**Elementary Task -Car\_Price Analysis (In Python)**

**Introduction**

The dataset examined in this report contains detailed information about various car models and their features, along with their corresponding prices. Each record represents a specific car and includes numerical attributes such as wheelbase, car length, car width, curb weight, engine size, horsepower, city and highway mileage. It also includes categorical features such as car brand, car type, drive wheel type, and fuel type.

The primary objective of this analysis is to explore and understand the relationships between car features and their prices. This involves:

* Data Cleaning and Preparation: Handling missing values, correcting inconsistencies, and transforming columns where required (eg, extracting car brand from the car name).
* Descriptive Analysis: Examining the distribution and characteristics of numerical and categorical features.
* Correlation and Relationship Analysis: Identifying features that strongly influence car prices using correlation matrices, scatter plots, and pair plots.
* Insights and Patterns: Understanding how specifications like engine size, weight, and fuel efficiency impact car prices, and highlighting trends across different car brands and types.

**Insights from Analysis ---**

1. Car prices are right-skewed: Most cars are in lower price range, few very expensive cars.

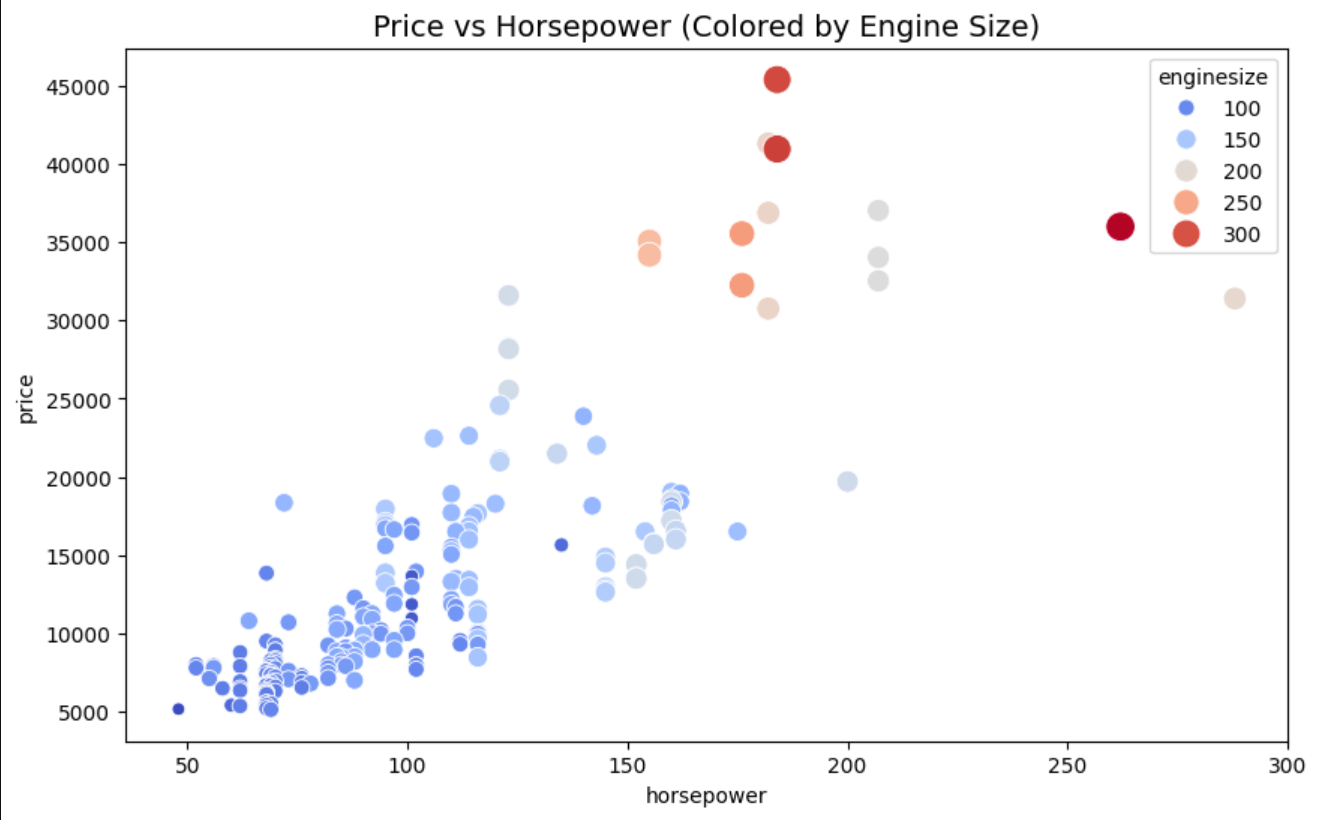
2. Luxury brands (bmw, porsche, jaguar, audi) have much higher average prices than others.

