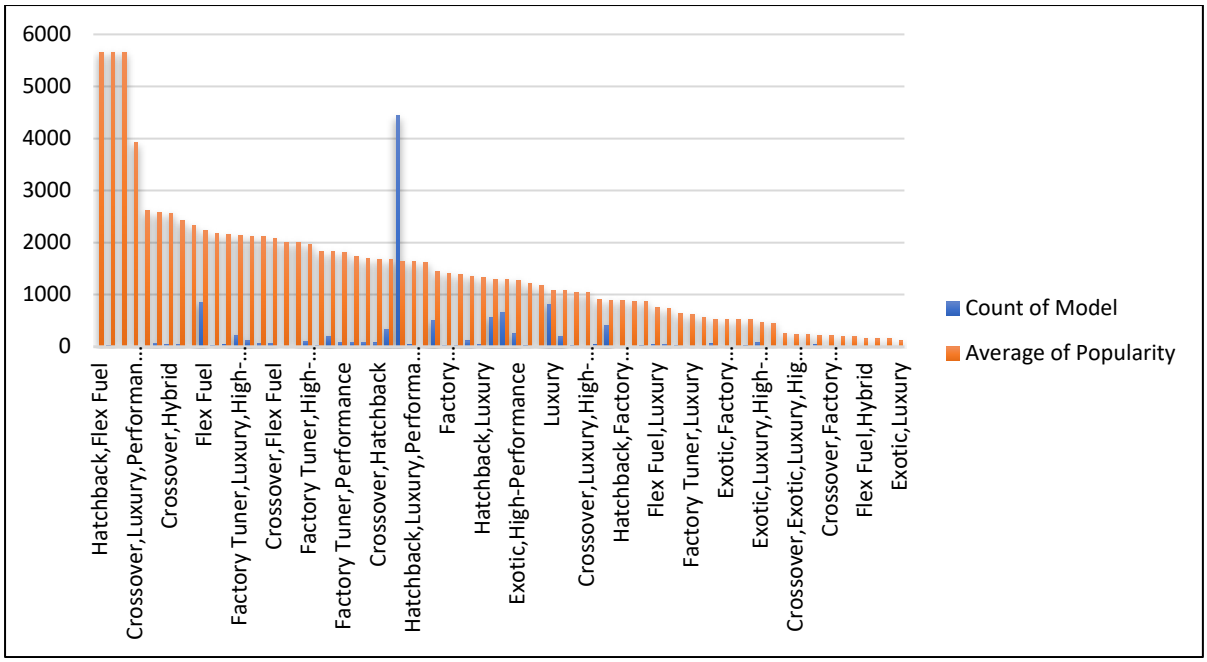
- **Excel Scatter Plots:** Used for visualizing the relationship between two continuous variables, such as engine power and price.

- **Excel Bar Charts:** Used for comparing categorical data, such as market categories or fuel types.

- **Excel Bubble Charts:** Used for visualizing two-dimensional data, where the size of the bubble represents a third variable, such as popularity.

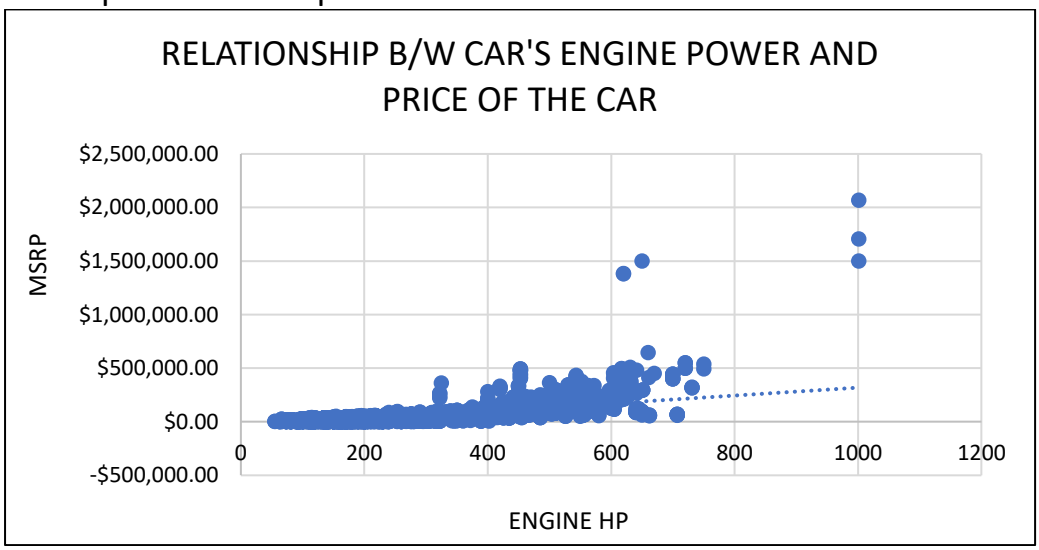
TASKS AND INSIGHTS:

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



- HIGHEST MODEL COUNT: VALUE-> 4435
  Crossover
- HIGHEST AVERAGE POPULARITY: VALUES-> 5657,
  Hatchback, Flex Fuel

