PROJECT-7

Analyzing the Impact of Car Features on Price and Profitability

Project Description:

Overview:

The goal of this research is to investigate how various car attributes impact the cost and profitability of automobiles within the automotive sector. through the examination of a dataset that includes details on vehicle models, their specs, and market statistics. The intention is to provide information that will help automakers make more informed decisions about pricing and new product development.

Business problem:

Manufacturers must comprehend what drives consumer demand in the rapidly evolving automobile market of today to properly set pricing and optimize profits. The difficulty of striking a balance between satisfying customer wants, adding appealing features, and determining profitable prices is tackled by this project.
As a data analyst, I used this dataset to gain insights into various aspects of the automotive industry, 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:

The dataset contains information on various car models and their specifications, and is titled "Car Features and MSRP".

- **Number of observations:** 11,159
- **Number of variables:** 16

Information on car models and their specifications, including details on 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). the dataset was last updated in 2017.

Data cleaning and preprocessing steps:

I started by cleaning the data to ensure its accuracy and reliability before conducting the analysis. This involved managing missing data, removing duplicates, standardizing data formats, and addressing any discrepancies in the dataset.
During this process, I found 102 empty rows and 715 duplicate values, which I then removed from the dataset.

Approach:

Analytical method:

- Descriptive statistics were used to summarize the dataset's salient aspects. These numbers provide information on the distribution and average values of several characteristics, such as MPG, MSRP, and Engine output.

- Visualization techniques, such as scatter plots, line graphs, and clustered column charts, were used to visually analyze the dataset's relationships and patterns. This facilitated the identification of trends, anomalies, and potential connections among disparate data sets.

- The application of analytical techniques was directed by the project's goal, which included developing effective models for price analysis and gaining insights into many aspects of the automotive industry.

Modeling techniques:

- Regression analysis, a modeling technique, was used to forecast car prices. more exact linear regression with feature selection and market category considerations. This project's objective was to establish relationships between independent variables (car pricing) and dependent variables (market category, engine power, and MPG) to fit a linear model to the data and conduct predictive analysis**.**

Challenges:

- One of the obstacles I encountered during the project was managing missing data. Other challenges were ensuring data quality and reliability, selecting an appropriate regression model, and optimizing its features for optimal performance in predicting car prices.

Tech-Stack Used:

Microsoft® Excel® 2021 MSO (Version 2404 Build 16.0.17531.20152) 64-bit

Excel's built-in functions, pivot tables, and charting capabilities made it easier to explore and interpret the dataset. Therefore, I used Excel as my primary tool for data manipulation, analysis, and visualization.
Microsoft Word was used to write a detailed report on this project. No other additional libraries or packages were used in the project.

Insights:

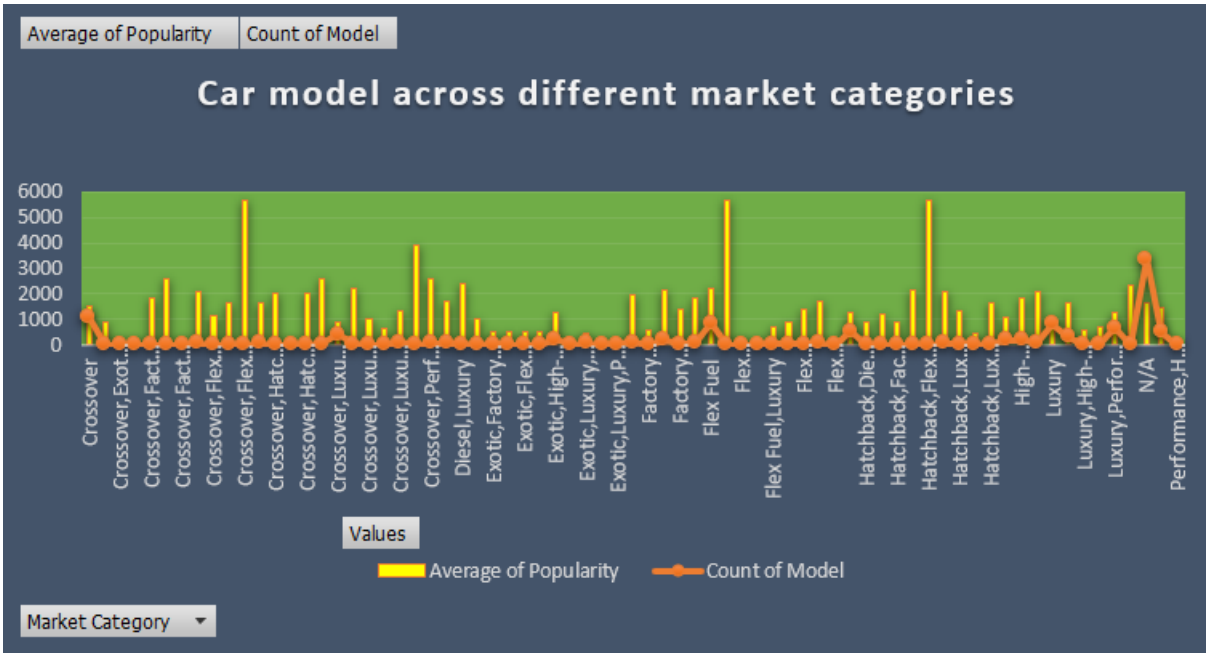
Tasks: Analysis

1. 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.
  Output:

Market category	Average of Popul	Count of Model
Crossover	1533.475655	1068
Crossover,Diesel	873	7
Crossover,Exotic,Luxury,High-Performance	238	1
Crossover,Exotic,Luxury,Performance	238	1
Crossover,Factory Tuner,Luxury,High-Performance	1823.461538	26
Crossover,Factory Tuner,Luxury,Performance	2607.4	5
Crossover,Factory Tuner,Performance	210	4
Crossover,Flex Fuel	2073.75	64
Crossover,Flex Fuel,Luxury	1173.2	10
Crossover,Flex Fuel,Luxury,Performance	1624	6
Crossover,Flex Fuel,Performance	5657	6
Crossover,Hatchback	1675.694444	72
Crossover,Hatchback,Factory Tuner,Performance	2009	6
Crossover,Hatchback,Luxury	204	7
Crossover,Hatchback,Performance	2009	6
Crossover,Hybrid	2563.380952	42
Crossover,Luxury	889.2142857	406
Crossover,Luxury,Diesel	2195.848485	33
Crossover,Luxury,High-Performance	1037.222222	9
Crossover,Luxury,Hybrid	630.3166667	24
Crossover,Luxury,Performance	1349.089286	112
Crossover,Luxury,Performance,Hybrid	3316	2
Crossover,Performance	2585.356522	69
Diesel	1730.304762	84
Diesel,Luxury	2416.106383	47
Exotic,Factory Tuner,High-Performance	1046.380952	21
Exotic,Factory Tuner,Luxury,High-Performance	523.0196078	51
Exotic,Factory Tuner,Luxury,Performance	520	3
Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance	520	13
Exotic,Flex Fuel,Luxury,High-Performance	520	11
Exotic,High-Performance	1270.326531	245
Exotic,Luxury	112.6666667	12
Exotic,Luxury,High-Performance	473.025974	77
Exotic,Luxury,High-Performance,Hybrid	204	1
Exotic,Luxury,Performance	217.0277778	36
Factory Tuner,High-Performance	1966.442308	104
Factory Tuner,Luxury	617	2
Factory Tuner,Luxury,High-Performance	2133.367442	215
Factory Tuner,Luxury,Performance	1413.419355	31
Factory Tuner,Performance	1818.049383	81
Flex Fuel	2225.71345	855
Flex Fuel,Diesel	5657	16
Flex Fuel,Factory Tuner,Luxury,High-Performance	258	1
Flex Fuel,Hybrid	155	2
Flex Fuel,Luxury	746.5384615	39

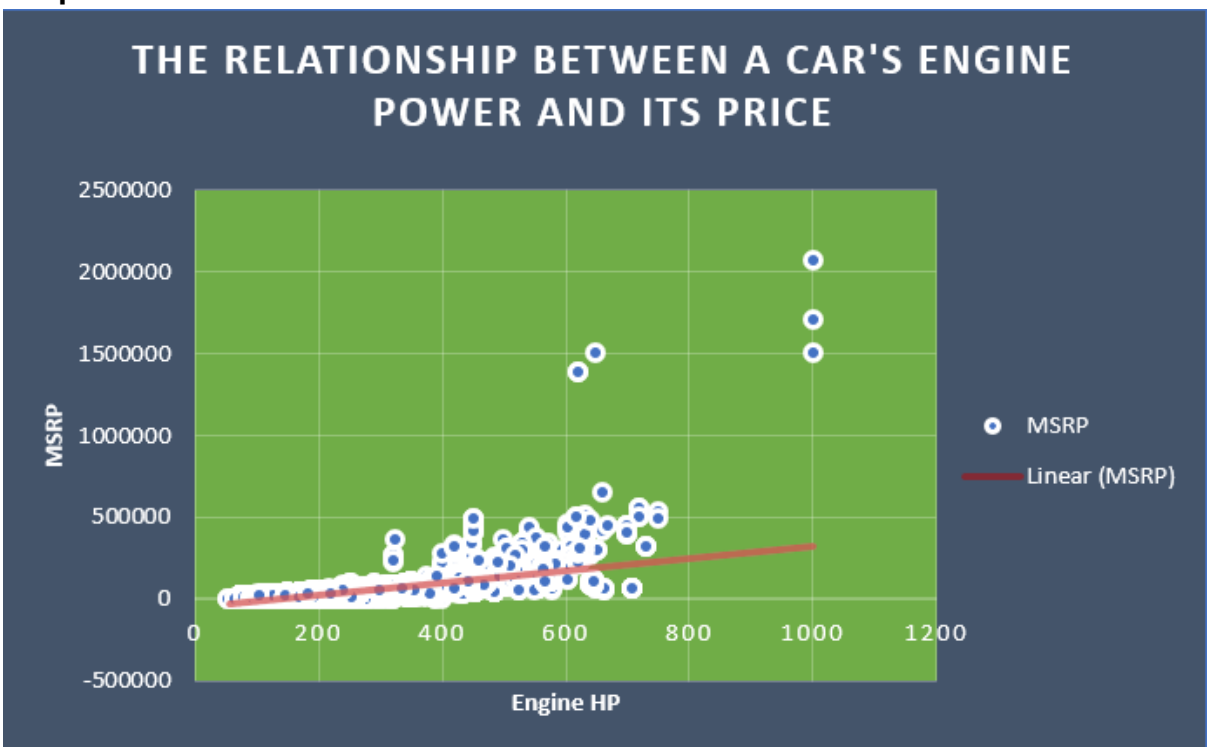
- **Task 1. B:** Create a combo chart that visualizes the relationship between market category and popularity.
  Output:



Insights: The market categories with the highest popularity score (5657) are "crossover, flexfuel, performance" "flexfuel, diesel", and "hatchback, flexfuel". The market category with the most models is N/A (3362), followed by crossover (1068).

2. 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.
  Output:



Insights: The price of an automobile is directly correlated with the engine power of that vehicle. In other words, the price of the car will rise in tandem with an increase in engine power. furthermore, there is a discernible positive link between engine power (hp) and car pricing.

3. 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.

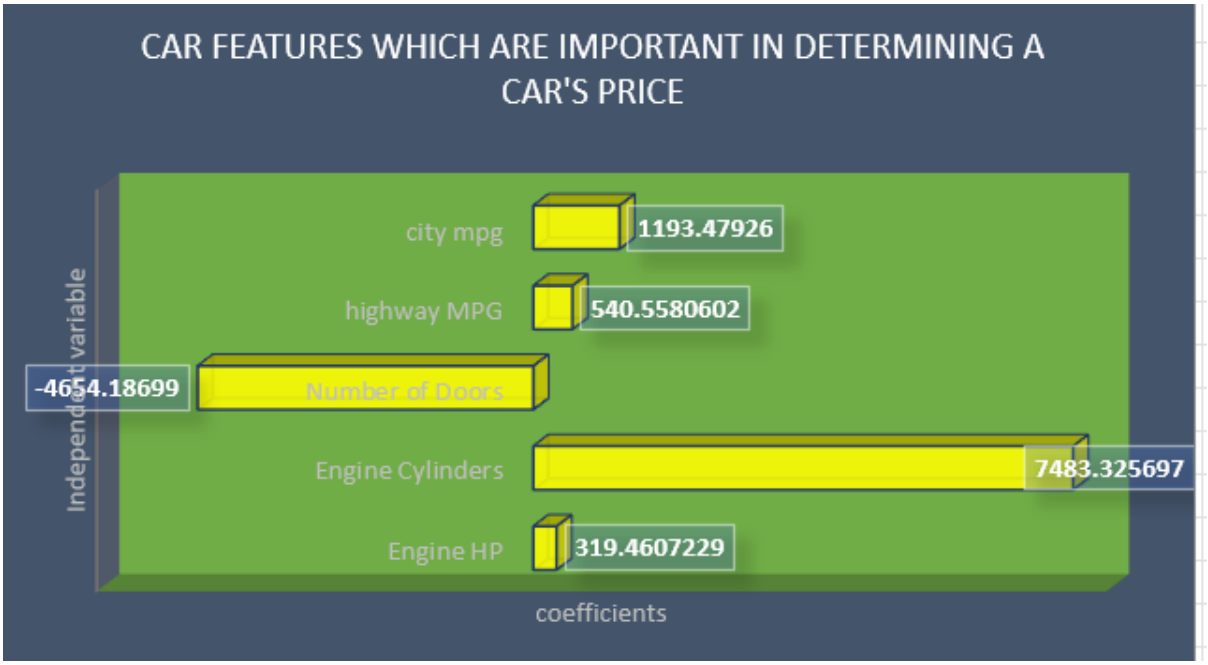
Summary:

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.678347
R Square	0.460155
Adjusted R	0.459912
Standard E	45366.26
Observatio	11097

