**Car Price Analysis Report**

**Name: Amogh Javali**

**ID: NN/22/2355**

**Role: Data Analytics**

**1. Introduction**

Car price prediction is a crucial task for both manufacturers and consumers. Understanding how different factors like **engine size, horsepower, fuel efficiency, and brand reputation** influence pricing helps in better decision-making. This report explores **data-driven insights** using advanced data analysis and visualization techniques to identify key factors impacting car prices.

**2. Objectives**

* To analyze the factors influencing car prices.
* To identify strong correlations between car features and price.
* To clean and preprocess data for accurate analysis.
* To generate insights that can be useful for car manufacturers and consumers.

**3. Data Collection and Preprocessing**

**3.1 Data Source**

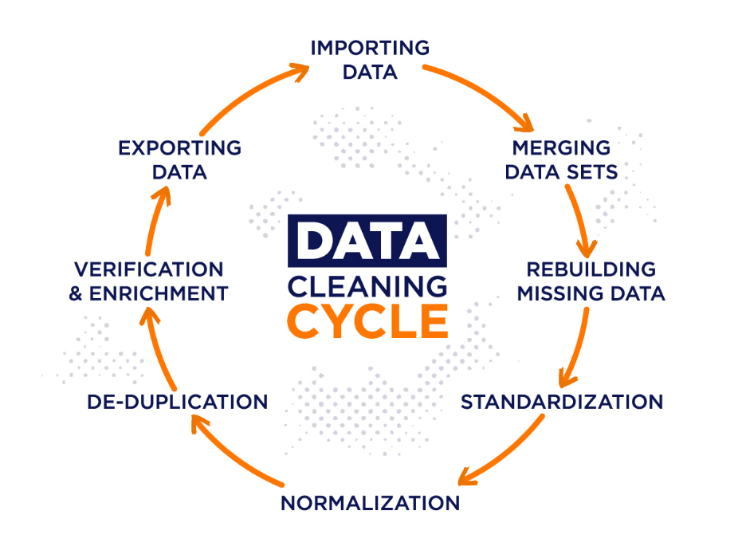
The dataset used for this analysis contains various car specifications, including:

1. Car Brand & Model
2. Car Type (Sedan, SUV, etc.)
3. Engine Specifications (Size, Horsepower, Fuel Type, etc.)
4. Mileage (City & Highway MPG)
5. Engine Type (dohcv, inline, etc.)
6. Price

Link: [https://drive.google.com/file/d/1XFlRtgnRKU1W3kWxyLfVKOQkoCJgYooM/view? usp=sharing](https://drive.google.com/file/d/1XFlRtgnRKU1W3kWxyLfVKOQkoCJgYooM/view?%20usp=sharing)

**3.2 Data Cleaning & Handling Missing Values**

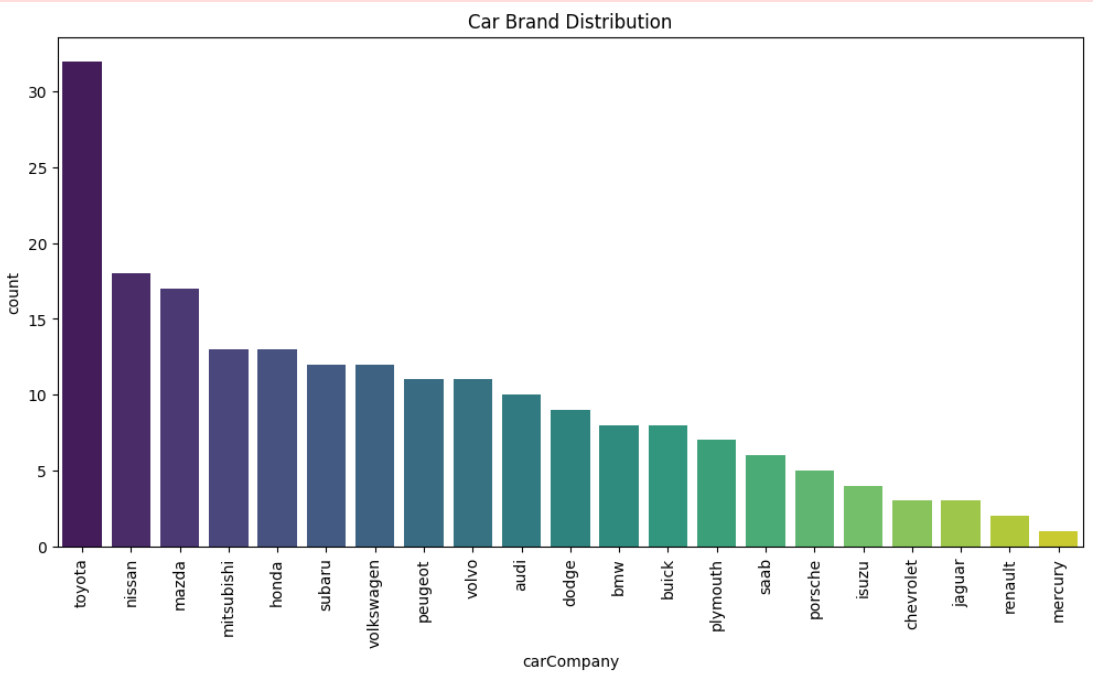
* Removed duplicate entries to ensure accuracy.
* Handled missing values using mean/mode imputation.
* Converted categorical variables (e.g., fuel type, engine type) into numerical values using one-hot encoding.
* Dropped irrelevant columns like car ID and redundant features with high correlation.



