Exploratory Data Analysis on the Automobile Dataset

Introduction

This report aims to analyse the automobile dataset to gain insight into the type of vehicles, their manufacturers, and how price, horsepower, engine size and fuel-efficiency (measured in miles per gallon / MPG) affect each other.

By examining how these variables interact and influence each other, the report seeks to provide an analysis of what drives the cost of a vehicle.

Data cleaning

- The columns ‘normalized-losses’ and ‘symboling’ were removed as they were not relevant to the analysis.
- Duplicate records were deleted to ensure the accuracy and consistency of the data.
- Numeric values were converted from strings to integers to enable proper analysis.
- A new variable representing average MPG was created by combining ‘city MPG’ and ‘highway MPG’ for improved evaluation of fuel efficiency.

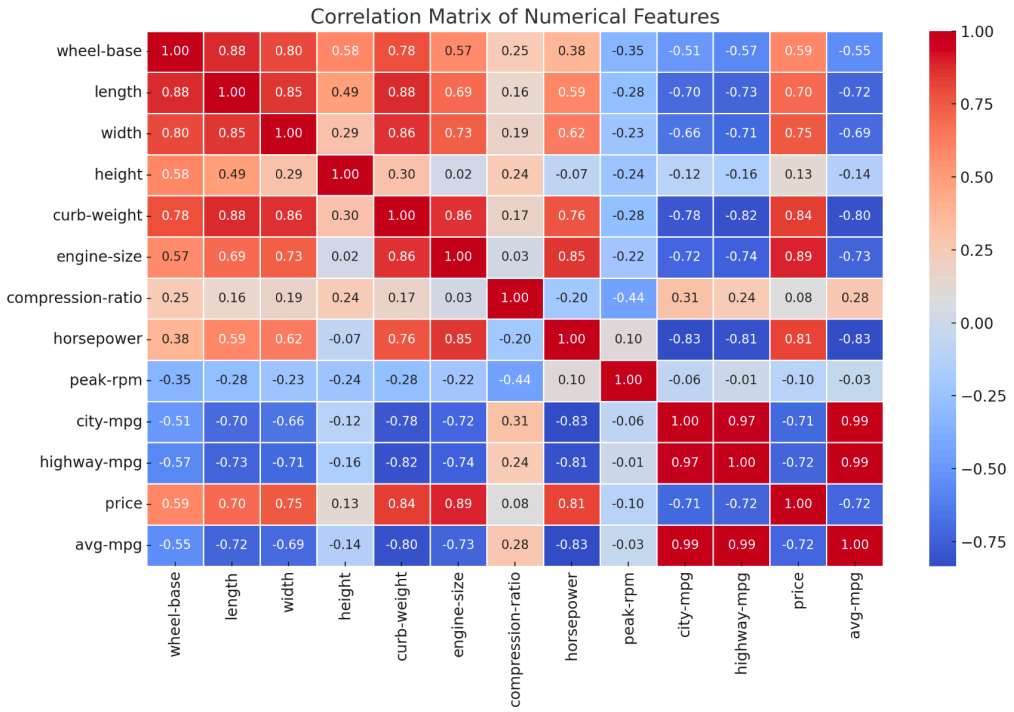
Missing data

- Placeholder symbols (‘?’) were replaced with proper missing value indicators (‘NaN’).
- Rows containing missing values were removed to maintain data quality.

