

2023 | By: ATTULAYA SINGH



Analyzing the Impact of Car Features on Price

Project description

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.

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?

This problem could be approached by analyzing the relationship between a car's features, market category, and pricing, and identifying which features and categories are most popular among consumers and most profitable for the manufacturer. By using data analysis techniques such as regression analysis and market segmentation, the manufacturer could develop a pricing strategy that balances consumer demand with profitability, and identify which product features to focus on in future product development efforts. This could help the manufacturer improve its competitiveness in the market and increase its profitability over time

PROJECT PROBLEM:

Analyzing the relationship between a car's features and its popularity is a crucial project in order to assist manufacturers in making informed decisions about product development and marketing strategies. By examining the popularity variable in the dataset, data analysts can determine which features are most favored by consumers and how these features impact a car's popularity in the market.

Another significant aspect of the project involves predicting the price of a car based on its features and market category. By utilizing the various features and market category variables present in the dataset, data analysts can develop a model to accurately forecast the price of a car. This predictive model can prove invaluable for both manufacturers and consumers, as it enables a better understanding of how different features influence the overall price of a car. This information can then be used to make well-informed decisions regarding pricing and purchasing.

Overall, the dataset of cars provides a valuable resource for data analysts seeking to explore various aspects of the automotive industry. The insights derived from this analysis can play a pivotal role in shaping decisions related to product development, marketing strategies, and pricing. Prior to conducting the analysis, it is essential to thoroughly comprehend the dataset and perform data cleanup procedures. This involves removing rows with blank cells and eliminating duplicate entries to ensure the dataset is accurate and reliable.

Approach:

In this project I have used regression analysis to find the relation between the car prices and many different variables as to find the results for the other features and price related analysis I have used the count and mean of the features

Tech Stack Used:

