PROJECT – 7 Impact of Car Features

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.

THE PROBLEM

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.

The project description file containing specific analysis tasks was unfortunately deleted from website. To proceed with the analysis, I relied on my own knowledge, internet and expertise in the subject matter. While this posed a challenge, I conducted the analysis based on my understanding of the project requirements and relevant methodologies. Although the original tasks from the file were not available, I ensured that the analysis was thorough and aligned with the project's objectives to the best of my ability.

DESIGN

DATA CLEANING: This step involves preprocessing the data to make it suitable for analysis. It includes handling missing values, dropping unwanted columns, removing duplicates, converting data types if necessary, and possibly feature engineering. Dropped 715 duplicate records.

HIGHWAY-CITY MPG: Created a new column with difference of higway mps and city mpg for outlier detection.

KNN-IMPUTATION: For imputing null values in Market Category column employed a Knn-classifier in Python. Null values in other columns were filled with online information.

EDA: Performed EDA on the data to understand the relationships between different variables. Used various Excel functions for descriptive statistics (COUNTIF, AVERAGE, MEDIAN, MODE, MAX, MIN, VAR, STDEV), and correlation analysis (CORREL). Used various visualizations tools to showcase the insiders story properly.

DUMMY VARIABLES: Created dummy variables from various categorical columns like Engine Fuel Type, Transmission Type, Driven Wheels etc. for regression model.

MULTIPLE LINEAR REGRESSION: Employed a multiple linear regression model for MSRP. Highlighted Multicollinear variables.

REPORT AND DATA STORY: Visual narratives take centre stage, employing scatter plots and charts to illuminate hidden relationships and reveal captivating trends.

MS EXCEL - Excel is a spreadsheet editor developed by Microsoft for Windows. It features calculation or computation capabilities, graphing tools, pivot tables, etc.

PYTHON - Python is a versatile and powerful programming language widely used for data analysis.

DESIGN (Contd.)

Handling Missing data

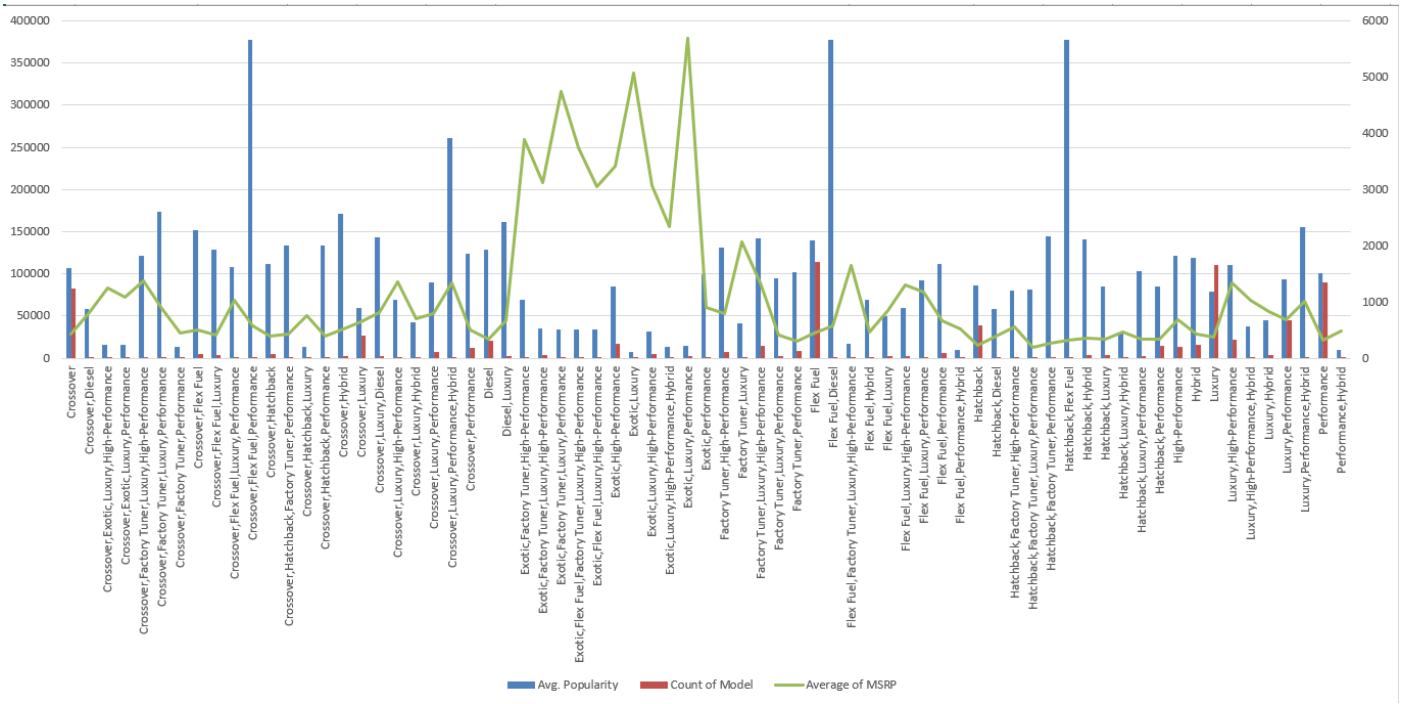
- For columns like Engine Fuel Type, Engine HP, Engine Cylinders, Transmission Type, Number of Doors null values were replaced with information present online.
- For Market Category column, for null values with same make and model, checked for other records with same make and model.
- For remaining null values in Market Category deployed a KNN-Classifier model with an accuracy of 84.6%.