Data stories and visualisations Report

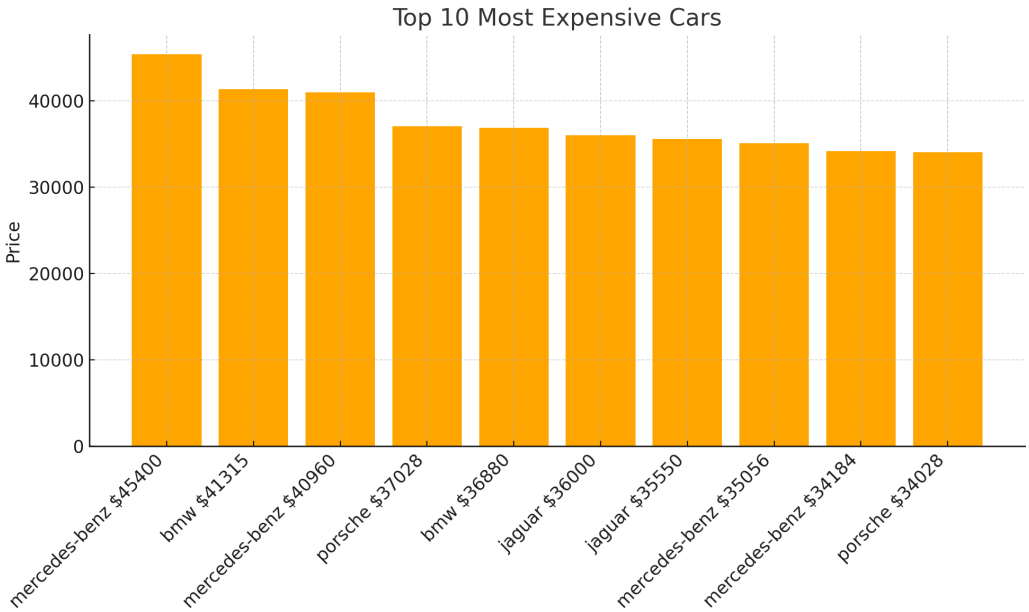
Correlation between the key metrics:

- Engine size, curb weight, and price show a strong positive correlation. Larger cars tend to cost more.
- Fuel efficiency is negatively correlated with engine size, weight, and price. Cars with bigger engines generally consume more fuel.
- City MPG and highway MPG are strongly positively correlated, indicating that cars tend to perform consistently across different driving environments



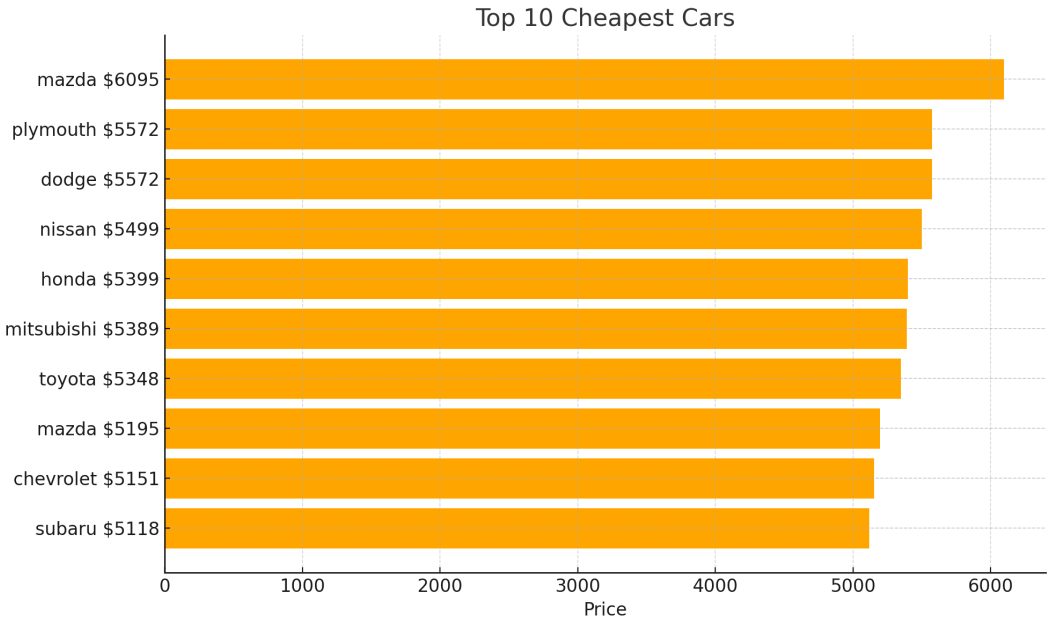
The Most Expensive Cars

- High-priced vehicles are typically equipped with large, powerful engines and offer higher horsepower.
- Expensive vehicles often include more advanced technology and luxury features, contributing to increased curb weight and cost.
- Performance-focused models, such as sports or luxury cars, prioritise speed and power over fuel economy.



