ANALYZING THE IMPACT OF CAR FEATURES ON PRICE AND PROFITABILITY

Project Description: This project focuses on analyzing a car dataset to extract meaningful insights related to pricing, fuel efficiency, and brand performance. The goal is to understand patterns in car prices, how features like engine size and body style affect market value, and how fuel efficiency varies across models.

The dataset used contains information on over 11,900 cars, including variables like Make, Model, Year, Engine HP, Transmission Type, MSRP, Fuel Type, MPG, and more.

Approach:

The analytical approach for this project was centered entirely around Excel Pivot Tables and data visualizations. Pivot tables were used to summarize and analyze the dataset by aggregating values such as average MSRP, total MSRP, fuel efficiency, and other metrics across various categories like car brand, body style, transmission type, and model year. This method allowed for dynamic data exploration and quick extraction of meaningful patterns without the need for complex formulas or additional functions.

We utilized a variety of chart types—such as stacked column charts, clustered column charts, scatter plots, and line graphs—to visualize relationships and trends in the data. Each chart was made interactive using slicers and filters, enabling users to drill down into specific aspects of the dataset easily. The combination of pivot tables and interactive charts formed the core analytical technique throughout the project, providing a user-friendly way to answer business questions and generate actionable insights.

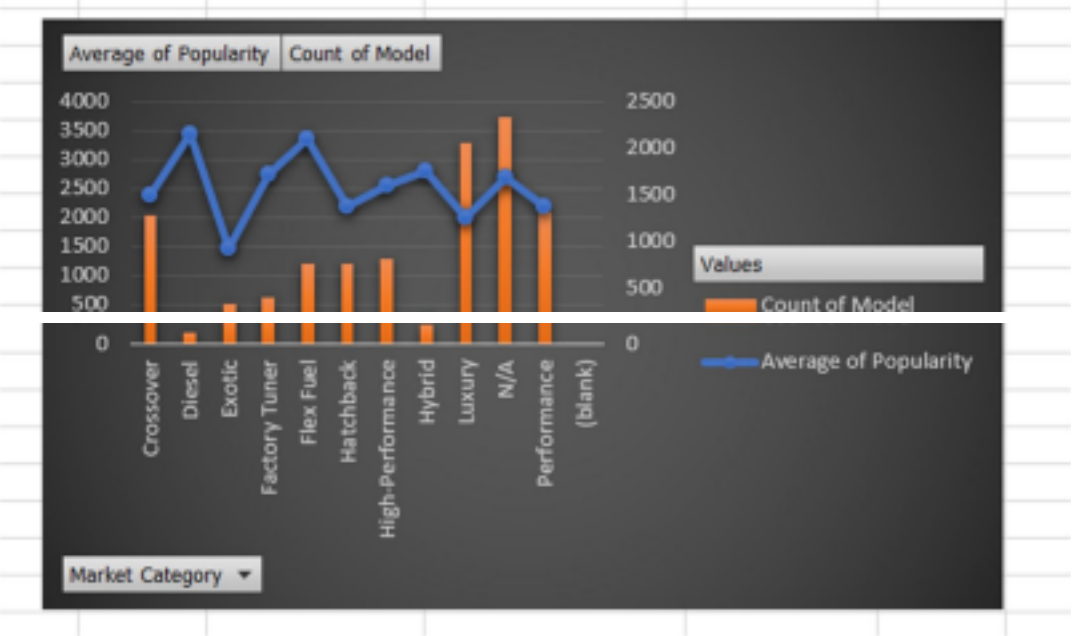
Tech stack used:

Microsoft Excel for the entire workflow: data cleaning, pivot tables, and visualization which provides:

