

Project Summary

Automotive Data Analysis Project

By Abdellah Boudlal

Course: 44-413, 613 Data Visualization

Date: February 2026

SECTION 1 — Introduction

This project analyzes car sales performance and vehicle specifications using an Excel dataset. The goal of this analysis is to explore how vehicle characteristics such as price, horsepower, and fuel efficiency influence sales and market performance. Data visualization techniques were used to transform raw data into meaningful insights that can support business and marketing decisions in the automotive industry.

Data Source:

https://github.com/Sonya-7/Car_Sales_Dashboard/blob/main/Car_Sales.xlsx

SECTION 2 — Data Description

The dataset belongs to the automotive sales domain and contains detailed information about car manufacturers and models.

The analysis focuses on how different membership types perform across several months using Excel visualizations.

The data file includes sales performance, pricing, resale value, retention rates, engine specifications, and physical vehicle attributes.

Describe your data source:

The data source was obtained from GitHub and is publicly available for analytics and visualization practice.

Number of data columns:

19 columns

Number of data records:

155 records

Data Features Summary

Column Name	Description
Manufacturer	Car brand name
Model	Car model
Unit Sales	Units sold
Price (\$)	Sale price
Year Resale Value (\$)	Resale price
Retention %	Value retention
Engine Size	Engine capacity
Horsepower	Engine power
Vehicle Type	Car category
Fuel Efficiency (MPG)	Fuel economy
Curb Weight	Vehicle weight

SECTION 3 — Data Cleaning Strategies

Although the dataset was already structured, several cleaning and preparation steps were applied to improve clarity and focus on the project goals.

The following columns were removed because they were not relevant to the selected business objectives:

- Power Performance Factor
- Wheelbase
- Width
- Length
- Curb Weight
- Fuel Capacity
- Latest Launch

These attributes were more technical specifications and were not necessary for sales, pricing, fuel efficiency, or retention analysis.

Additionally:

- Currency fields (Price, Year Resale Value) were formatted properly.
- Retention percentages were standardized.
- Horsepower was grouped into categories (Low, Medium, High).
- Retention was classified as GOOD or POOR.

These steps ensured a focused and efficient dataset for pivot table analysis and visualization.

SECTION 4 — Clean Dataset

The cleaned dataset contains 155 rows and 12 attributes ready for visualization.

An excerpt of the dataset includes:

- Manufacturer: Acura, Audi, BMW...
- Price range: \$9,235 – \$85,500
- Sales range: 110 – 540,561 units
- Fuel efficiency: 15 – 45 MPG

This dataset allows analysis of relationships between vehicle performance and market success.

Manufacturer	Model	Unit Sales	Price	Year	Resal	Retention	Retention	Engine Size	Horsepower	HP Level	Vehicle_type	Fuel Efficiency
Acura	Integra	16919	21500	16360	0.76093	GOOD		1.8	140	Medium HP	Passenger	28
Acura	TL	39384	28400	19875	0.699824	GOOD		3.2	225	Medium HP	Passenger	25
Acura	RL	8588	42000	29725	0.707738	GOOD		3.5	210	Medium HP	Passenger	22
Audi	A4	20397	23990	22255	0.927678	GOOD		1.8	150	Medium HP	Passenger	27
Audi	A6	18780	33950	23555	0.693814	GOOD		2.8	200	Medium HP	Passenger	22
Audi	A8	1380	62000	39000	0.629032	POOR		4.2	310	High HP	Passenger	21
BMW	323i	19747	26990	24500	0.907744	GOOD		2.5	170	Medium HP	Passenger	26
BMW	328i	9231	33400	28675	0.858533	GOOD		2.8	193	Medium HP	Passenger	24
BMW	528i	17527	38900	36125	0.928663	GOOD		2.8	193	Medium HP	Passenger	25
Buick	Century	91561	21975	12475	0.567691	POOR		3.1	175	Medium HP	Passenger	25
Buick	Regal	39350	25300	13740	0.543083	POOR		3.8	240	Medium HP	Passenger	23
Buick	Park Avenue	27851	31965	20190	0.631628	POOR		3.8	205	Medium HP	Passenger	24

SECTION 5 — Visualization Tools

For this project, Microsoft Excel was selected as the primary data visualization and analysis tool.

Excel was chosen due to its strong capabilities in data aggregation, pivot table analysis, and chart visualization.