ANOVA						
	df	SS	MS	F	Significance F	
Regression	5	1.94568E+13	3.89137E+12	1890.759	0	
Residual	11091	2.28264E+13	2058097531			
Total	11096	4.22832E+13				

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95%	Upper 95%
Intercept	-102835	3893.782161	-26.4099912	4E-149	-110467	-95202.2	-110467	-95202.2
Engine HP	319.4607	6.417508778	49.77955371	0	306.8813	332.0402	306.8813	332.0402
Engine Cyli	7483.326	464.1300157	16.12333925	7.93E-58	6573.548	8393.103	6573.548	8393.103
Number of	-4654.19	498.8130353	-9.33052399	1.25E-20	-5631.95	-3676.42	-5631.95	-3676.42
highway M	540.5581	109.9295838	4.917311986	8.9E-07	325.0765	756.0396	325.0765	756.0396
city mpg	1193.479	126.3627723	9.444864481	4.27E-21	945.7858	1441.173	945.7858	1441.173

Output:



Insights: The above bar chart displays each variable's coefficient values, allowing us to visualize their relative importance.
The engine cylinder has the highest correlation with car pricing, at 7483.326, while the number of doors has the lowest correlation (-4654.19).
As a result, "engine cylinders" is the variable that has the strongest relationship with car pricing.

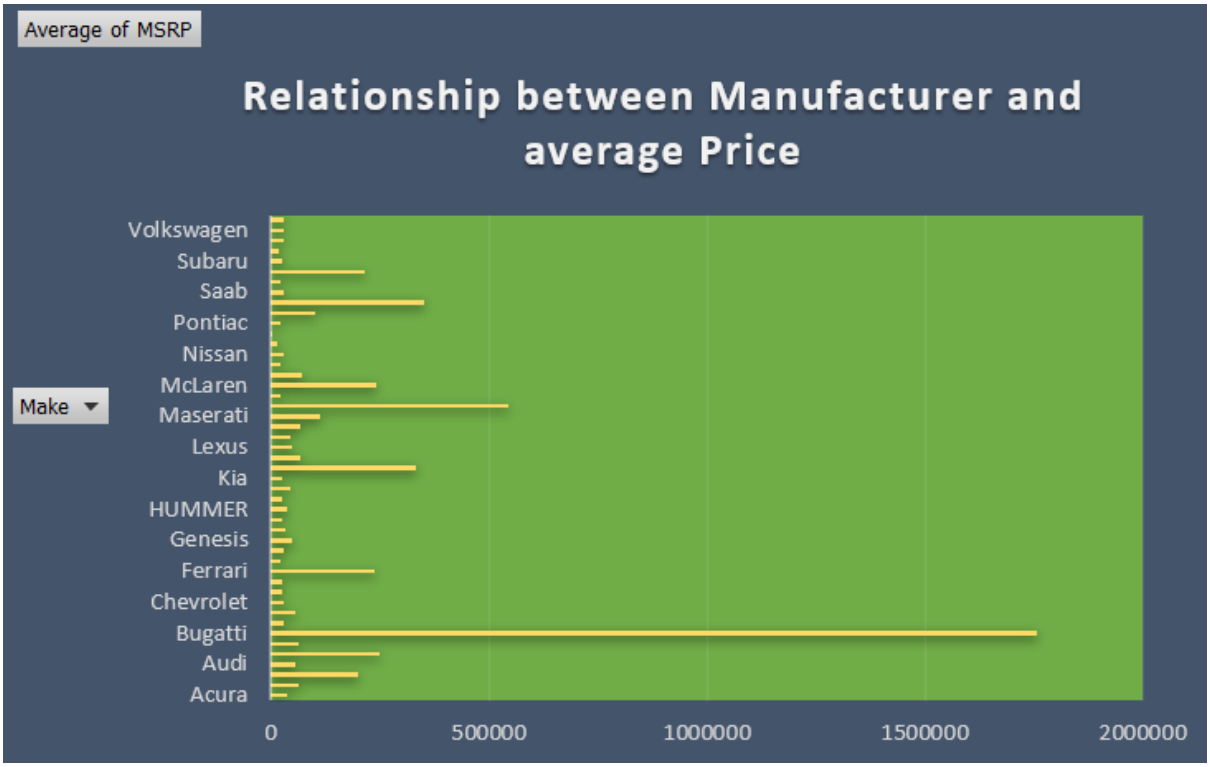
This suggests that a car's price is primarily influenced by its cylinder count, with the number of doors having the least bearing on the price.