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



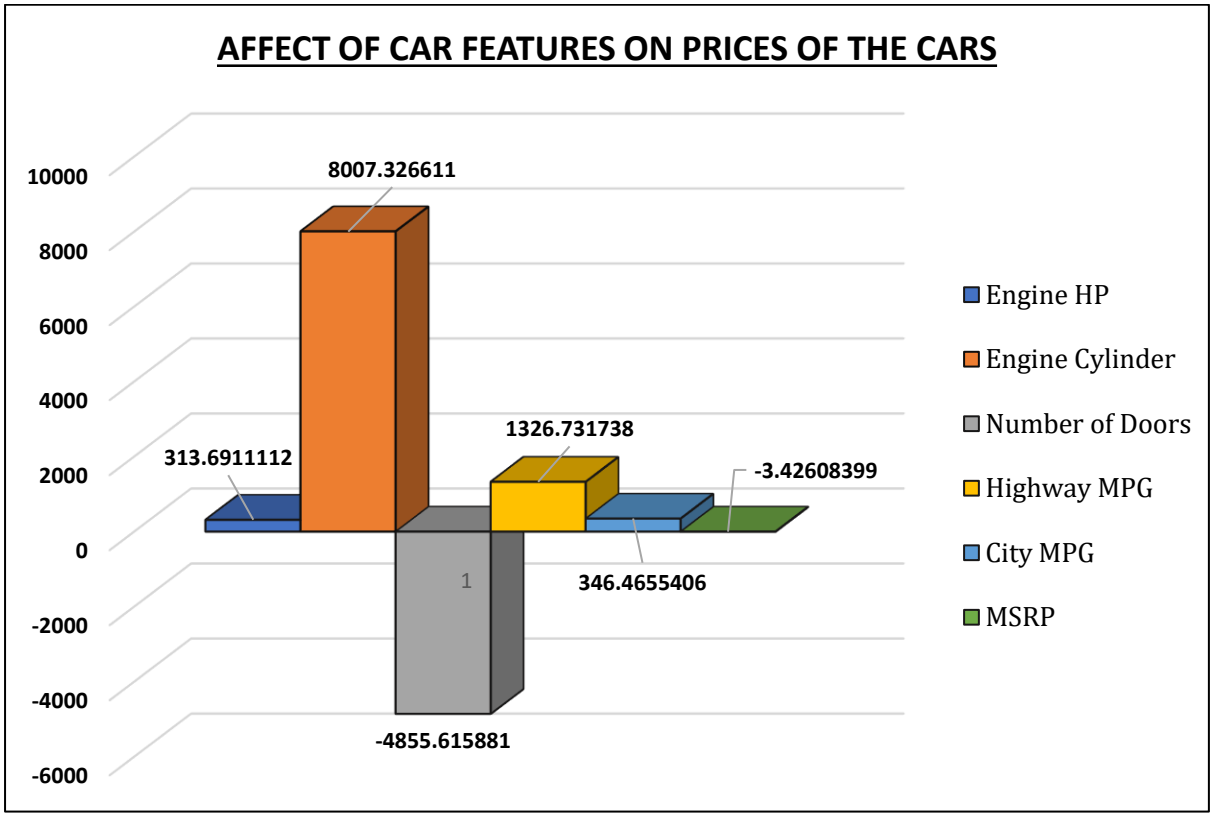
We see a positive relationship between ENGINE HP and MSRP. As engine HP increases, so does MSRP.

3.) INSIGHT REQUIRED: Which car features are most important in determining a car's price?

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.683632855
R Square	0.467353881
Adjusted R Square	0.467066353
Standard Error	45050.16436
Observations	11122