Handling Outliers

- Engine HP - Engine HP have 3 outliers of 1001, but all these were of Buggati Veyron which is a high performing exotic car, so 1001 is justifiable.
- Highway-City mpg
    - For outlier in Highway mpg , created a new column with difference of Highway mpg and city mpg. Negative values in this column were outliers as City mog cannot be larger then Higway mpg ,so replaced these values of Highway mpg with City mpg values.
    - One outlier contained 330 so replaced it with online information.
    - Other outliers in Highway and city mpg were electric cars so didn’t changed.
- MSRP – Outliers in MSRP have High Performance, Exotic, Luxury cars in market category so didn’t changed.

FINDINGS

Avg Popularity and Model Count according to Market Category



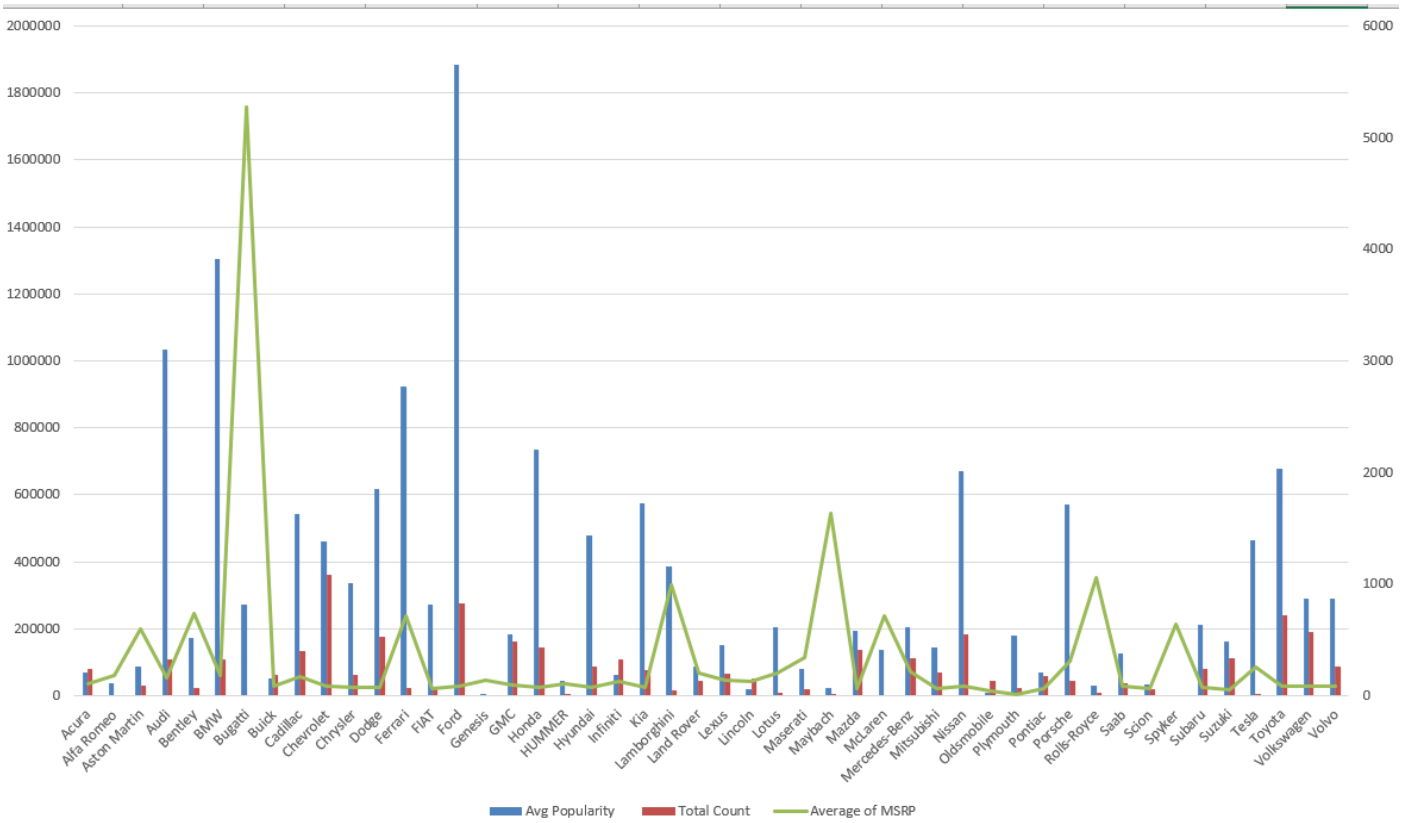
Highest popularity Market Categories – (Crossover,Flex Fuel, Performance) , (Flex Fuel, Diesel) , (Hatchback, Flex Fuel) with 5657 avg popularity score.

Highest Model Count – Flex Fuel , Luxury , Performance , Crossover

Highest Avg MSRP – (Exotic,Luxury,Performance) , (Exotic, Luxury), (Exotic, Factory Tuner,Luxury, Performance)

FINDINGS (Contd.)