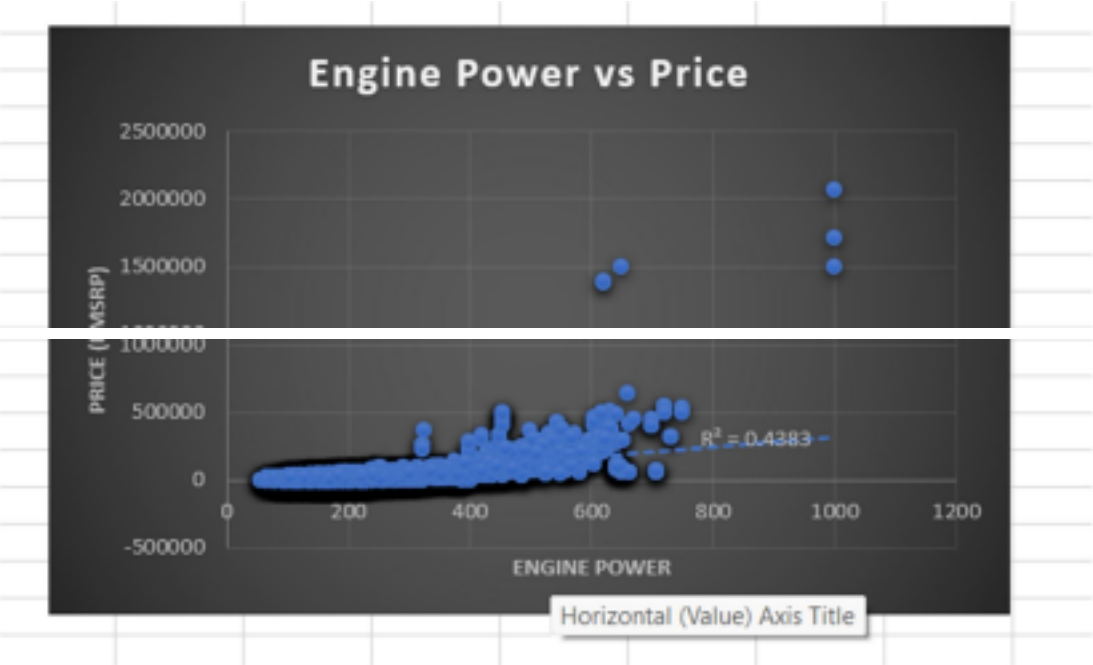
1. Highly accessible and widely used.
2. Offers dynamic features like slicers, filters, and interactive dashboards.

Excel's Data Analysis Toolpak for correlation coefficient.

Insights and Results:



Above figure shows the relationship between different market categories with that of popularity of the model and the count of model. Popularity peaks for diesel market category whereas highest count of model turns out to be unknown followed by luxury market category.



Above figure shows the scatter plot between engine power and MSRP (price). Relationship is found to be linearly increasing.

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.674145591
R Square	0.454472278
Adjusted R Square	0.454151541
Standard Error	44409.54215
Observations	11914

The above figure shows the regression statistics.

Multiple R value suggests a moderate positive relationship between predictors and target.

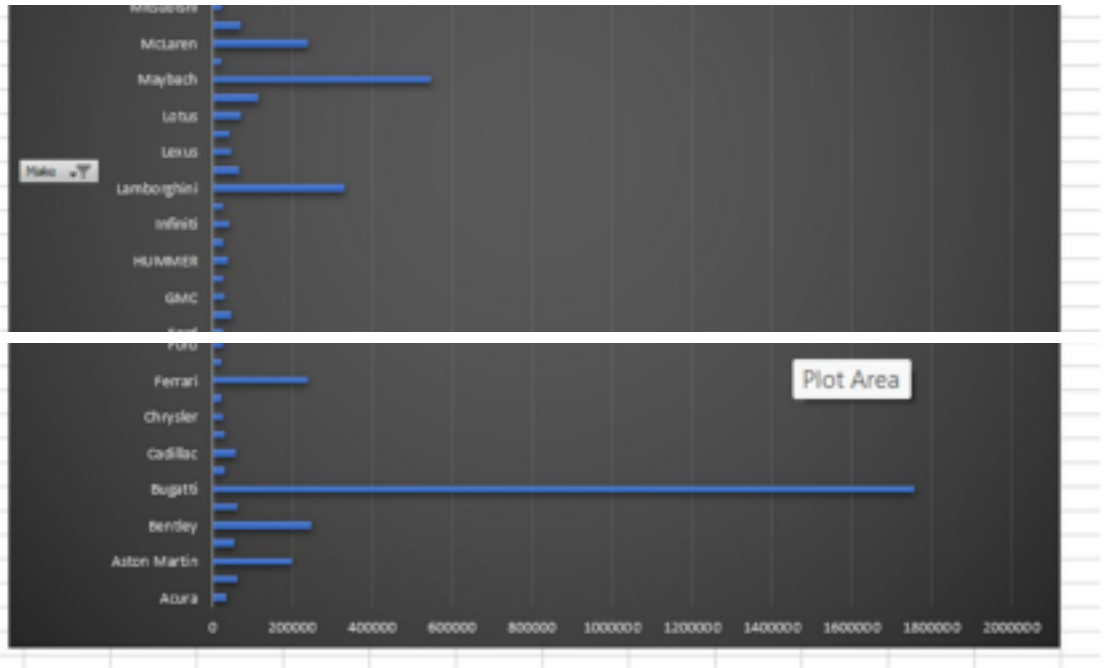
Adjusted R square close to R square indicates a good number of relevant predictors.