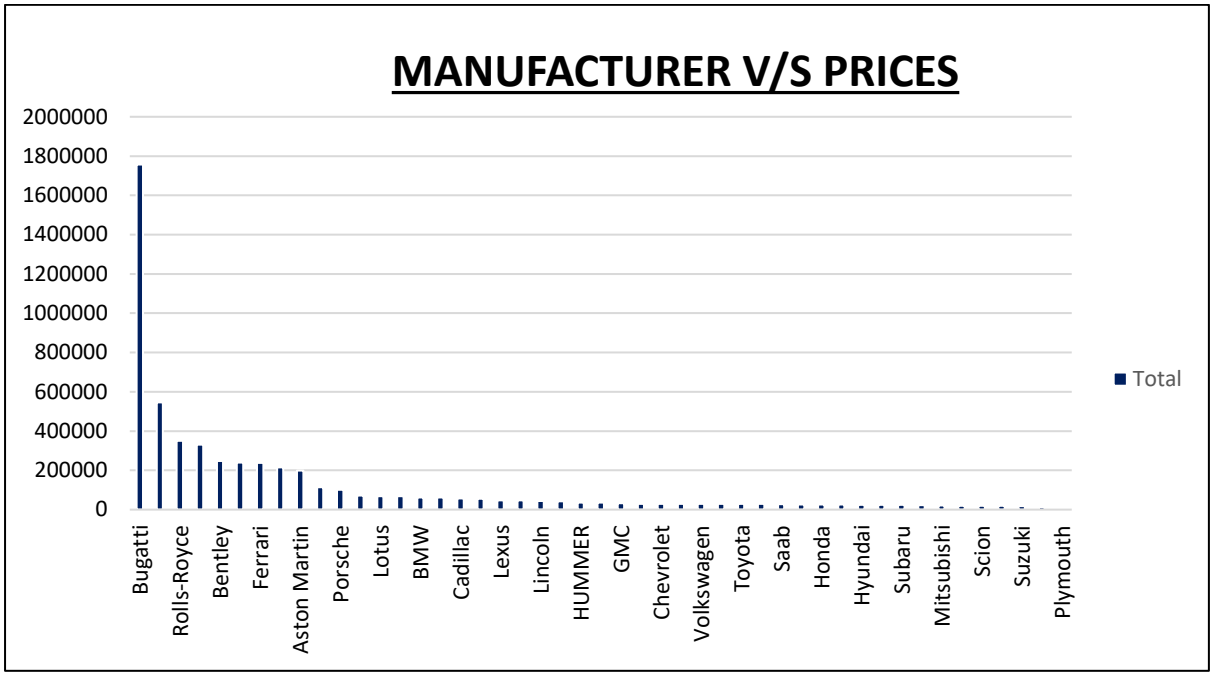
ANOVA	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	6	1.97929E+13	3.3E+12	1625.419	0
Residual	11115	2.25581E+13	2.03E+09		
Total	11121	4.2351E+13			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-102632.3296	4247.791679	-24.1613	9.6E-126	-110958.755	-94305.9042	-110958.755	-94305.90417
Engine HP	313.6911112	6.417472726	48.88079	0	301.111726	326.270496	301.111726	326.2704965
Engine Cylinder	8007.326611	474.9070156	16.86083	5.28E-63	7076.424594	8938.22863	7076.424594	8938.228627
Number of Doors	-4855.615881	495.9000533	-9.79152	1.51E-22	-5827.667976	-3883.56379	-5827.667976	-3883.563785
Highway MPG	1326.731738	172.5238255	7.690136	1.59E-14	988.5544276	1664.90905	988.5544276	1664.909048
City MPG	346.4655406	158.9342156	2.17993	0.029284	34.92627724	658.004804	34.92627724	658.0048039
MSRP	-3.42608399	0.296678217	-11.5481	1.13E-30	-4.007625936	-2.84454204	-4.007625936	-2.844542043



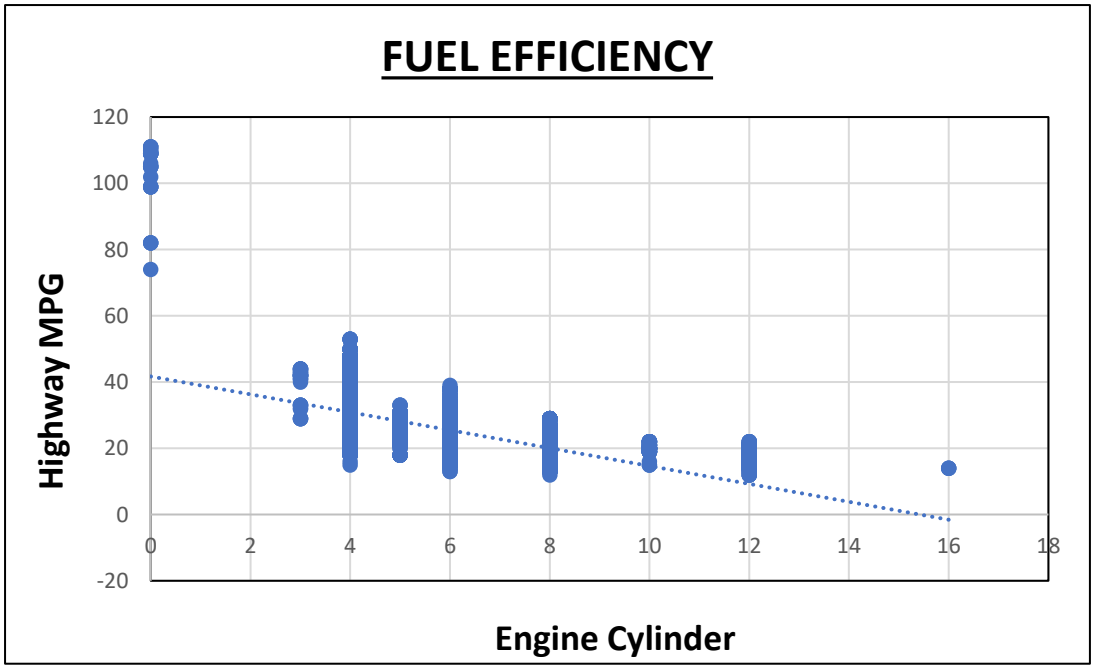
The most influential variables seem to be the **ENGINE CYLINDER variable (the number of cylinders in a car's engine)** and the least important variable for the MSRP is the **NUMBER OF DOORS** variable.