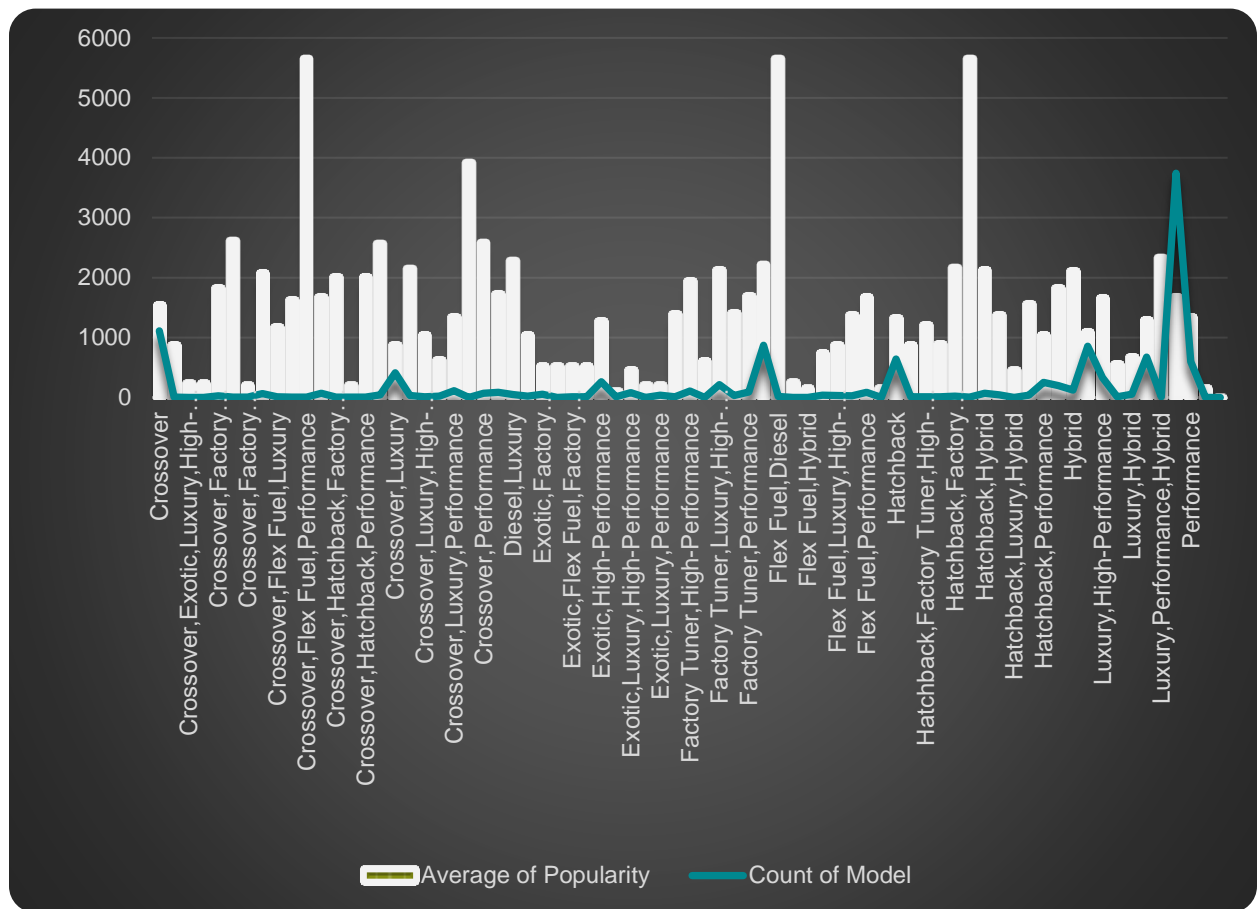
- MS excel
- Tableau
- Data tabs

Tasks: Analysis

To ensure accurate and reliable results in the analysis of the given dataset, it is crucial to perform thorough data cleaning. This process involves several steps to eliminate any inconsistencies, missing values, or duplicate entries that could potentially affect the analysis

- 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

Market Category	Average of Popularity	Count of Model
Crossover	1545.263063	1110
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	884.5487805	410



Insights:

From this we get to know that the most popular market categories are :

- Crossover,Flex Fuel,Performance
- Flex Fuel,Diesel
- Hatchback,Flex Fuel

But the count of these model were very less

So we observe that the model with a decent popularity and a high number of sales was

Crossover 1545.263063 1110

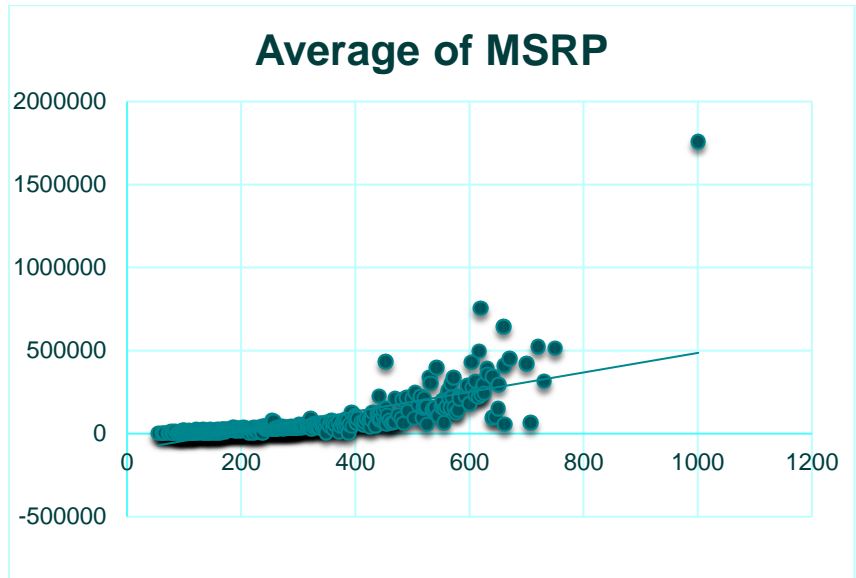
And the other category was not classified as this had the highest popularity and also the highest count i.e

N/A 1676.889364 3742

So we can infer that the cars with decent numbers of count also has decent popularity and also if the popularity score is high then the count of the model is also less reason might be many for the popularity one of which can be the rarity

- 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.