Avg Popularity and MSRP according to Make



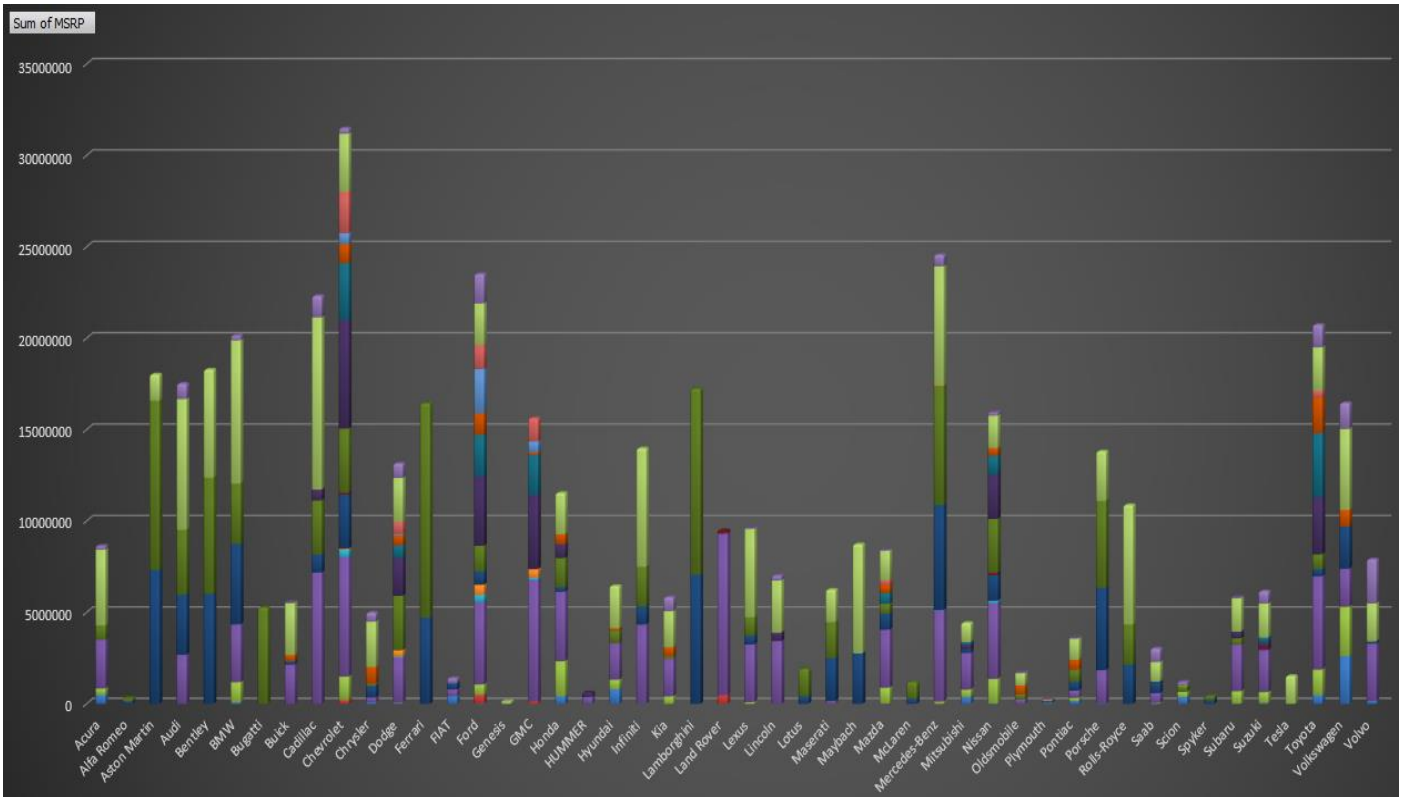
Highest Avg Popularity – Ford tops the chart with 5657, followed by BMW and Audi with 3916 and 3105 respectively.

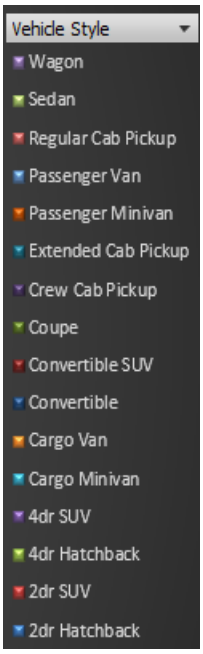
Highest Model Count – Chevrolet, Ford, Toyota

Highest Avg MSRP – Bugatti , Maybach, Rolls-Royce

FINDINGS (Contd.)