4. 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.
  Output:

Make	Average of MSRP
Acura	35087.4878
Alfa Romeo	61600
Aston Martin	198123.4615
Audi	54574.1215
Bentley	247163.3243
BMW	62162.55864
Bugatti	1757223.667
Buick	29034.18947
Cadillac	56368.26515
Chevrolet	29000.2214
Chrysler	26722.96257
Dodge	24857.04537
Ferrari	237383.8235
FIAT	22206.01635
Ford	28522.86207
Genesis	46616.66667
GMC	32444.08506
Honda	26608.88399
HUMMER	36464.41176
Hyundai	24926.26255
Infiniti	42640.27134
Kia	25318.75
Lamborghini	331567.3077
Land Rover	68067.08633
Lexus	47549.06931
Lincoln	43560.01316
Lotus	68377.14286
Maserati	113684.4909
Maybach	546221.875
Mazda	20106.55612
McLaren	239805
Mercedes-Ben	72135.02647
Mitsubishi	21316.35122
Nissan	28856.42329
Oldsmobile	12843.79545
Plymouth	3296.873239
Pontiac	19800.0442
Porsche	101622.3971
Rolls-Royce	351130.6452
Saab	27879.80734
Scion	19932.5
Spyker	214990
Subaru	24240.67364
Suzuki	18021.0531
Toyota	28758.76676
Volkswagen	28947.36879
Volvo	29724.68421
Grand Total	**41901.119**

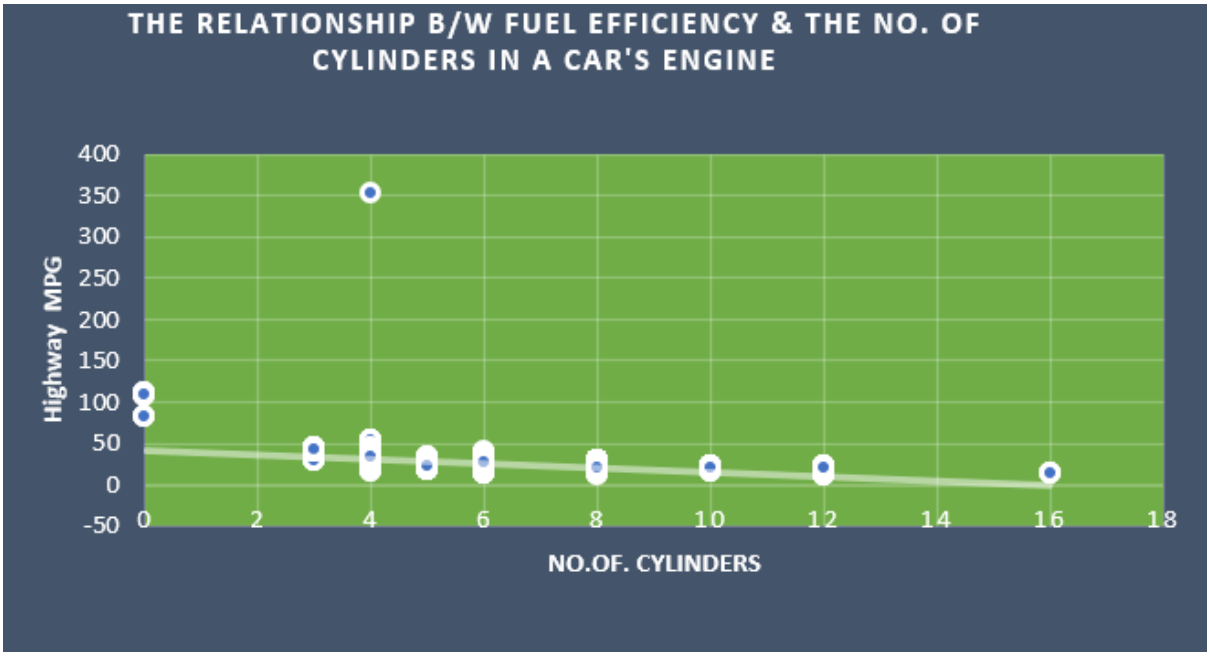
- **Task 4. B:** Create a bar chart or a horizontal stacked bar chart that visualizes the relationship between the manufacturer and the average price.
  Output:



Insights: By analyzing this data, I discovered that Bugatti has the highest average price of any brand, at $175723.7, followed by Maybach at $546221.9.
These details can support pricing and positioning strategies for auto manufacturers.