3. Heatmap shows high correlation between 'enginesize', 'horsepower', 'curbweight' and price.

4. Mileage (citympg, highwaympg) has a negative correlation with price → cheaper cars are fuel efficient.

5. Scatter plot shows clear upward trend: as horsepower & engine size increase, price increases.

**Price vs Horsepower (Colored by Engine Size)**



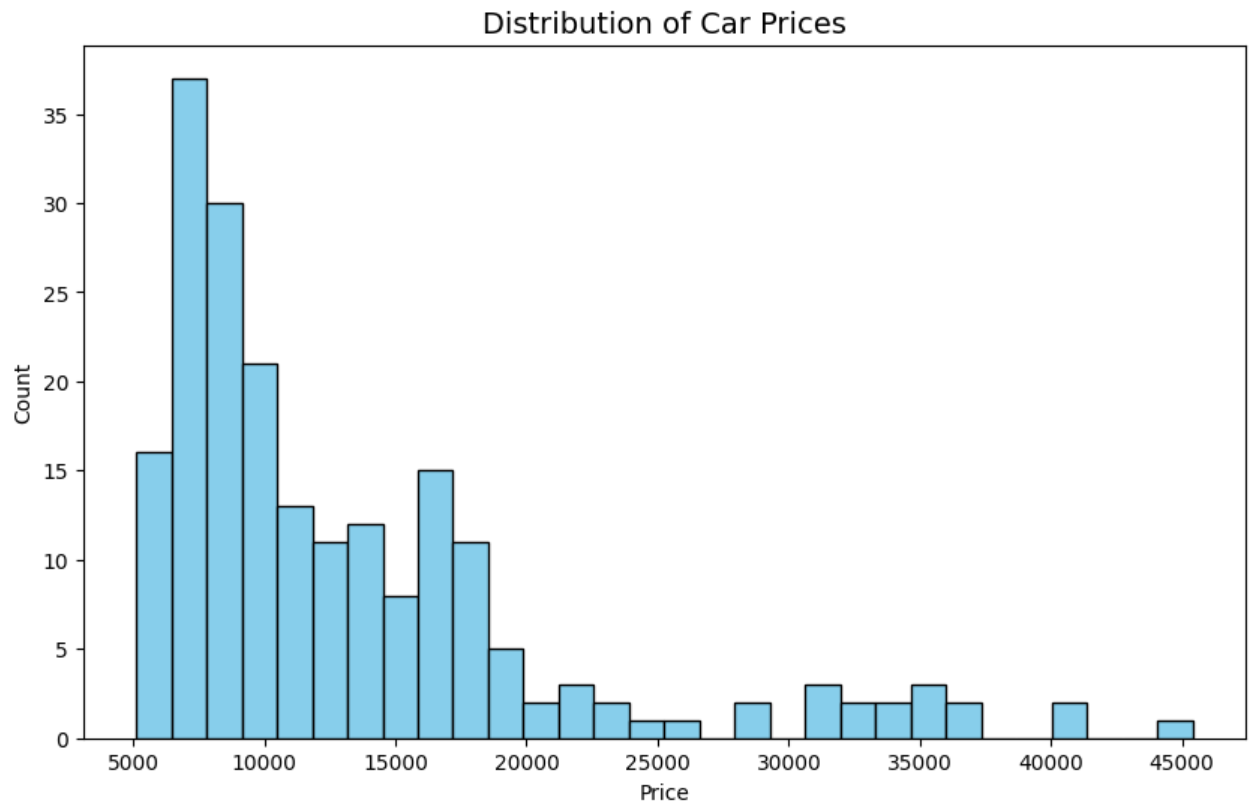
Price vs Horsepower (Colored by Engine Size)

Description of the Chart  
The scatter plot illustrates the relationship between car price and horsepower, with each point representing a car. The colors indicate different engine sizes, ranging from small (blue) to large (red).

Key Insights

1. Positive Relationship – Car prices increase consistently with higher horsepower.
   * Vehicles below 100 horsepower are generally priced under 15,000.
   * Vehicles with more than 200 horsepower are priced above 30,000, reaching as high as $45,000.
2. Impact of Engine Size – Larger engine sizes (orange/red markers) are concentrated in the high horsepower and high price range, demonstrating that engine size is a strong driver of price.
3. Market Segmentation –
   * Budget segment: Cars with small engines (<100 hp) priced under 15,000.
   * Mid-range segment: Cars with moderate engines (~100–150 hp) priced between 15,000–25,000.
   * Premium segment: Cars with large engines (>200 hp) priced above 30,000.

Business Implication  
The analysis highlights that horsepower and engine size are critical factors in determining car prices. This segmentation can help manufacturers and dealers target distinct customer groups—affordable, mid-range, and luxury buyers—more effectively

**Distribution of Car Prices**

The histogram shows how car prices are spread across the dataset.