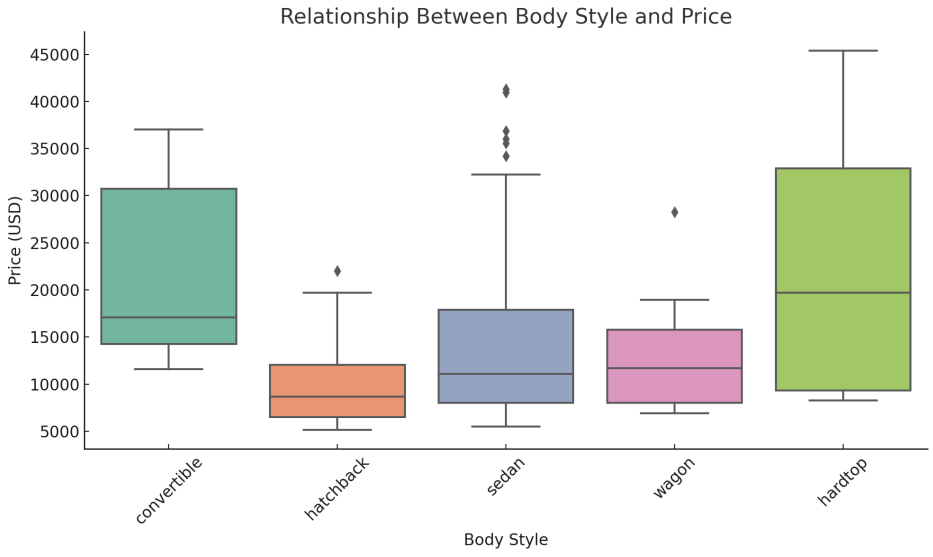
The Cheapest Cars

- Low-cost vehicles usually have small, less powerful engines and fewer features.
- Their lightweight design helps reduce manufacturing costs and improves fuel-efficiency.
- These cars usually have lower horsepower and performance, making cost-effectiveness a major selling point.



Body Style To Price

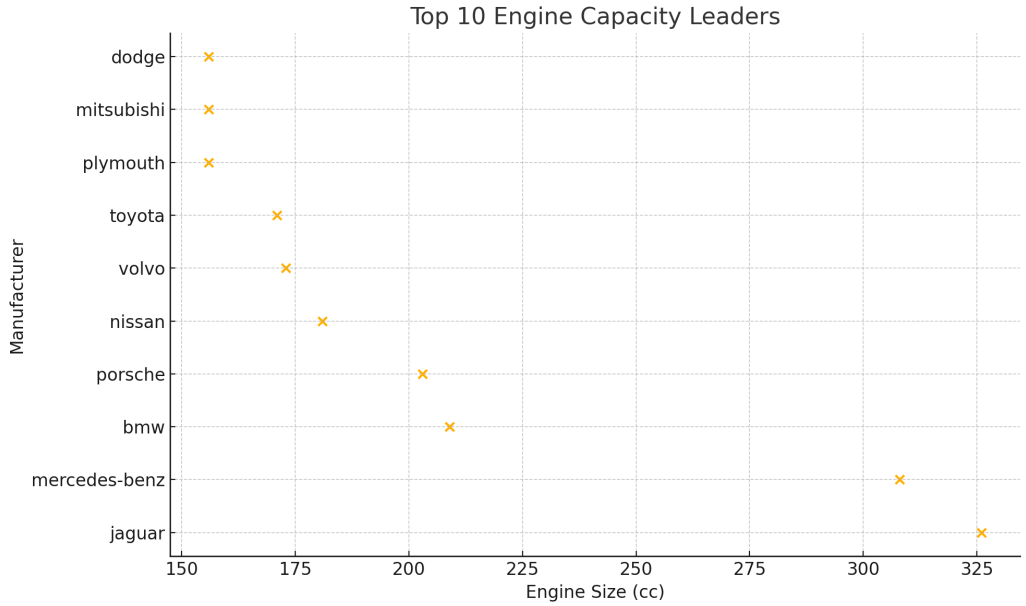
- Hatchbacks, when compared to hardtops and other body styles, tend to have simpler designs, features and smaller size.
- As a result, they are lighter, more fuel-efficient, and generally less expensive.