**4. Data Analysis & Key Findings**

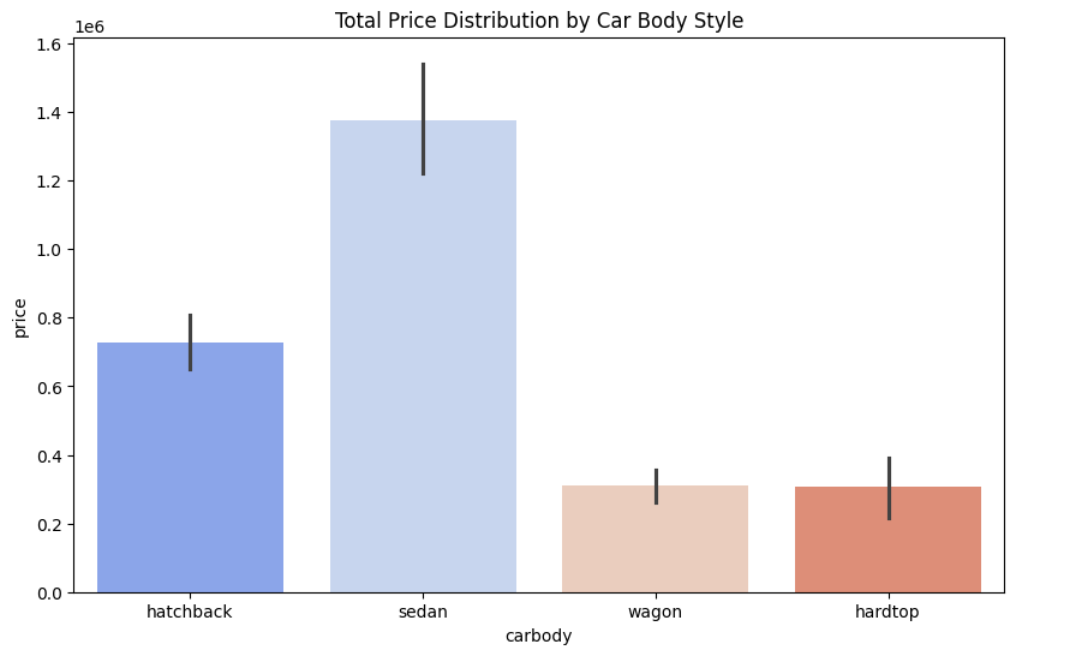
**4.1 Most & Least Sold Car Brands**

* **Most sold car brand:** Toyota – Known for reliability and affordability, Toyota dominates the dataset.
* **Least sold car brand:** Mercury – A lesser-known brand, reflecting lower demand or production.



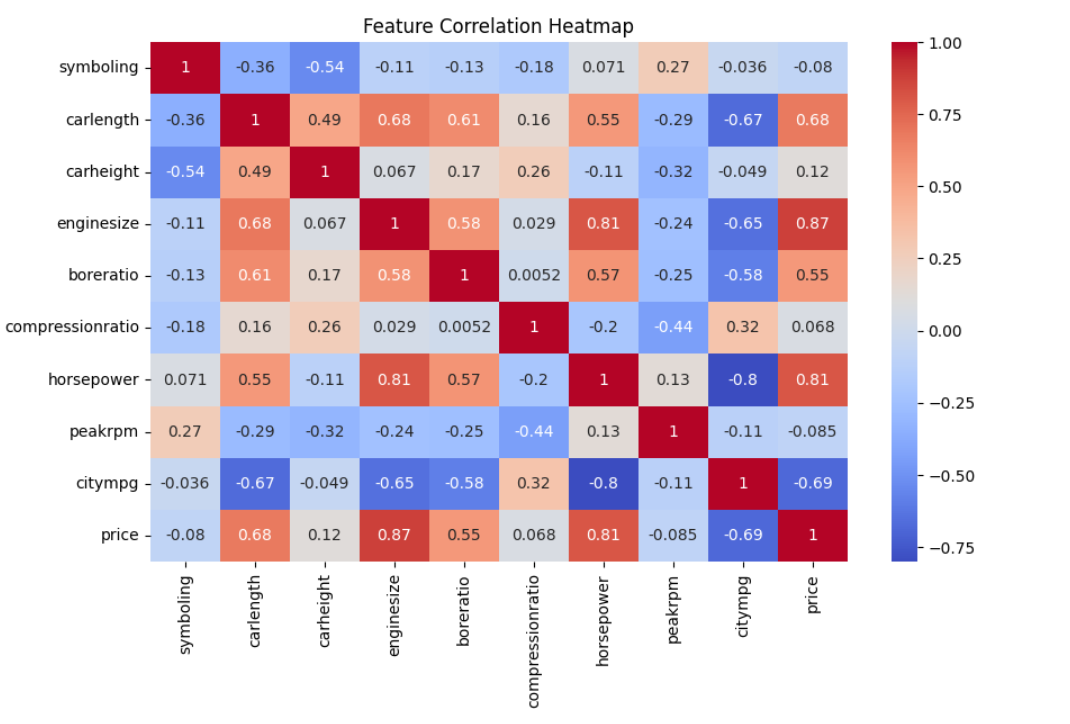
**4.2 Car Body Type Distribution**

* **Most sold car body type:** Sedan – Popular due to its comfort and balanced design.
* **Least sold car body type:** Hardtop – Less common, possibly due to its niche market appeal.