It provides an efficient environment to clean, transform, and analyze structured datasets such as automotive sales data.

Several Excel features were utilized throughout the project, including:

- Pivot Tables for data aggregation and grouping
- Calculated averages and maximum values
- Sorting and Top N filtering
- Classification grouping (Retention categories, HP levels)
- Data formatting (currency, percentages)
- Chart creation and customization

Multiple chart types were created to match the analytical goals, including:

- Bar charts
- Column charts
- Pie charts

Excel was particularly effective for comparing manufacturer sales performance, analyzing pricing structures, and evaluating resale value retention.

All visualizations and dashboards were created and organized within a single Excel workbook to ensure consistency and clarity.

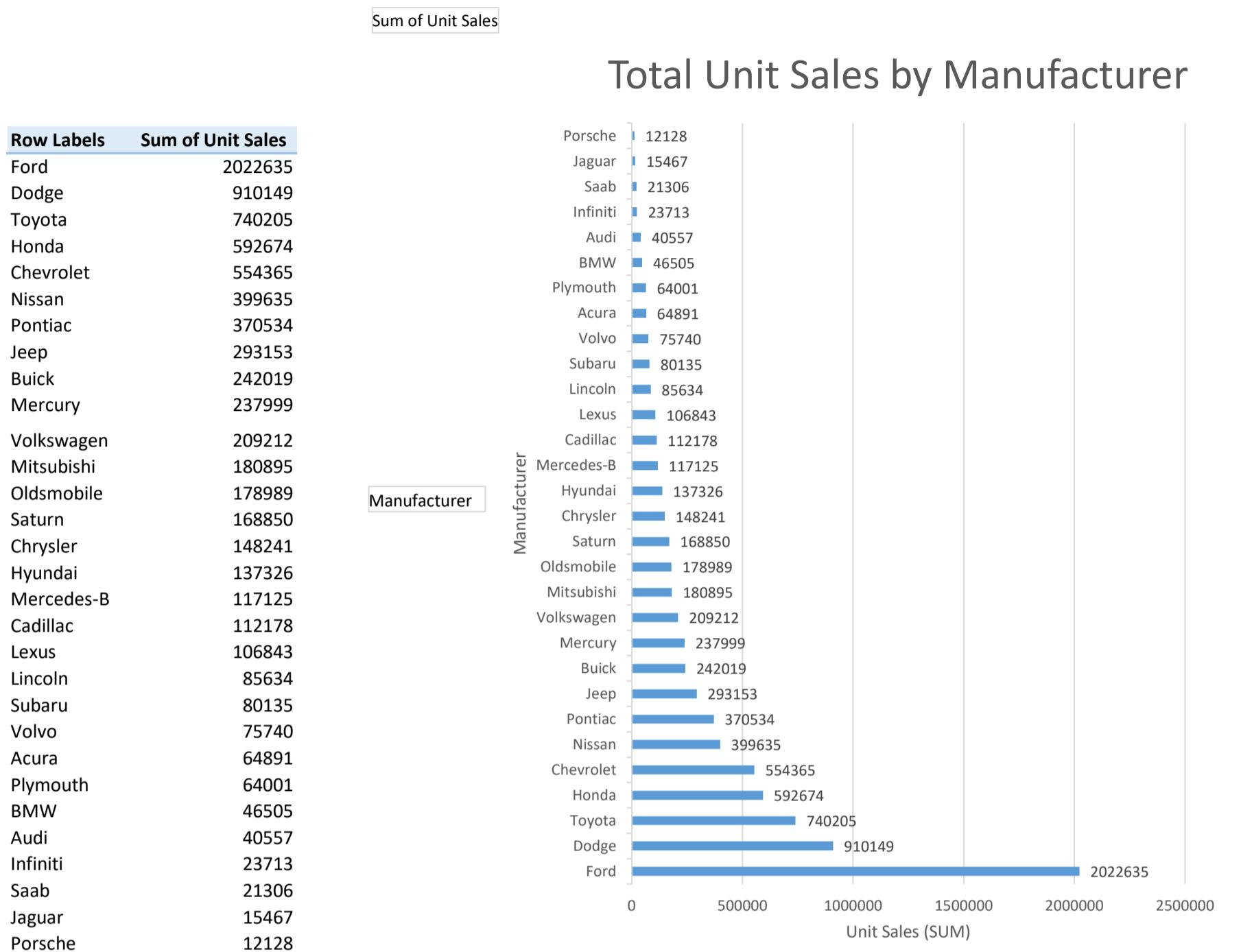
GitHub Data Source:

https://github.com/Sonya-7/Car_Sales_Dashboard/blob/main/Car_Sales.xlsx

SECTION 6 — Visualizations & Stories

Goal 1: Total Unit Sales by Manufacturer

Prepared by Abdellah Boudlal



Goal

Identify which manufacturer generates the highest total unit sales.