Standard error indicates error in predictions.

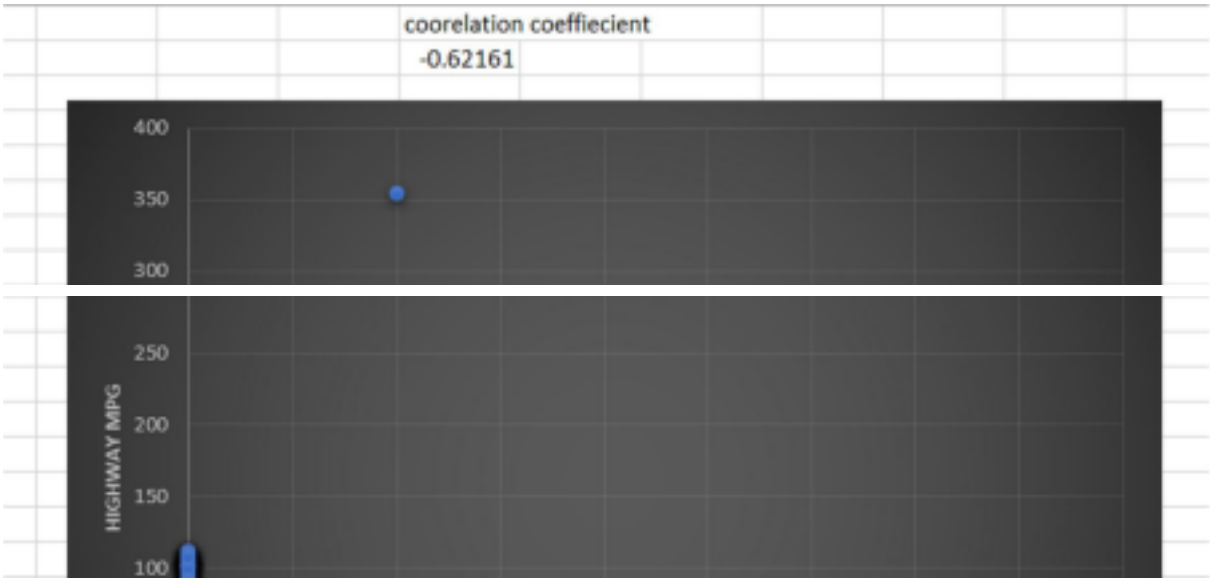
Observations refer to sample size which in this case is a solid amount.

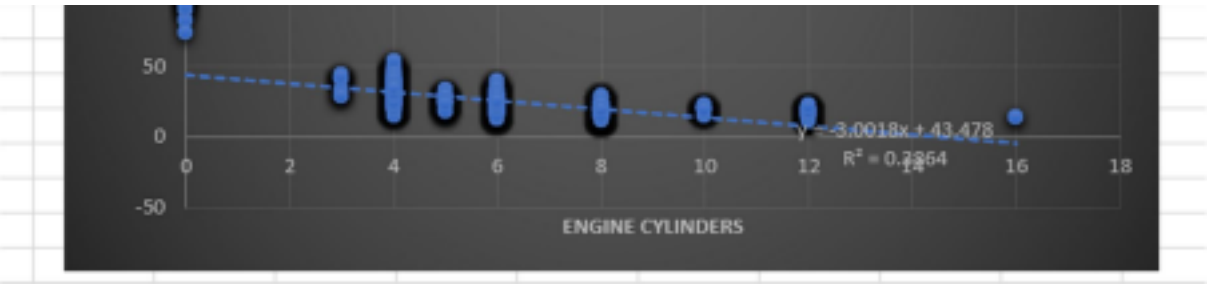
Model's p-value comes out to be near zero from further analysis, which is highly significant and f-score turns out to be very high, indicating overall regression model is statistically significant.

The only insignificant predictor: city MPG.



The above figure shows average of MSRP based on the brand (make). Highest price goes to the Bugatti model.





The above figure shows the relationship between highway mpg with engine cylinders which turns out to be linearly decreasing, i.e. inversely proportional. The correlation coefficient is close to -1 suggesting the same.

Hyperlink: car_features_.xlsx