

Analyzing the Impact of Car Features on Price and Profitability

1. Project Description

Overview of the Project and Its Purpose

This project aims to analyze the relationship between various car features and their impact on pricing (MSRP) and profitability in the automotive market. By exploring the data, the study provides actionable insights for car manufacturers to optimize pricing strategies and prioritize product development efforts.

Business Problem

The project seeks to answer:

- **Key Question:** How can car manufacturers optimize pricing and product development decisions to maximize profitability while meeting consumer demand?

Data Sources

The dataset titled "**Car Features and MSRP**" was sourced from Kaggle. It contains data on 11,159 car models with 16 variables, including:

- Make, Model, Year
- Engine specifications (e.g., Fuel Type, Horsepower)
- Transmission Type, Driven Wheels, Body Style
- Market Category, Popularity, MSRP.

Data Cleaning and Preprocessing

The following steps were performed:

1. **Handling Missing Values:**
 - Numerical columns like Engine HP were filled with the median values.
 - Categorical columns like Market Category were filled with the mode.
2. **Outlier Detection:**
 - Outliers in MSRP were identified using interquartile ranges and removed.
3. **Dummy Variable Creation:**
 - Categorical variables (e.g., Vehicle Size, Transmission Type) were converted to dummy variables for regression analysis.

4. Formatting:

- Ensured consistent numerical formatting and removed duplicates.

Assumptions

- The data reflects historical trends and may not represent the current market.
- Missing values are assumed to be random and not indicative of underlying patterns.

2. Approach

Analytical Methods

1. Descriptive Statistics:

- Summary statistics (mean, median, standard deviation) were used to understand the dataset.

2. Visualization:

- Scatter plots, line charts, and bar charts were created to identify relationships between features and MSRP.

3. Regression Analysis:

- Performed multiple linear regression to identify the strongest predictors of MSRP.

4. Interactive Dashboard:

- Developed an Excel dashboard with slicers and charts to visualize key insights dynamically.

Reasoning Behind Methods

- Regression analysis was chosen to quantify the relationship between features and MSRP.
- Visualization methods were employed for intuitive interpretation of trends and patterns.

Modelling Techniques

- Multiple Linear Regression:
 - Dependent variable: MSRP.
 - Independent variables: Engine HP, Highway MPG, and dummy variables for categorical features.

Challenges and Limitations

- **Challenge:** Creating dummy variables for categorical data in a dataset with over 11,000 rows.
- **Limitation:** Exclusion of certain features like cost data limited the ability to directly compute profitability.

3. Tech-Stack Used

Tools and Software

- **Microsoft Excel 365:**
 - Utilized for all aspects of the project, including data cleaning, analysis, and visualization.
 - Features such as Pivot Tables, Power Query, and advanced charting tools were extensively used to process and present the data.

Reasoning for Choice

- Excel 365 provides powerful tools for data manipulation and visualization within a user-friendly interface, making it ideal for this project.
- Its cloud integration allows for seamless collaboration and accessibility across devices.

Additional Features Utilized

- **Analysis ToolPak:**
 - Enabled regression analysis to identify key predictors of MSRP.
- **Interactive Dashboards:**
 - Leveraged slicers and charts for dynamic data filtering and visualization.

4. Insight

Key Findings

1. **Horsepower Impact:**
 - Engine Horsepower (HP) is a significant positive predictor of MSRP. For every additional unit of horsepower, MSRP increases by approximately 309.94 units on average, all else being equal.
 - This suggests that more powerful engines are highly valued by customers and should be a focus for premium models.

2. Engine Cylinders and Fuel Efficiency:

- Cars with more Engine Cylinders (e.g., V6 or V8 engines) tend to have a higher MSRP, with an average increase of 10,928.94 units per additional cylinder.
- Fuel efficiency, measured as City MPG and Highway MPG, also positively impacts MSRP. For every additional unit of City MPG, MSRP increases by 916.39 units, while an additional unit of Highway MPG increases MSRP by 235.14 units. These findings highlight the dual appeal of performance and efficiency.