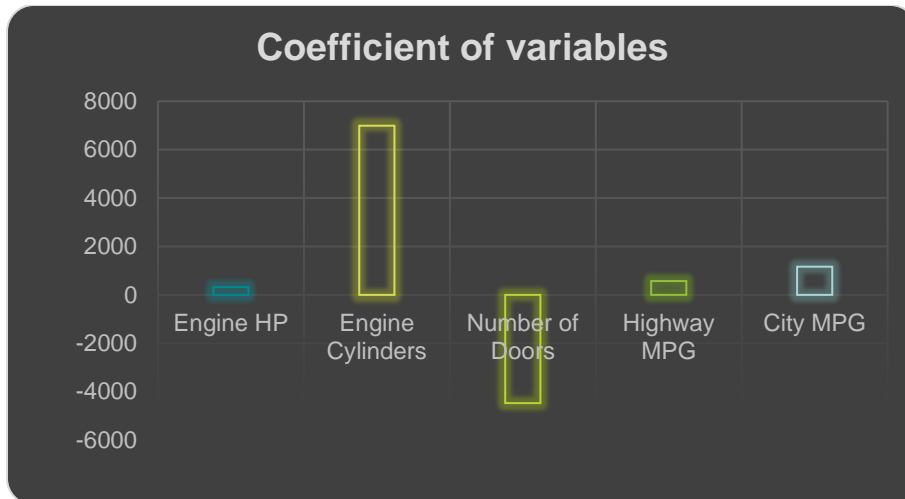
Engine hp	Average of MSRP
55	\$ 2,000.00
62	\$ 2,000.00
63	\$ 2,000.00
66	\$ 2,000.00
73	\$ 2,000.00
74	\$ 8,116.94
78	\$ 15,082.50
79	\$ 5,584.25
81	\$ 2,000.00
82	\$ 2,000.00
84	\$ 14,493.33
88	\$ 2,000.00
90	\$ 2,000.00
92	\$ 2,000.00
93	\$ 2,013.54
94	\$ 7,667.50
95	\$ 2,005.73
96	\$ 2,000.00
97	\$ 2,000.00
98	\$ 18,931.56
99	\$ 21,377.69
100	\$ 5,685.78
101	\$ 18,061.36
102	\$ 2,000.00
103	\$ 13,152.50
105	\$ 2,000.00
106	\$ 14,055.39
107	\$ 10,857.50
108	\$ 11,488.39
109	\$ 15,747.50
110	\$ 7,424.22
111	\$ 10,500.00
113	\$ 2,000.00
114	\$ 2,000.00
115	\$ 12,758.48
116	\$ 2,057.40
118	\$ 2,000.00



Insights: from this analysis we can infer that as the engine hp increases the price of the car also increases linearly so this is one of the factors on which the price of the car depends upon horse power of the engine

- 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 visualize their relative importance

Engine HP	Engine Cylinders	Number of Doors	Highway MPG	City MPG	MSRP
335	6	2	26	19	46135
300	6	2	28	19	40650
300	6	2	28	20	36350
230	6	2	28	18	29450
230	6	2	28	18	34500
230	6	2	28	18	31200
300	6	2	26	17	44100
300	6	2	28	20	39300
230	6	2	28	18	36900
230	6	2	27	18	37200
300	6	2	28	20	39600
230	6	2	28	19	31500
300	6	2	28	19	44400
230	6	2	28	19	37200
230	6	2	28	19	31500
320	6	2	25	18	48250
320	6	2	28	20	43550
172	6	4	24	17	2000
172	6	4	24	17	2000
172	6	4	20	16	2000
172	6	4	24	17	2000
172	6	4	21	16	2000
172	6	4	24	17	2000
172	6	4	20	16	2000
172	6	4	24	17	2000
172	6	4	24	17	2000
172	6	4	21	16	2000
172	6	4	21	16	2000
172	6	4	22	16	2000
172	6	4	22	17	2000
172	6	4	22	16	2000
172	6	4	21	16	2000
160	4	2	35	26	27495
160	4	2	35	26	24995
160	4	2	35	26	28195
130	4	4	26	18	2000
158	6	4	25	17	2000
158	6	4	25	17	2000
130	4	4	26	18	2000