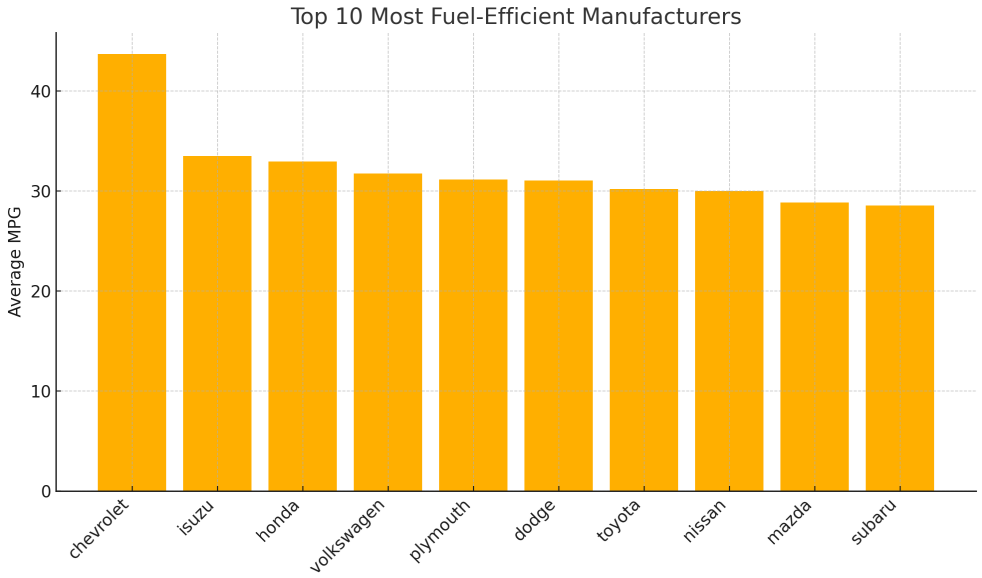


Top Engine Manufacturers In Contrast with Fuel-efficient Manufacturers

- Manufacturers with high engine capacity build cars with big, powerful engines. These cars are usually fast and heavy, but they use a lot of fuel. They are often more expensive because of their size and performance.

- In contrast, fuel-efficient brands prioritise smaller, lighter vehicles designed for affordability and fuel savings.





City MPG vs Highway MPG

- Chevrolet, Honda, and Volkswagen produce vehicles with high fuel-efficiency.
- Mercedes-Benz and Porsche focus more on making powerful and luxury cars. As a result, their cars use more fuel in comparison.
- The difference is marginal for the two areas across all the manufacturers.

Manufacturer	Avg. City MPG	Avg. Highway MPG
Chevrolet	41.0	46.3
Honda	31.5	35.5
Volkswagen	30.6	34.9
Mercedes-Benz	17.0	18.8
Porsche	17.4	25.0

Manufacturer with the most car models

- Toyota has the most car models (32 counts). They have more range across the different price points.

Manufacturer	Number of Models
Toyota	**32**
Nissan	18
Mazda	17
Mitsubishi	13
Honda	13

In Conclusion

- The most expensive cars have big engines, are heavy, and use a lot of fuel.
- The cheapest cars have small engines, are light, save fuel, and usually have front-wheel drive.
- Fuel-efficient brands mostly make small cars that don’t cost a lot.
- Cars with big engines cost more and use more fuel.

This report was written by : Jessica Ramathibe