4.) INSIGHT REQUIRED: How does the average price of a car vary across different manufacturers?



We see that BUGATTI with an MSRP value of 1757223.667 is by far the highest value in the MSRP variable. On the contrary, PLYMOUTH happens to have the lowest value of 12843.79.

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



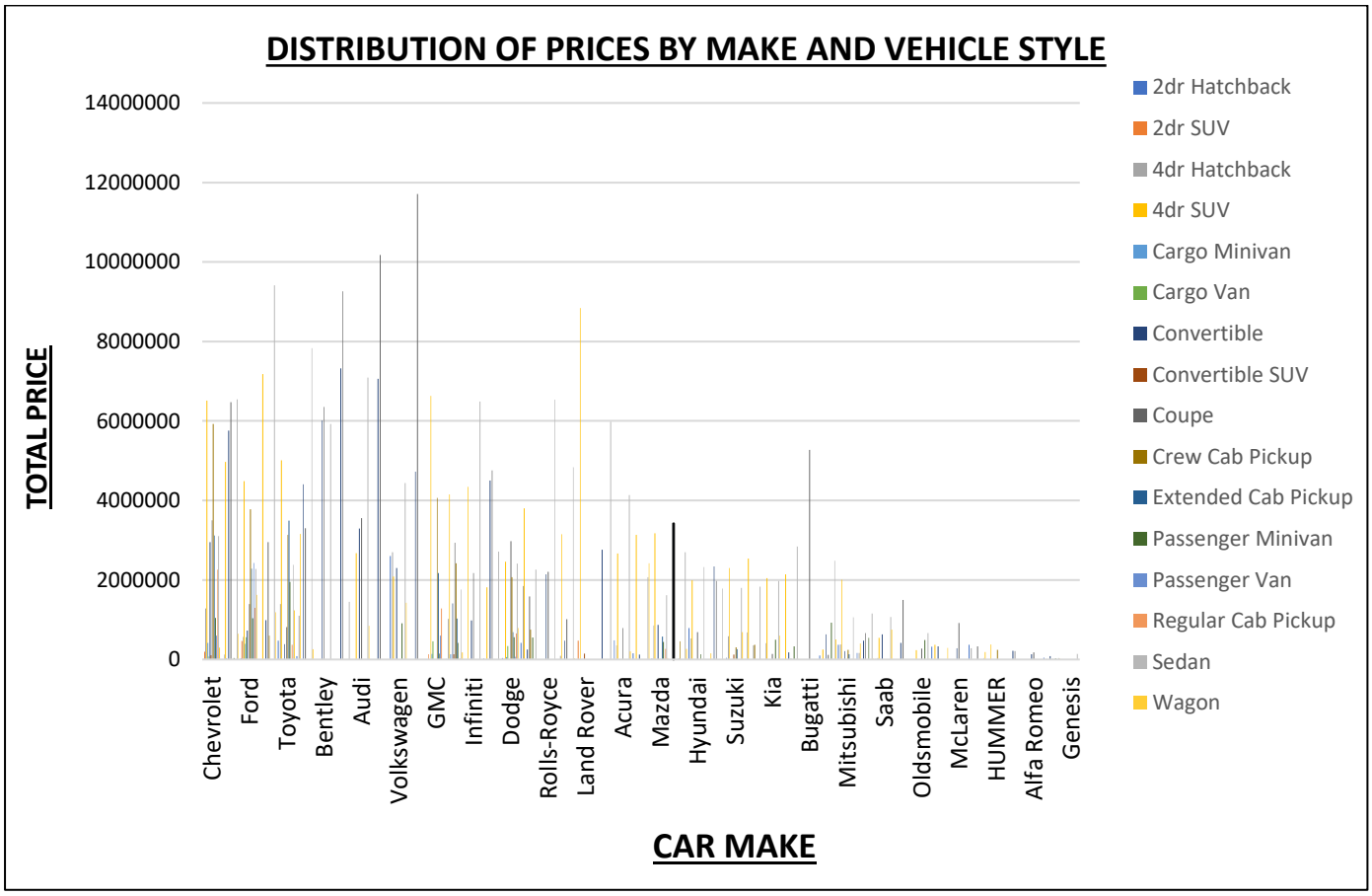
We see a negative relationship between fuel efficiency and number of cylinders in a car's engine.

CORRELATION COEFFICIENT	-0.663587792

Therefore, the correlation coefficient between the two variables is **-0.663587792**

DASHBOARD TASKS:

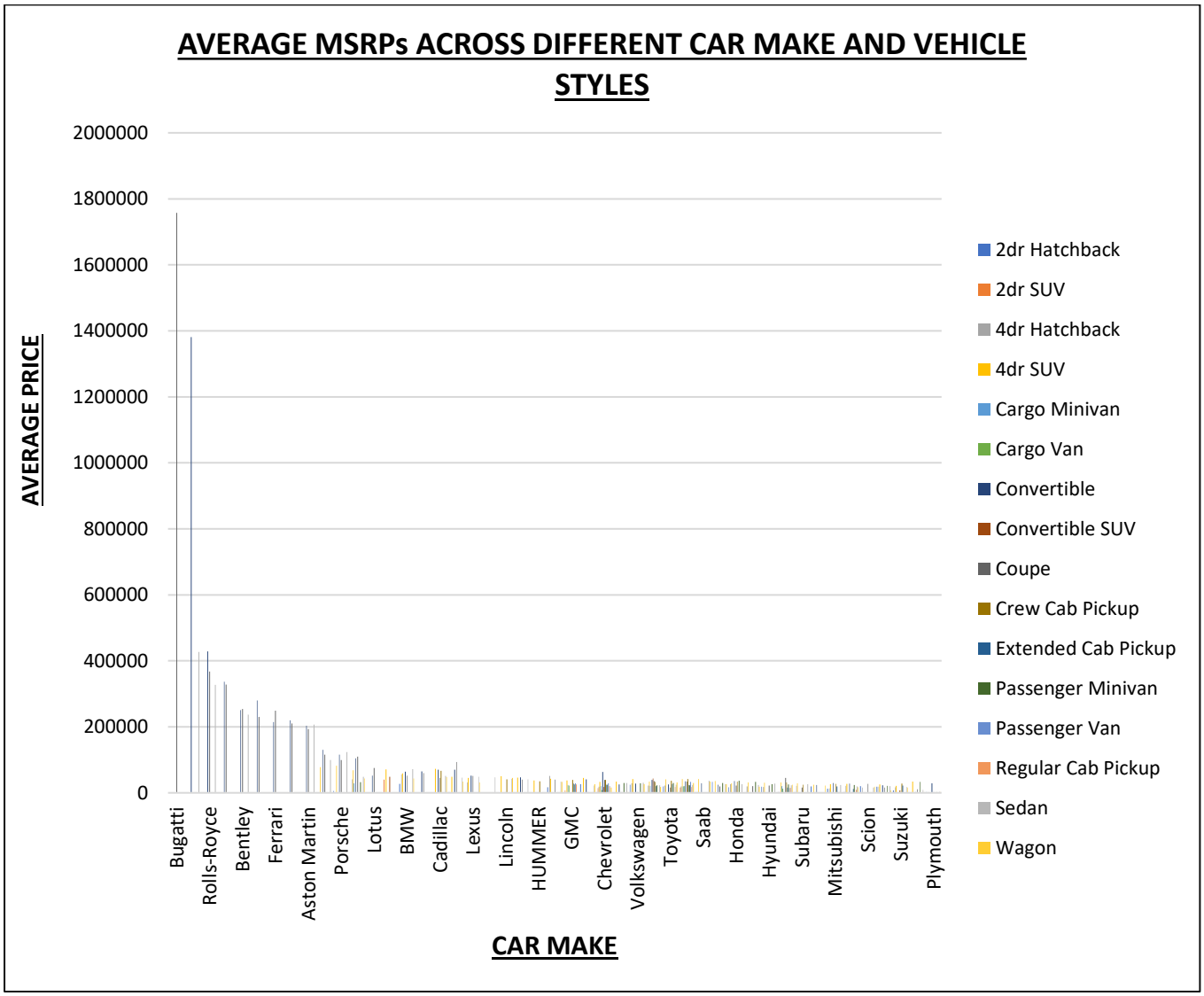
1.) TASK-1: How does the distribution of car prices vary by brand and body style?



HIGHEST POINT:

- Body style "Coupe"
- Car brand "Ferrari"
- MSRP 11,713,289

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

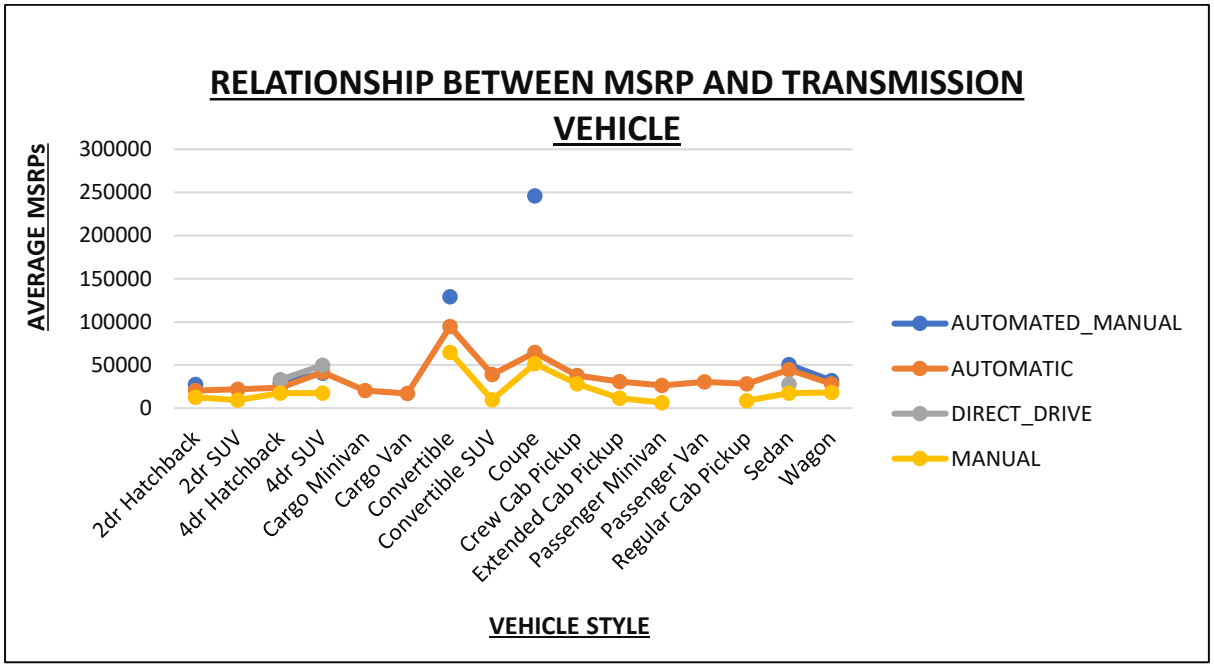


CAR BRAND WITH:

HIGHEST AVERAGE MSRP: BUGATTI (1,757,223.667), STYLE: “COUPE”

LOWEST AVERAGE MSRP: PLYMOUTH (28,543.66667), STYLE: “CONVERTIBLE”

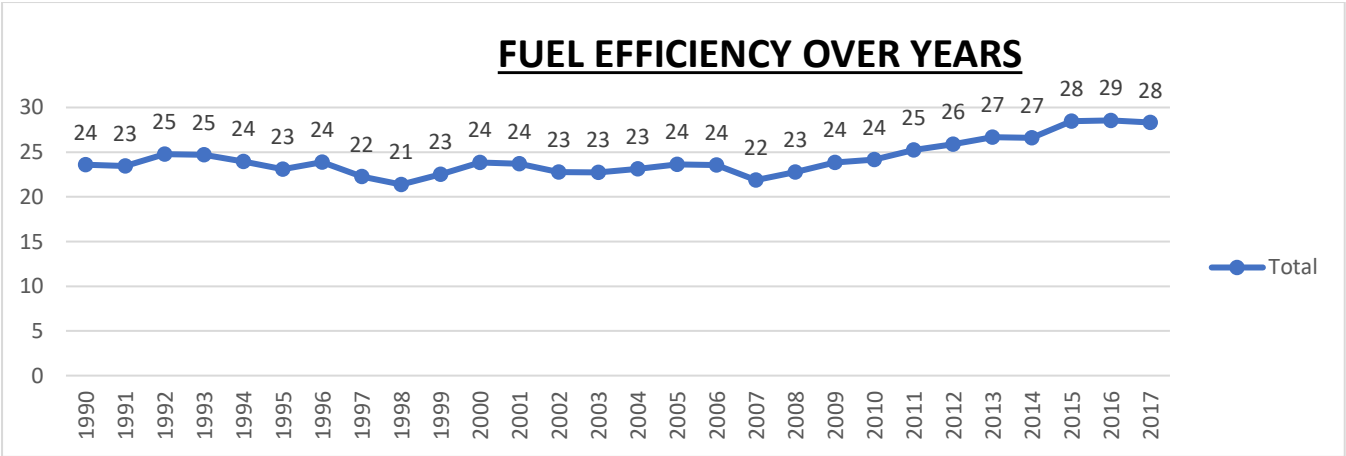
3.) TASK 3: How do the different features such as transmission type affect the MSRP, and how does this vary by body style?