Avg Popularity and MSRP according to Make



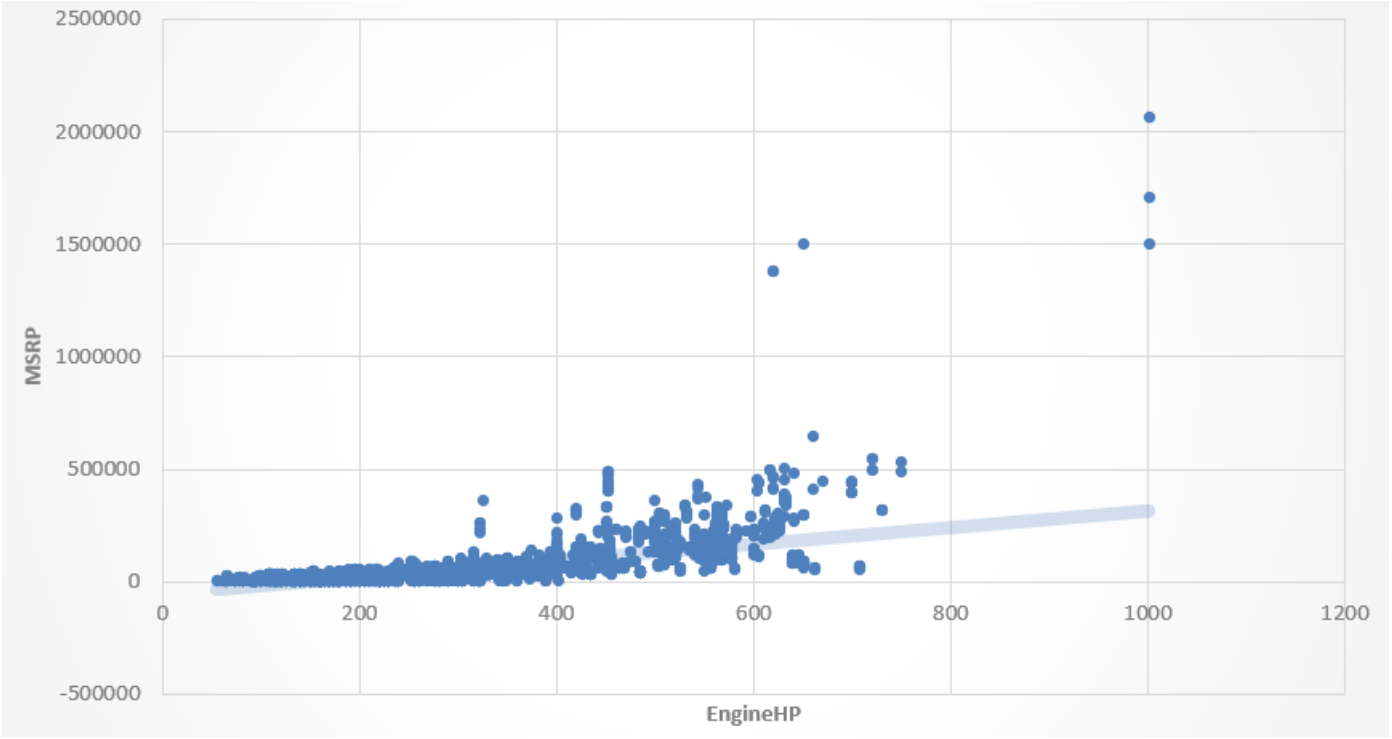


Chevrolet has the highest price distribution by body style.