5. 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.
  Output:



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

  Output:

correlation coefficient	**-0.6147**

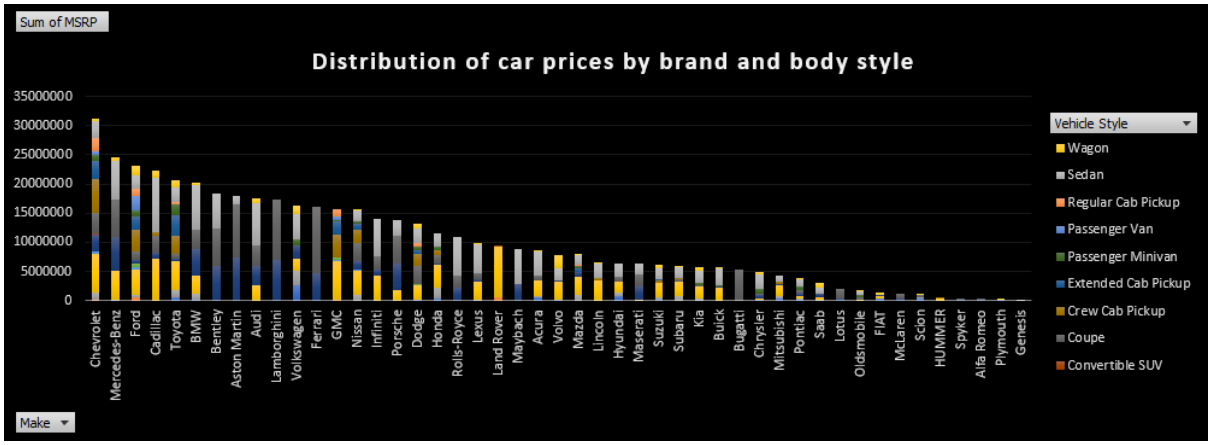
Insight:

- The analysis reveals a negative correlation between the two variables, with highway MPG tending to decrease as the number of engine cylinders increases.
- For example, cars with 16 cylinders had the lowest highway MPG of 14, illustrating the relationship between engine capacity and fuel efficiency.
- Furthermore, a correlation value of -0.6147 between the number of cylinders and highway MPG indicates a significant negative relationship.
- According to this statistical indicator, highway MPG tends to decline with an increasing cylinder count, quantifying a stronger and negative association.

Building the Dashboard

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

Output:

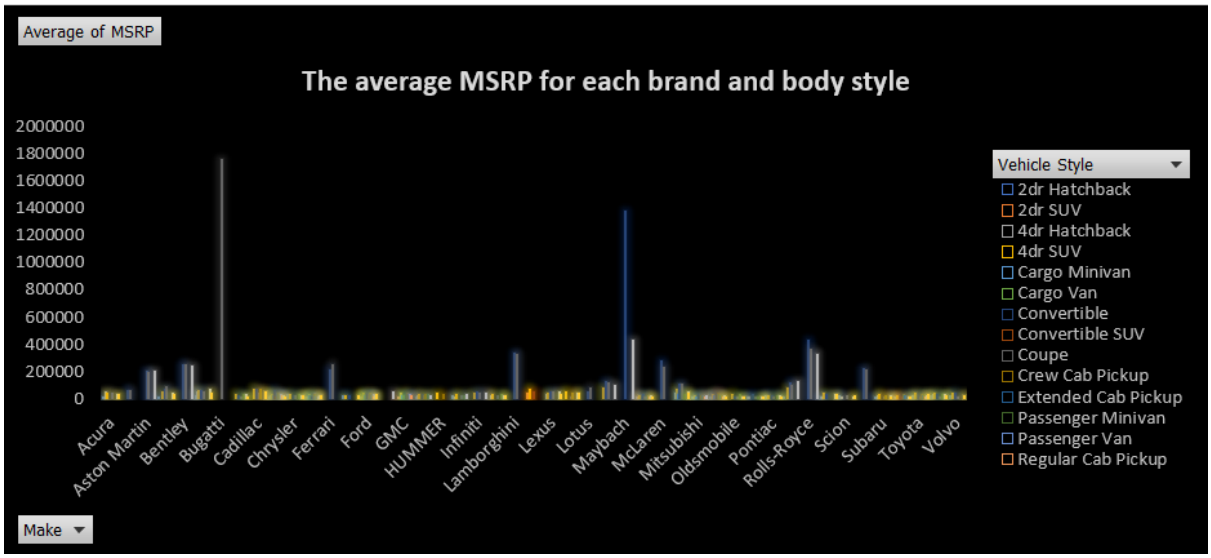


Insight:

- Upon conducting this investigation, I discovered that, regarding overall prices, Chevrolet ranks first with $31175238, followed by Mercedes Benz with $24525909.
- Sedans are the most expensive car type, costing $112811251, followed by four-door SUVs at $98826386.
- These observations offer valuable guidance for the automotive sector's pricing, product development, and marketing strategies.