We see the highest point in our chart belongs to

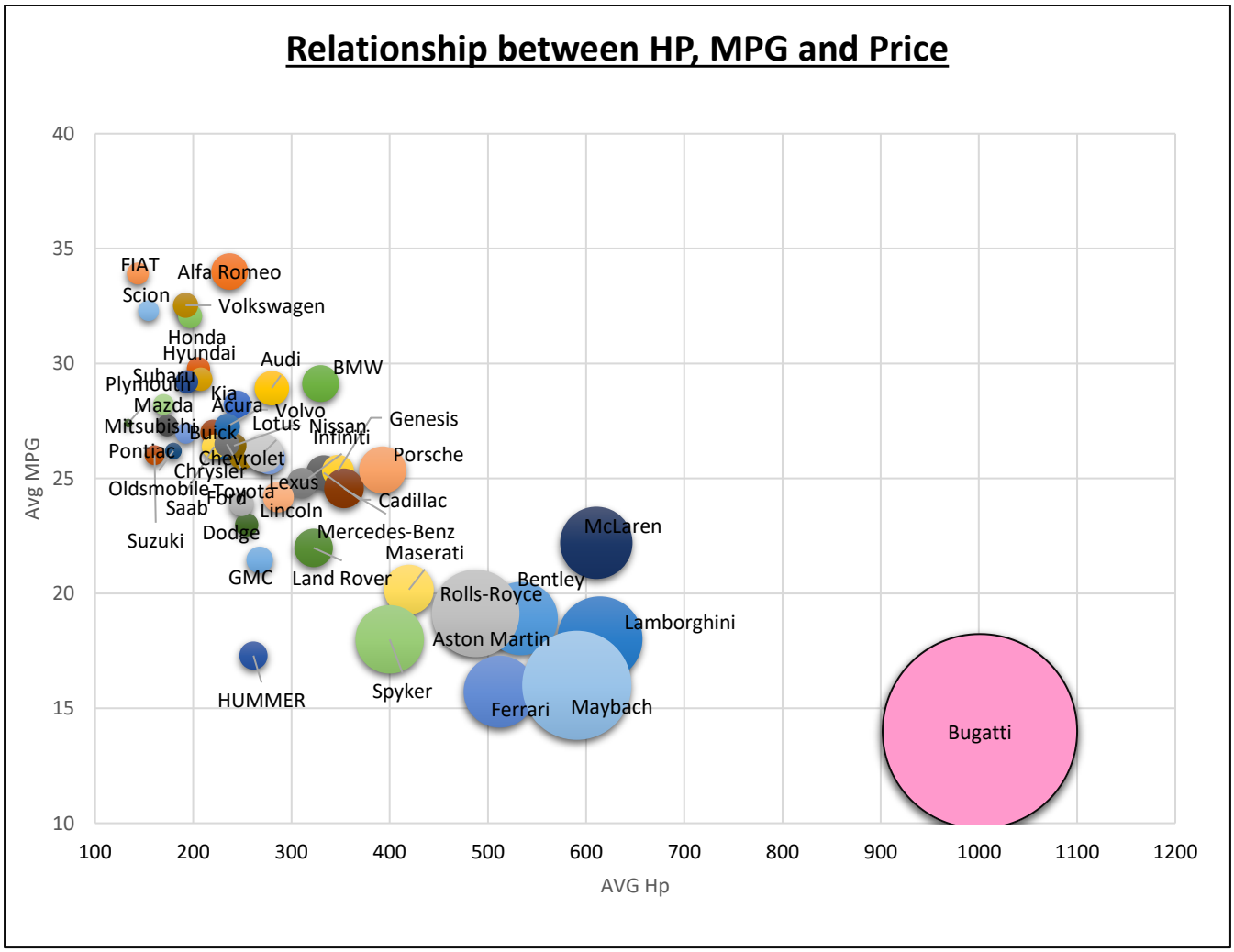
- transmission type: "AUTOMATED_MANUAL"
- vehicle style: "COUPE"
- MSRP: 245,977.4252

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



WE SEE THE HIGHEST POINT OF AVERAGE HIGHWAY MPG IS "29" IN THE YEAR 2016.

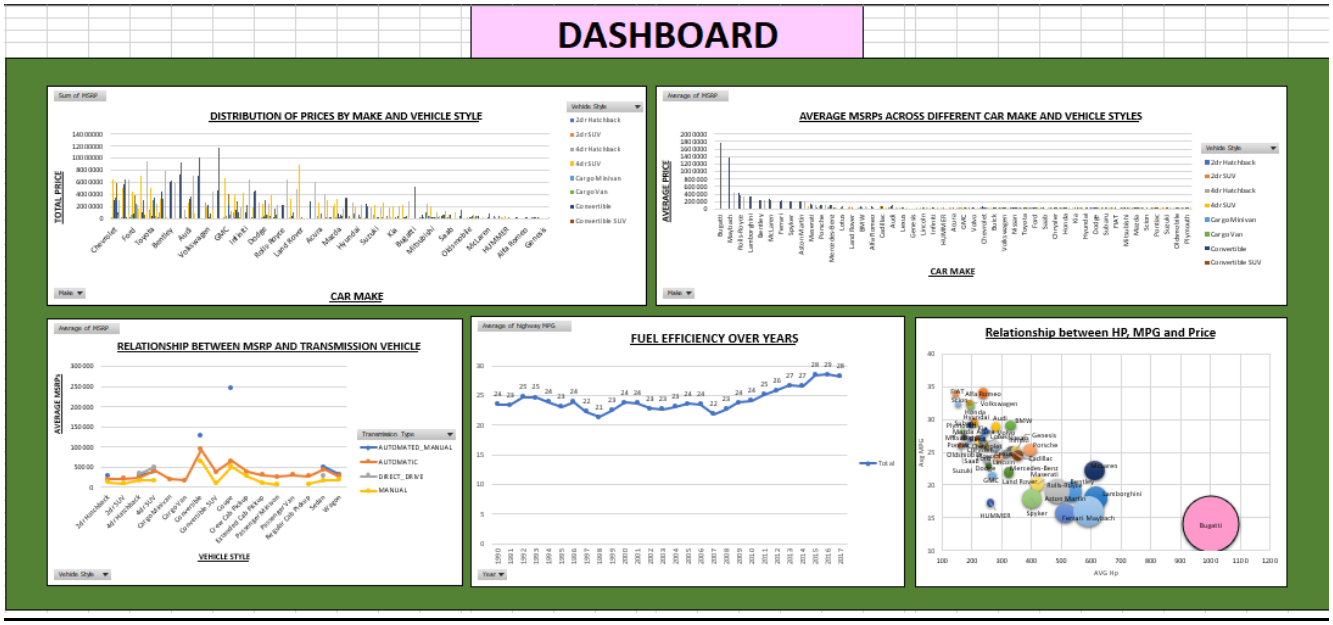
5.) TASK 5: How does the car's horsepower, MPG, and price vary across different Brands?