Followed by Mercedes-Benz but Chevrolet deal with all 16 body styles.

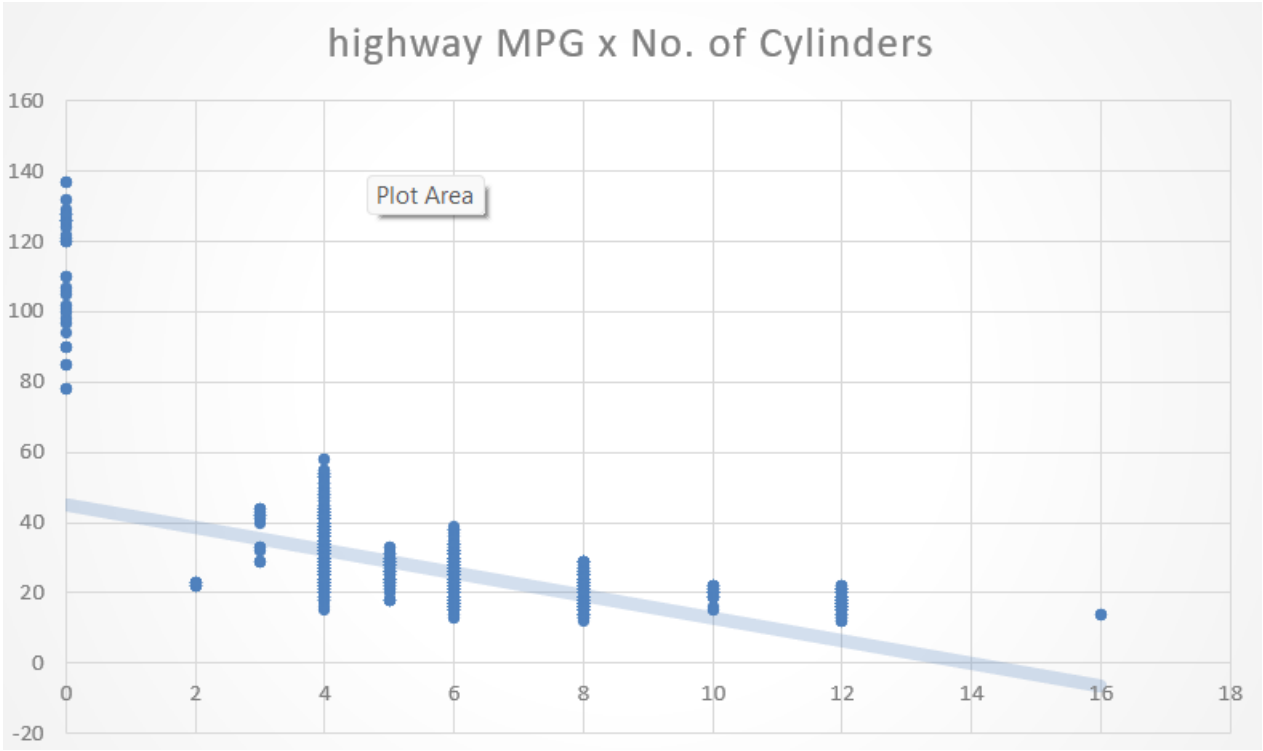
FINDINGS (Contd.)