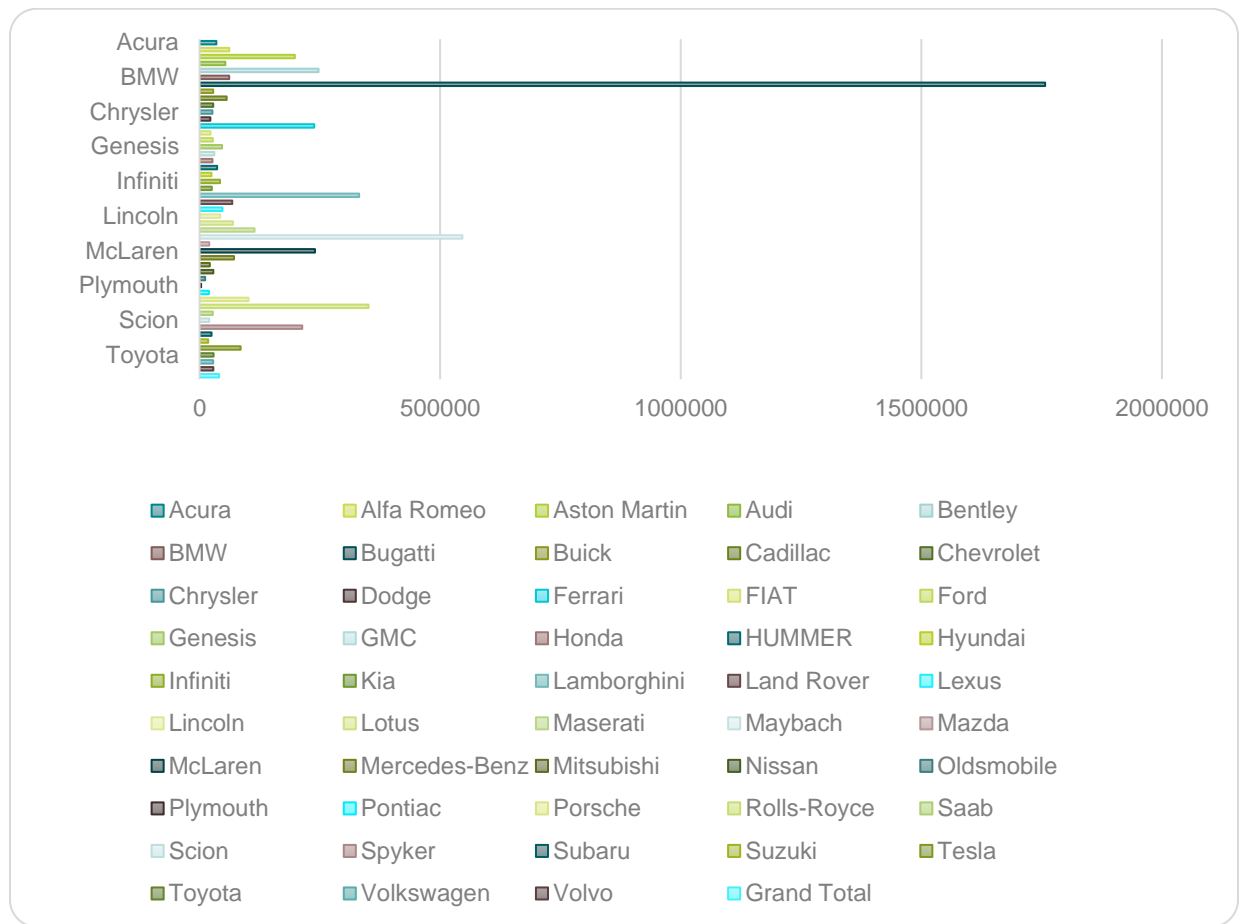
the r value of engine cylinder is the highest

Insight:

From the analysis we infer that the engine cylinders has the highest coefficient of variable and also we can say that it has the best relationship with the price of the car

- **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.