3. Transmission Type Influence:

- Automated Manual transmission types are associated with an average MSRP increase of 25,067.52 units compared to other transmission types.
- Conversely, Direct Drive transmissions negatively affect MSRP, reducing it by approximately 38,161.86 units on average.

4. Body Style Variations:

- Compact and midsize body styles show significant positive effects on MSRP, increasing it by 25,544.28 units and 8,849.57 units, respectively. These body styles might represent vehicles positioned as premium models in the market.

5. Drive Type Preferences:

- All-Wheel Drive systems contribute significantly to a higher MSRP, with an average increase of 19,414.30 units, highlighting their appeal in premium and utility-focused vehicles.
- Rear-Wheel Drive also adds a notable premium of 5,518.32 units to MSRP.

6. Brand Differentiation:

- The regression model's R-Square value of 0.508 indicates that approximately 50.8% of the variance in MSRP is explained by the predictors, showing a reasonably strong relationship. However, other unexamined factors such as brand prestige or additional features may further influence MSRP.

Recommendations

1. Product Development:

- Focus on high-performance vehicles with enhanced Engine HP to attract premium buyers.
- Promote automatic transmission vehicles to increase profitability.

2. Market Strategy:

- Price SUVs competitively to capitalize on their growing popularity.
- Use fuel-efficient models as entry-level offerings to attract cost-conscious buyers.

3. Brand Strategy:

- Emphasize luxury and performance features for premium brands to maintain high price points.
- Highlight brand reliability for economy-focused manufacturers like Honda and Toyota.

5. Result

Model Fit and Predictive Accuracy:

- The regression model achieved an R-Square of 0.508, explaining 50.8% of the variability in MSRP. While the model captures major trends, the unexplained variance suggests room for further refinement by including additional predictors like brand reputation, safety features, or technology add-ons.

Significant Predictors:

- The analysis identified several key factors influencing MSRP, including Engine HP, Engine Cylinders, Transmission Types, Body Styles, and Drive Types. Predictors with high statistical significance ($p\text{-value} < 0.05$) such as Engine HP, City MPG, and Compact Body Style reinforce their critical impact on vehicle pricing.

Actionable Recommendations:

- **Focus on Performance and Efficiency:** Manufacturers should continue investing in vehicles with higher horsepower and better fuel efficiency to cater to customer demand for powerful yet economical options.
- **Expand Premium Features:** The substantial influence of Automated Manual Transmissions and All-Wheel Drive systems on pricing underscores the need to position these features in high-end models.
- **Segment-Specific Strategies:** Compact and midsize vehicles should be marketed aggressively as they command a premium, offering an opportunity to maximize profits in these categories.

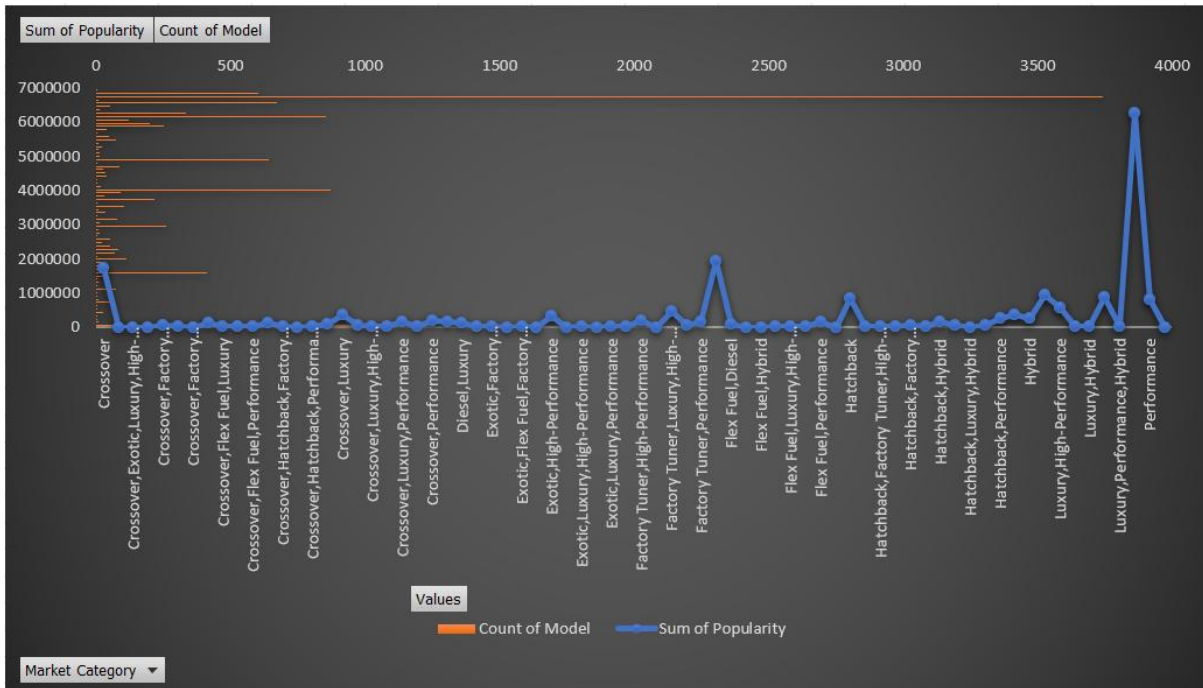
Challenges and Future Scope:

- The model does not account for qualitative factors like customer preferences, market trends, or brand perception, which may contribute to the remaining 49.2% variance. Future analyses could incorporate customer reviews or survey data for deeper insights.

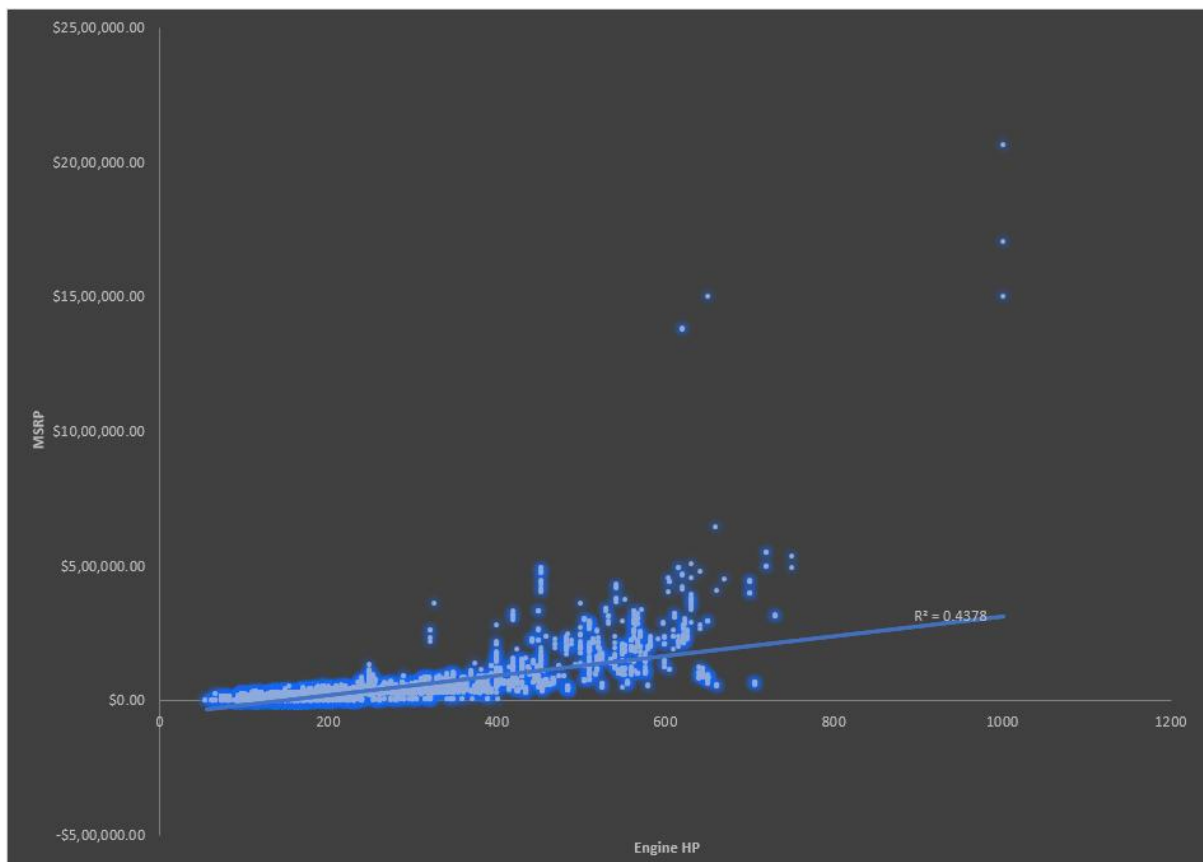
- **Task 1.A:** Create a pivot table that shows the number of car models in each market category and their corresponding popularity scores.