MSRP increases with increase in Engine HP.



Highway MPG decrease with increasing number of Cylinders.

0 cylinders that is electric cars have highest MPG.



FINDINGS (Contd.)

Correlation of MSRP with dummy variables

premium unleaded (required)	0.455205667
AUTOMATED_MANUAL	0.247398764
rear wheel drive	0.155629194
all wheel drive	0.14142809
flex-fuel (premium unleaded required/E85)	0.13309711
electric	0.007529223
DIRECT_DRIVE	0.006891437
flex-fuel (premium unleaded recommended/E85)	0.005265129
premium unleaded (recommended)	-0.000694414
flex-fuel (unleaded/natural gas)	-0.001027879
diesel	-0.002795375
AUTOMATIC	-0.004251568
natural gas	-0.006704721
flex-fuel (unleaded/E85)	-0.026914273
four wheel drive	-0.031347287
MANUAL	-0.123091592
front wheel drive	-0.240975506
regular unleaded	-0.356039417

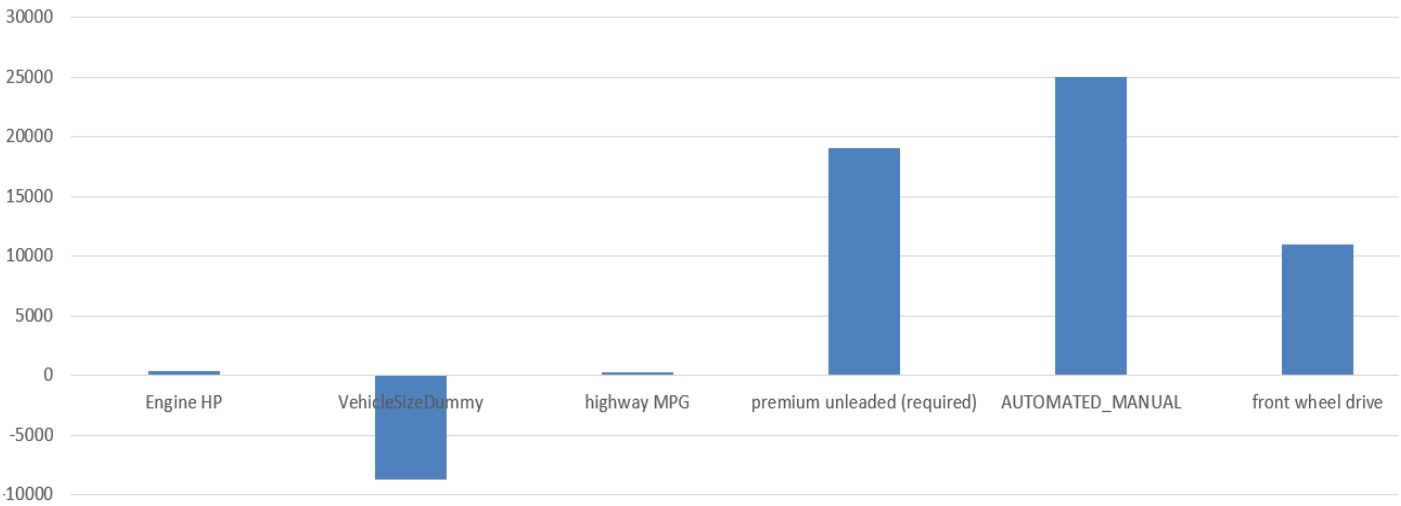
Multicollinearity

Engine Cylinders x Engine HP — Correlation = 0.774

premium unleaded(required) x regular unleaded — Correlation = -0.557

rear wheel drive x front wheel drive — Correlation = -0.506

Correlation



ANALYSIS

Popularity Trends

- Market categories like "Crossover," "Flex Fuel," and "Hatchback" tend to have higher average popularity scores compared to others, indicating a stronger consumer preference for these types of vehicles.
- "Exotic" and "Factory Tuner" categories generally exhibit lower average popularity scores, suggesting a niche market or limited consumer demand for these types of vehicles.
- Ford, Bugatti, and Chevrolet have the highest average popularity scores, indicating a strong consumer interest in these brands.
- Makes like Lotus, Maybach, and Spyker have relatively low average popularity scores, suggesting a niche or limited market presence.

MSRP Analysis

- Market categories associated with luxury and high-performance vehicles, such as "Luxury," "High-Performance," and combinations thereof, tend to have higher average MSRP values. This aligns with the expectation that luxury and high-performance vehicles typically command higher prices in the market.
- Categories like "Performance" and "Hybrid" also show moderate to high average MSRP values, reflecting the value proposition and technological advancements associated with these vehicle types.
- Bugatti, Maybach, and Lamborghini have the highest average MSRP values, reflecting their status as luxury car brands known for producing high-end, expensive vehicles.