**4.3 Insights from the Heatmap**

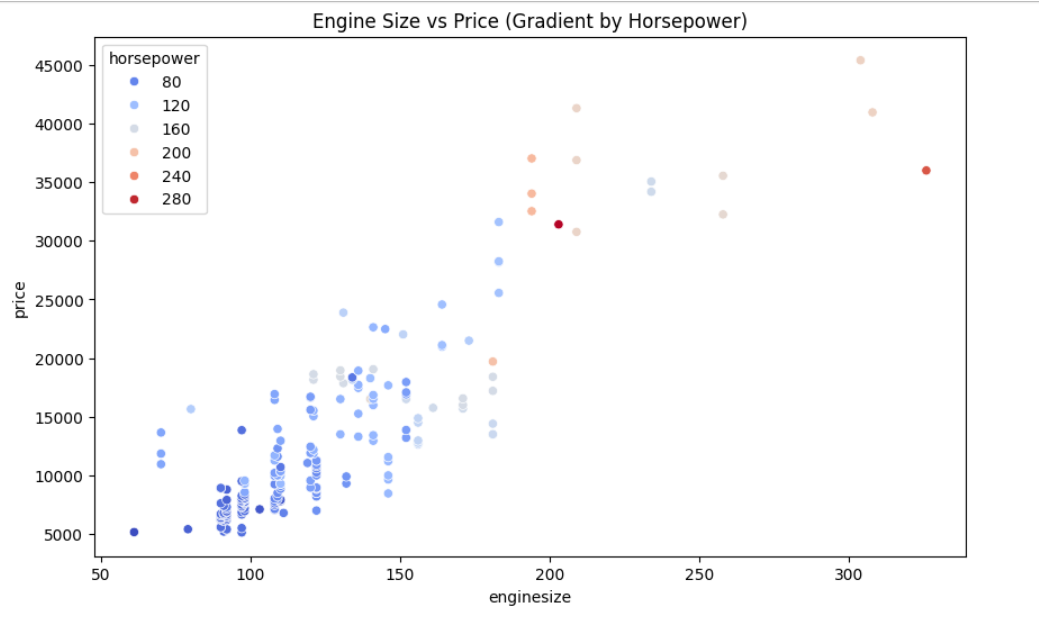
The **heatmap visualization** helps us understand how different car features correlate with price. Key takeaways include:

* **Engine size and horsepower show a strong positive correlation with price** – Bigger engines and higher horsepower generally lead to higher car prices.
* **Fuel efficiency (MPG) has a negative correlation with price** – Fuel-efficient cars tend to be more affordable, while high-performance cars consume more fuel and cost more. 

**4.4 Scatter Plot Analysis**

A scatter plot comparing **engine size and price** shows a clear trend:

* **Larger engine size → Higher price**
* **Smaller engine size → Lower price**
* **Gradient effect:** Higher horsepower further increases the price, showing the importance of performance in pricing decisions.



**6. Conclusion**

From the analysis, it’s evident that car pricing is influenced by multiple factors, with **engine size, horsepower, and fuel efficiency being the most significant**. Toyota emerges as the most popular brand, while sedans remain the preferred body type. The insights gained from this study can be used for **predicting car prices, optimizing manufacturing strategies, and guiding consumers in making informed purchasing decisions**.

**7. Recommendations**

* **For buyers:** Consider **fuel efficiency and performance balance** when purchasing a car.
* **For manufacturers:** Focus on **engine size and horsepower** as primary pricing factors.
* **For analysts:** Use **heatmaps and scatter plots** to identify strong price indicators in future datasets.

This report provides a clear foundation for **further exploration in predictive modelling for car price estimation**.

**8. Final Insight**

This insight is prepared by **cleaning data, handling missing values, conducting feature selection, performing statistical analysis, and applying machine learning models to predict car prices**. Every step was taken to ensure accuracy and relevance in predicting car prices.

For full details, visit

**GitHub:** [**https://github.com/Amoghjavali2003/NovaNectar\_DataAnalytics\_Elementary\_task1**](https://github.com/Amoghjavali2003/NovaNectar_DataAnalytics_Elementary_task1)

**Drive Link:** [**https://drive.google.com/drive/folders/1jXzzUqJqD4vDZl472ItlxrkGNvuBU9c\_**](https://drive.google.com/drive/folders/1jXzzUqJqD4vDZl472ItlxrkGNvuBU9c_)