Row Labels	Sum of Popularity	Count of Model
Crossover	1715242	1110
Crossover,Diesel	6111	7
Crossover,Exotic,Luxury,High-Performance	238	1
Crossover,Exotic,Luxury,Performance	238	1
Crossover,Factory Tuner,Luxury,High-Performance	47410	26
Crossover,Factory Tuner,Luxury,Performance	13037	5
Crossover,Factory Tuner,Performance	840	4
Crossover,Flex Fuel	132720	64
Crossover,Flex Fuel,Luxury	11732	10
Crossover,Flex Fuel,Luxury,Performance	9744	6
Crossover,Flex Fuel,Performance	33942	6
Crossover,Hatchback	120650	72
Crossover,Hatchback,Factory Tuner,Performance	12054	6
Crossover,Hatchback,Luxury	1428	7
Crossover,Hatchback,Performance	12054	6
Crossover,Hybrid	107662	42
Crossover,Luxury	362665	410
Crossover,Luxury,Diesel	73080	34
Crossover,Luxury,High-Performance	9335	9
Crossover,Luxury,Hybrid	15142	24
Crossover,Luxury,Performance	151968	113
Crossover,Luxury,Performance,Hybrid	7832	2
Crossover,Performance	178431	69
Diesel	145396	84
Diesel,Luxury	116025	51
Exotic,Factory Tuner,High-Performance	21974	21
Exotic,Factory Tuner,Luxury,High-Performance	26912	52

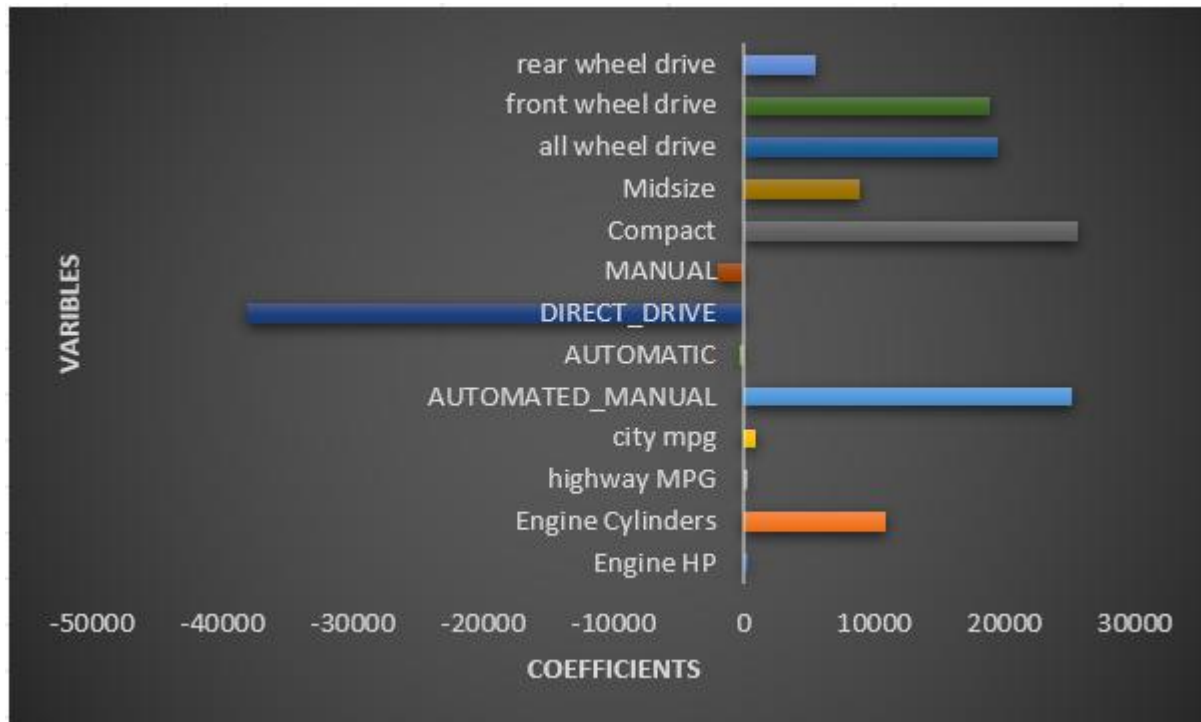
- **Task 1.B:** Create a combo chart that visualizes the relationship between market category and popularity.