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

Output:



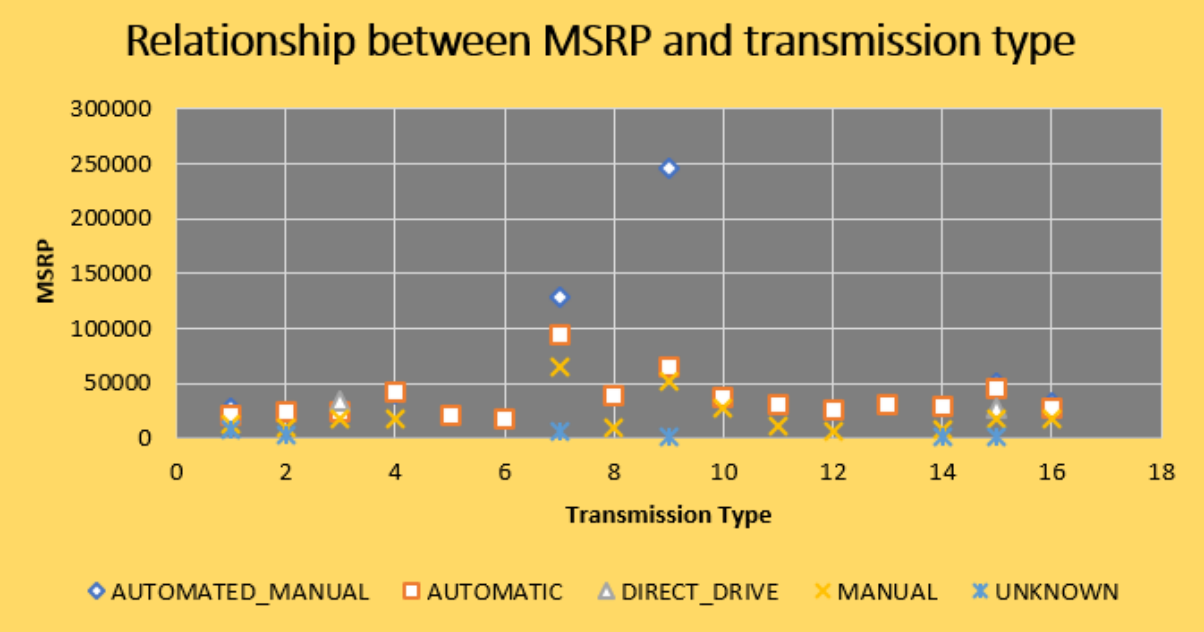
Insight:

- At $1757223.667, Bugatti has the highest average price according to brand analysis. whereas Maybach, at $546221.875, is the second most affordable.
- In terms of body style, convertibles have the highest average price at $88216.79, while coupes rank second at $78292.5.
- On the other hand, at $3296.87, Plymouth has the lowest average price per brand. Oldsmobile, at $12843.79, is in second place.
- In terms of body design, the cheapest two-door SUVs cost $14306.54, while the most expensive two-door hatchback costs $16063.15159.

- This study provides useful data on pricing trends across several auto markets.

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

Output:



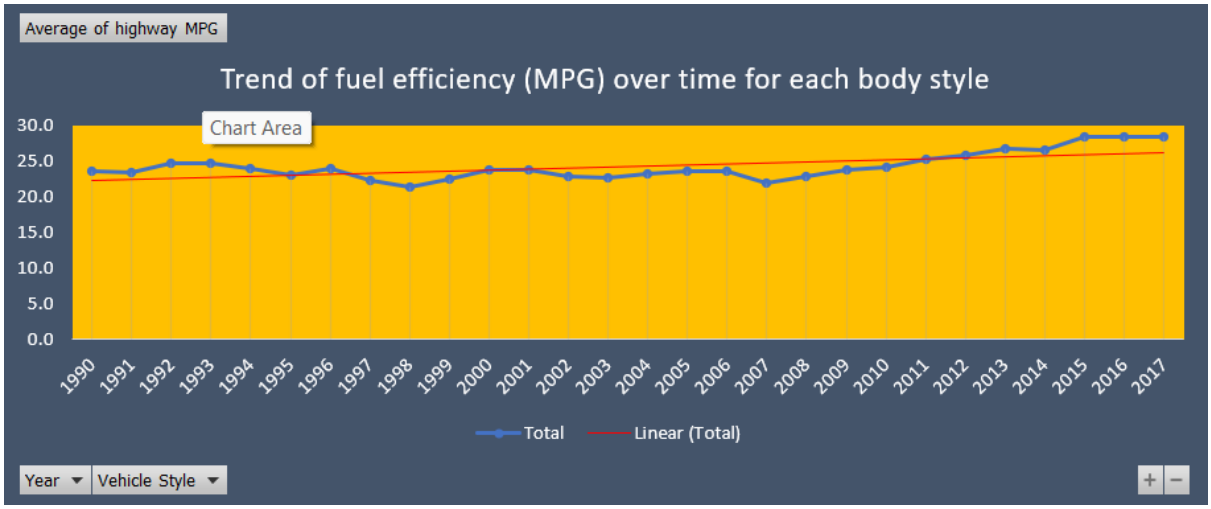
Insight:

- **'Automated-manual transmission'** is the most expensive type of transmission, with an average price of $108381.52., while **'UNKNOWN transmission'** is the least expensive, with an average price of $3647.833

- When it comes to body shape, "convertibles" have the highest average price at $88216.8, followed by "coupes," which come in second at $78292.5.
- The least expensive body type is a 2-door SUV, which costs $14306.54 on average.