- Makes like Plymouth, Oldsmobile, and Scion have significantly lower average MSRP values, indicating more affordable or budget-friendly options.
- Luxury car brands like Rolls-Royce, Bentley, and Aston Martin command high average MSRP values, reflecting their prestige and exclusivity in the market.

ANALYSIS (Contd.)

- Mainstream brands such as Toyota, Honda, and Volkswagen offer more affordable options with lower average MSRP values, appealing to a broader consumer base.
- Brands like Ferrari, Lamborghini, and Porsche exhibit a combination of high average popularity and high average MSRP values, indicating a strong association between performance, prestige, and price in the sports car segment.
- Tesla stands out with a high average popularity score and a relatively high average MSRP value, indicating strong consumer interest in electric vehicles despite their premium pricing.

- Highway mpg decreases with increase in number of cylinders. Vehicles with more engine cylinders often have larger, more powerful engines that consume more fuel, resulting in lower fuel efficiency. Additionally, larger engines typically generate more power, which may be desirable for performance but can lead to decreased fuel economy.
- Electric vehicles with 0 Engine Cylinders represent a highly efficient and environmentally friendly alternative to traditional internal combustion engine vehicles, offering the potential for significant reductions in greenhouse gas emissions and energy consumption in the transportation sector.

CONCLUSION

In conclusion, the analysis of car features in this project has provided valuable insights into the factors influencing vehicle popularity, pricing, and fuel efficiency etc. By examining a diverse range of car makes and attributes such as engine cylinders, fuel type, transmission type, driven wheels etc., we've gained a deeper understanding of consumer preferences and market trends within the automotive industry.

From the inverse relationship between highway mpg and the number of engine cylinders to the high efficiency of electric vehicles, these findings underscore the importance of technological advancements, environmental considerations, and consumer demand in shaping the landscape of the automotive market.

Moving forward, leveraging these insights can inform strategic decisions for manufacturers, policymakers, and consumers alike, driving innovation and sustainability in the future of transportation.