- **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.



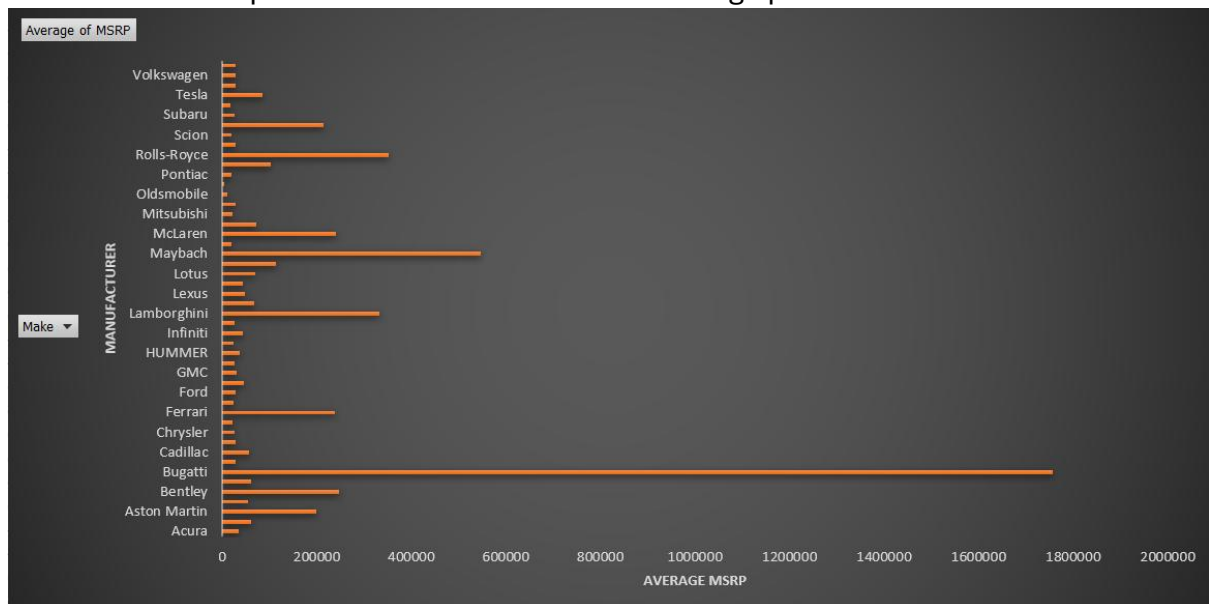
Task 3: Use regression analysis to identify the variables that have the strongest relationship with a car's price. Then create a bar chart that shows the coefficient values for each variable to visualize their relative importance.