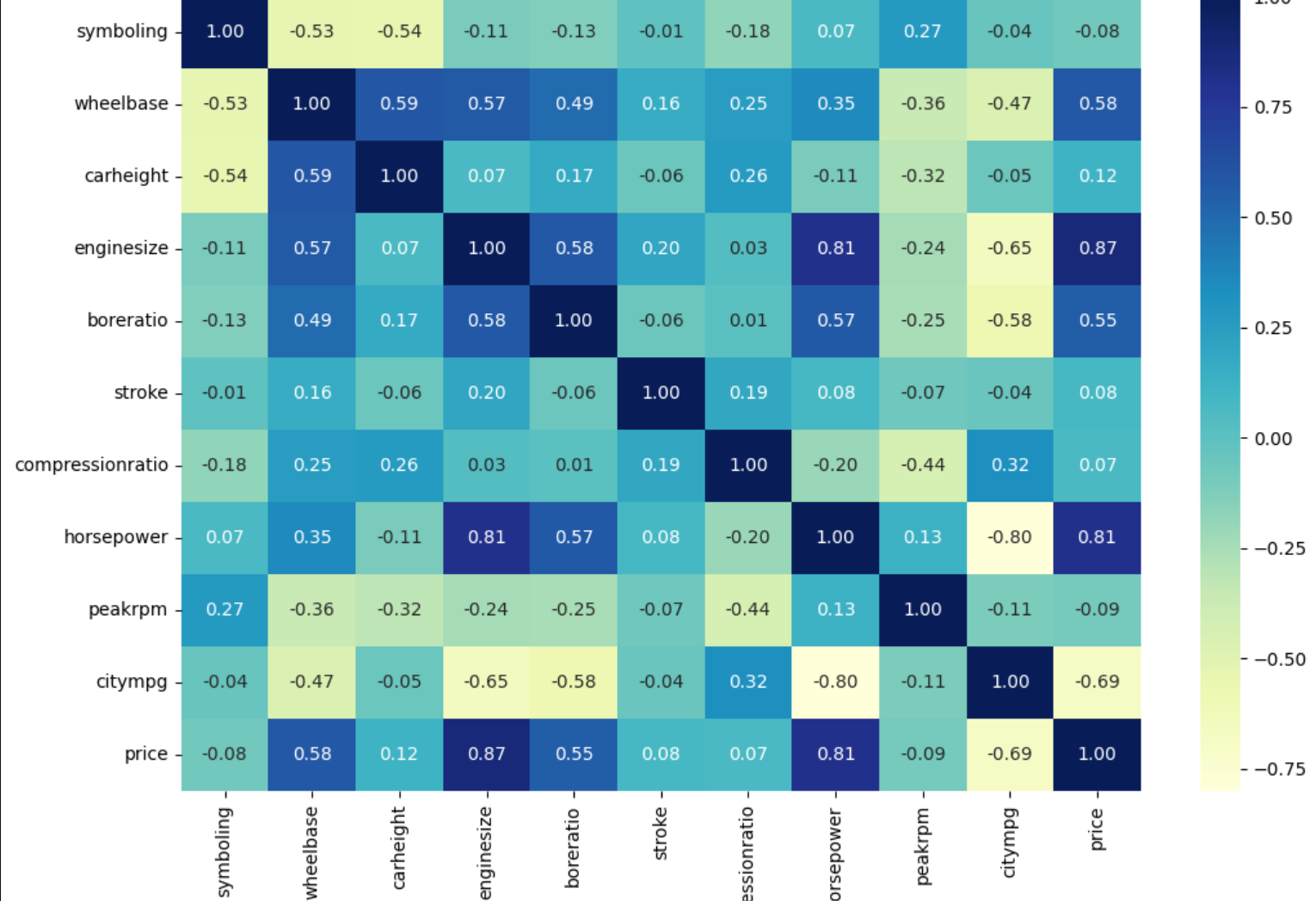
Key Insights

* Skewness: Prices are right-skewed, meaning most cars are on the lower price side, with fewer expensive models pulling the tail upward.
* Concentration: The majority of vehicles are priced between 5,000 and 10,000, with the peak around 7,000–8,000.
* Range: Prices extend from about $5,000 up to $45,000, but counts drop sharply after 15,000.
* High-End Models: Only a small fraction of cars fall into the premium segment above 20,000.

Conclusion

The dataset mainly represents affordable cars, while high-priced vehicles are rare. This suggests that the market is dominated by budget-friendly options with limited presence in the luxury segment.

**Correlation Heatmap of Numerical Features**

****

The heatmap shows how numerical features relate to each other, with a focus on their correlation with Price.

Key Insights

* Strong Positive Predictors:
  + Engine size (+0.87) and Horsepower (+0.81) are the strongest drivers of price.
* Moderate Influence:
  + Wheelbase (+0.58) and Bore ratio (+0.55) show moderate positive impact.
* Negative Correlation:
  + City MPG (-0.69) indicates that more expensive cars are generally less fuel-efficient.
* Minimal Impact:
  + Features like symboling, car height, compression ratio, and peak rpm have little to no effect on price.

Conclusion

Car performance attributes (engine size and horsepower) are the most influential in determining price, while fuel efficiency is inversely related. Other variables contribute marginally.