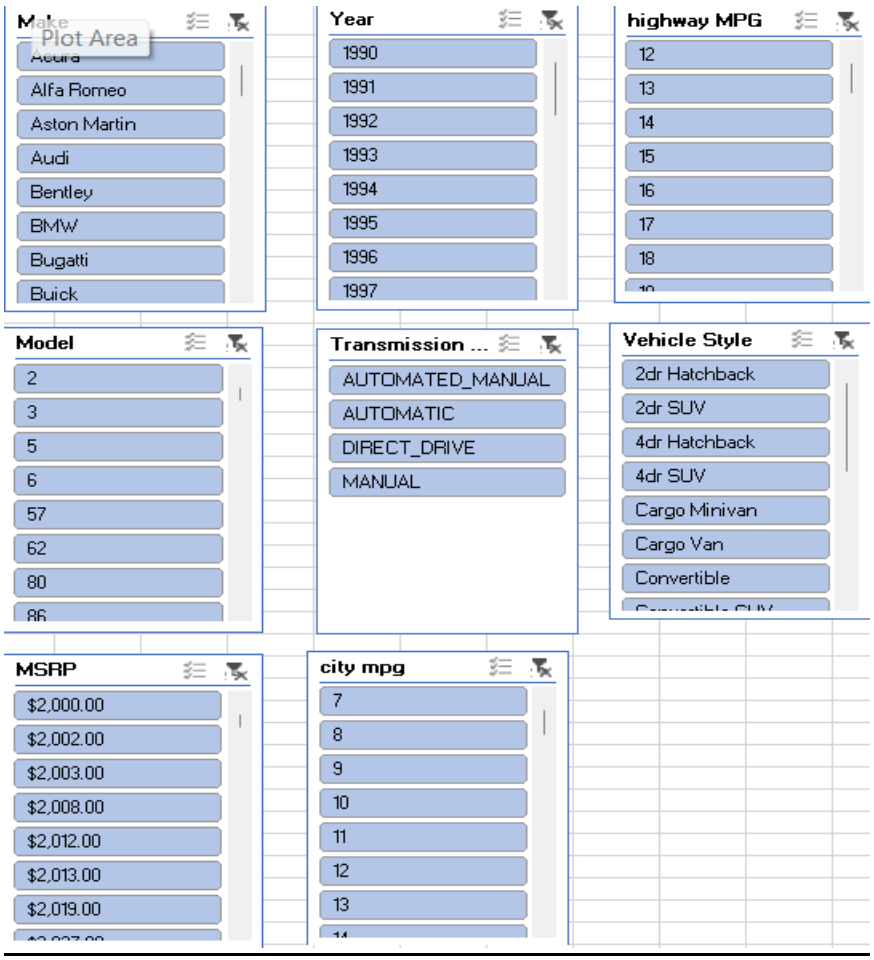


HIGHEST POINT:

- **CAR BRAND:** BUGATTI
- **AVERAGE HORSEPOWER**: 1001
- **AVERAGE MPG**: 14
- **MSRP**: 1757223.667

CREATING INTERACTIVE DASHBOARD





CONCLUSION:

The analysis of the dataset containing information on over 11,000 car models has provided valuable insights into various aspects of the automotive industry. Key findings include:

1. Trends in Car Features and Pricing: The analysis revealed trends showing how car features and prices have changed over time. This information can help manufacturers make informed decisions about product development and pricing strategies.

2. Fuel Efficiency Comparison: The comparison of fuel efficiency among different types of cars highlighted which types are the most efficient. This can assist consumers in making informed decisions about their purchases.

3. Popularity Analysis: The analysis of the relationship between a car's features and its popularity identified which features are most popular among consumers. This can help manufacturers tailor their marketing strategies to better appeal to consumers.

4. Price Prediction Model: The developed model to predict the price of a car based on its features and market category can help manufacturers set competitive prices and consumers make informed purchasing decisions.

Overall, these insights provide stakeholders in the automotive industry with valuable information that can inform their decision-making processes and help them stay competitive in a rapidly changing market. The data cleaning and processing steps ensured the accuracy and reliability of the analysis, laying a solid foundation for deriving meaningful conclusions from the dataset.

-END-

HYPERLINK TO EXCEL FILE:
https://docs.google.com/spreadsheets/d/1hHelD6DAw5w8uziekGKE37ITkyMjEz5h/edit?usp=sharing&ouid=102683227032029211056&rtpof=true&sd=true

HYPERLINK TO PPT:
https://drive.google.com/file/d/1NRdbjW_eajVV9D0I_B4xG5pMl8FUUbN8/view?usp=sharing

HYPERLINK TO VIDEO SUBMISSION:
https://drive.google.com/file/d/1G7fpcKV2ELBBfr_720NFJ7rHtqY85sts/view?usp=sharing