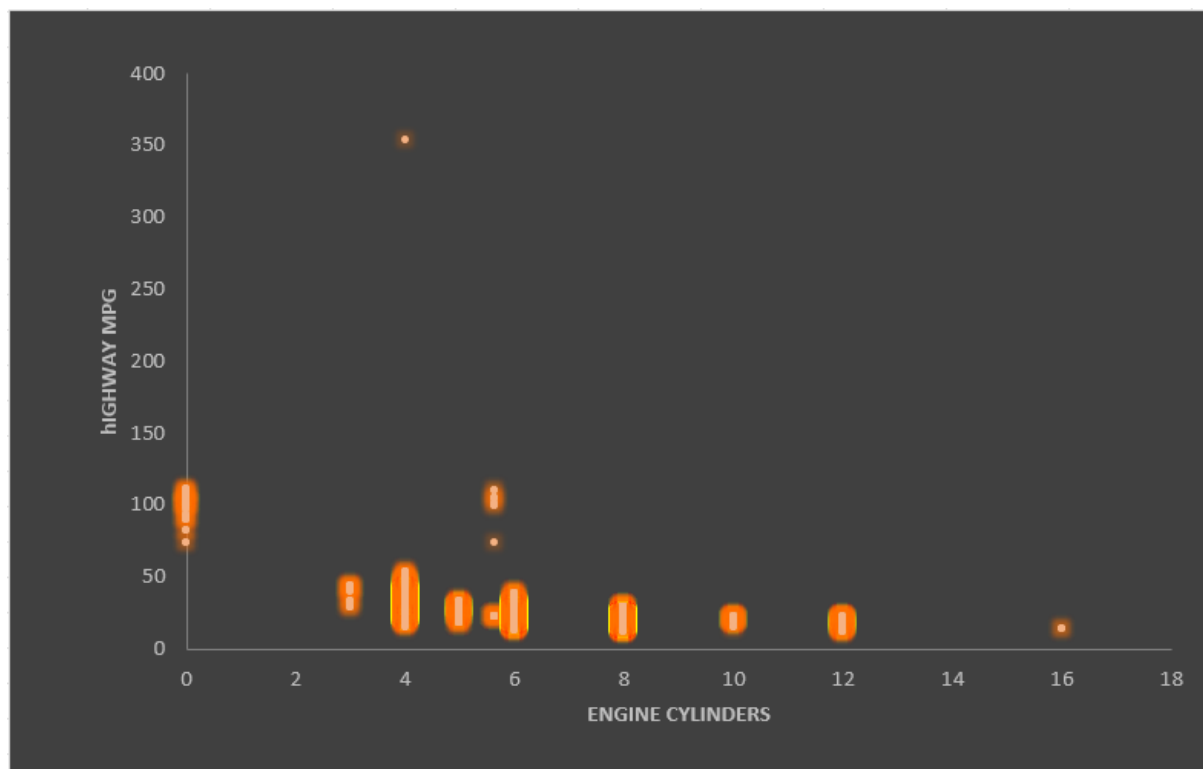
Task 4.A: Create a pivot table that shows the average price of cars for each manufacturer.

Manufacturer	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
Infiniti	42394.21212
Kia	25310.17316
Lamborghini	331567.3077
Land Rover	67823.21678
Lexus	47549.06931
Lincoln	42839.82927
Lotus	69188.27586
Maserati	114207.7069
Maybach	546221.875
Mazda	20039.38298
McLaren	239805
Mercedes-Benz	71476.22946
Mitsubishi	21240.53521
Nissan	28583.4319
Oldsmobile	11542.54
Plymouth	3122.902439
Pontiac	19321.54839
Porsche	101622.3971
Rolls-Royce	351130.6452
Saab	27413.5045
Scion	19932.5
Spyker	213323.3333
Subaru	24827.50391
Suzuki	17907.20798
Tesla	85255.55556
Toyota	29030.01609
Volkswagen	28102.38072
Volvo	28541.16014

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



- **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= -0.60229

Conclusion

This project successfully identified key factors influencing car prices and provided actionable insights for manufacturers. The interactive dashboard offers an efficient way to explore data and make informed decisions.

FOR DASHBOARD TASKS PLEASE REFER TO THE EXCEL SHEET: [CLICK HERE](#)