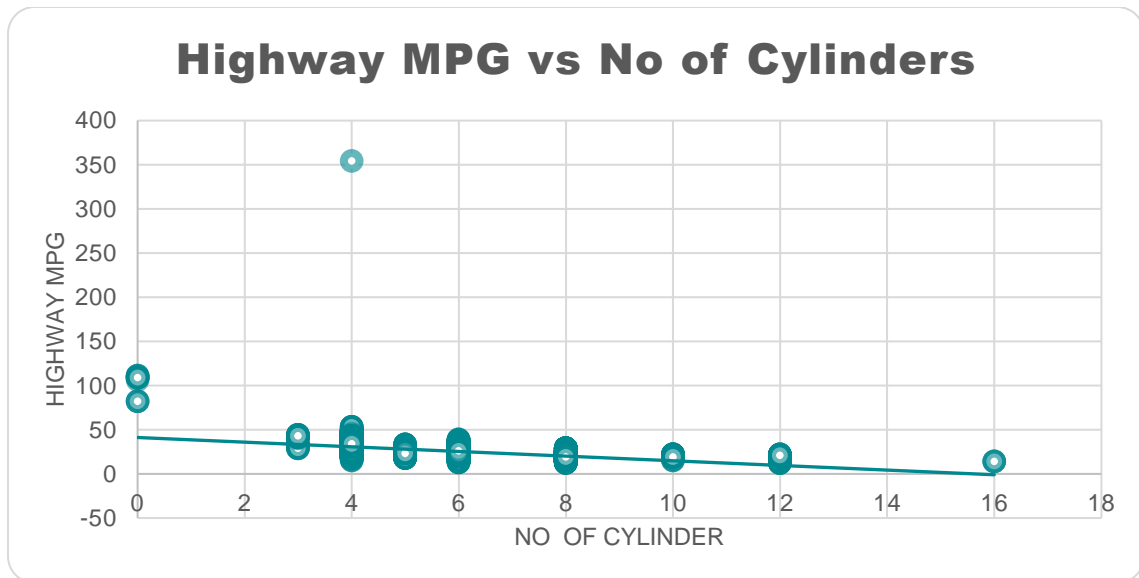
Make	Average of MSRP
Acura	34887.5873
Alfa Romeo	61600
Aston Martin	197910.3763
Audi	53452.1128
Bentley	247169.3243
BMW	61546.76347
Bugatti	1757223.667
Buick	28206.61224
Cadillac	56231.31738
Chevrolet	28350.38557
Chrysler	26722.96257
Dodge	22390.05911
Ferrari	238218.8406
FIAT	22670.24194
Ford	27399.26674
Genesis	46616.66667
GMC	30493.29903
Honda	26674.34076
HUMMER	36464.41176
Hyundai	24597.0363



Insight:

The highest average MSRP is of BUGGATI i.e \$1757223.667 this might be due to the reason that Buggati manufactures mostly performance oriented cars

- **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.



- Correlation coefficient= **-0.620312551**

Insight: The highway Milage is inversely proportional to the number of cylinders present in the car

Link to the Dashboard to visualize the relation between price and other features of car:
[Tableau dashboard](#)

Conclusion:

- The popularity of cars are based on different categories and people mostly prefer Crossover, Flex Fuel, Performance cars as it offers the best combination of efficiency and power at the same time
- The price of the Car depends on the various features such as the number of cylinder , milage and engine hp