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

Output:

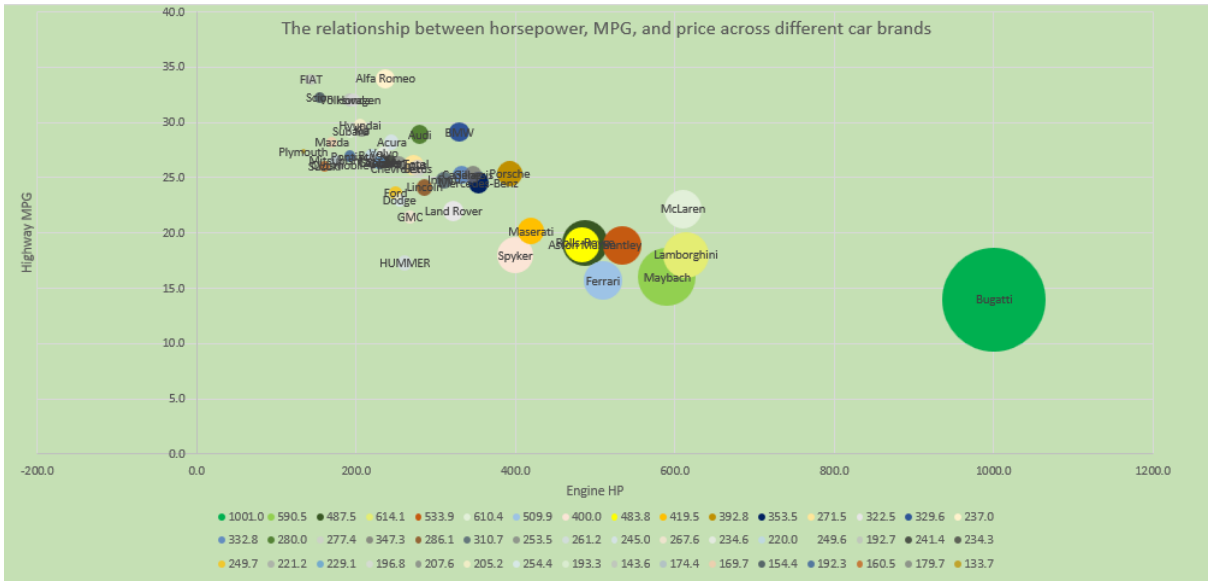


Insight:

- My analysis suggests that there may be a trend toward greater fuel economy as car model years go by.
- This trend could be linked to advancements in automotive technology and policies aimed at lowering emissions and boosting fuel economy.

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

Output:



Insight:

- The above bubble chart displays the average horsepower, MPG, and MSRP for each car brand.

The highest average engine horsepower is for the Bugatti brand (1001), the highest average highway MPG is for the Alfa Romeo, and the highest average MSRP is for the Bugatti (1757223.7).

Key insights:

We found several key insights, such as the significant impact of engine power on car prices, the varying popularity of car models across market categories, and the relationship between fuel efficiency and engine specifications.

Relevance to Business Problem:

By providing practical information to automakers, helping them prioritize product development, find lucrative market niches, and improve pricing strategies, these insights immediately address the business problem.

Recommendations:

Based on the knowledge acquired, we advise automakers to concentrate on creating fuel-efficient models, strategically determine the pricing of automobiles according to the value of their features, and modify their marketing plans to focus on niche markets.

Result:

Visualization:

To display the analysis's findings, we employed visual aids like bar charts, scatter plots, and pivot tables. Stakeholders can make decisions more easily and more effectively by using these visuals to help them understand the findings.

Discussion:

Car manufacturers will be greatly impacted by the analysis's findings, which offer insightful information on consumer preferences, price dynamics, and market trends. The results of this study can inform strategic choices that aim to improve profitability and competitiveness.

Future Directions:

- In the future, more research may examine dynamic pricing schemes, include current market data, and make use of cutting-edge machine learning methods for predictive modeling.
- Furthermore, maintaining competitiveness in the automobile sector requires constant observation of consumer preferences and market developments.
- Overall, the knowledge gathered from this analysis offers automakers insightful advice on how to best approach pricing and new product development in a market that is changing quickly.

EXCEL workbook link:
P7- Analyzing the impact of car features on price and profitability.xlsx

Project video link:
https://www.tella.tv/video/manasas-video-et1x