Chart

Bar Chart displaying total unit sales aggregated by manufacturer.

Story / Analysis

This visualization presents the total unit sales for each car manufacturer included in the dataset.

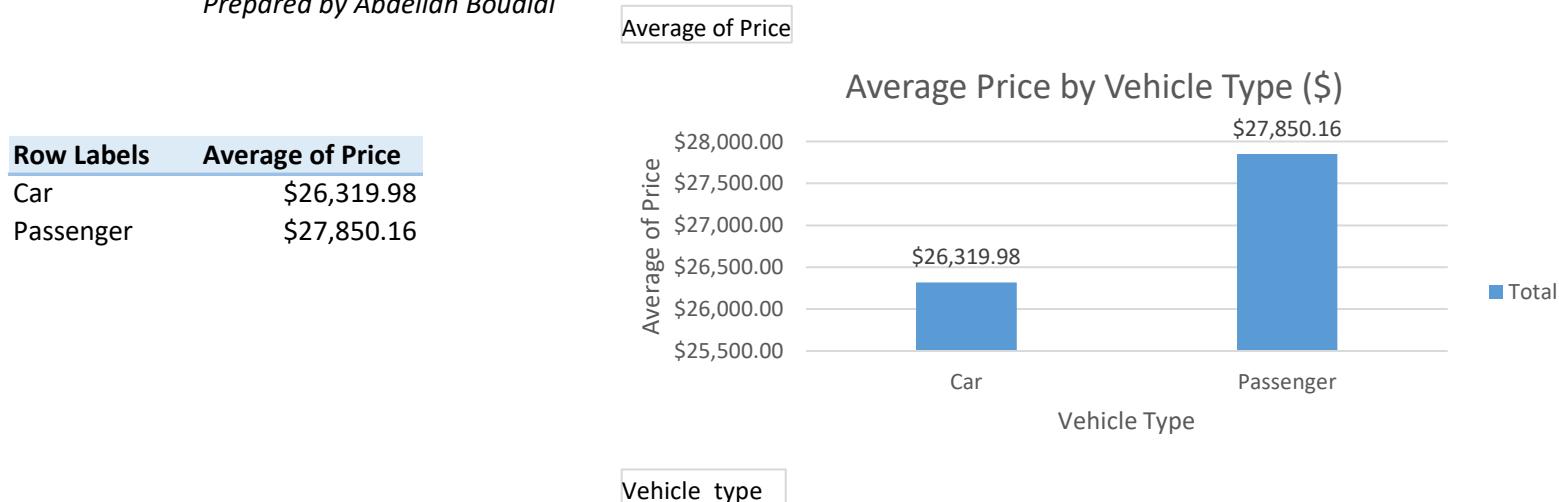
The analysis reveals that **Ford** has the highest total unit sales, significantly exceeding all other manufacturers. This indicates strong brand demand and a dominant market presence.

Dodge and **Toyota** also rank among the top manufacturers in sales volume, showing consistent performance and wide consumer reach.

On the other hand, luxury manufacturers such as **Porsche** and **Jaguar** record much lower unit sales. This difference is expected because luxury vehicles are sold in smaller quantities due to higher pricing and niche market targeting.

Goal 2: Average Price by Vehicle Type

Prepared by Abdellah Boudlal



Goal

Compare the average vehicle price between Passenger vehicles and Cars.

Chart

Column chart displaying the average price by vehicle type.

Story / Analysis

This visualization compares the average selling price between Passenger vehicles and Cars.

The analysis shows that **Passenger vehicles have a slightly higher average price (\$27,850)** compared to Cars (\$26,319).

This difference may be explained by the presence of premium sedans and luxury passenger models in the dataset, which increase the overall average price.

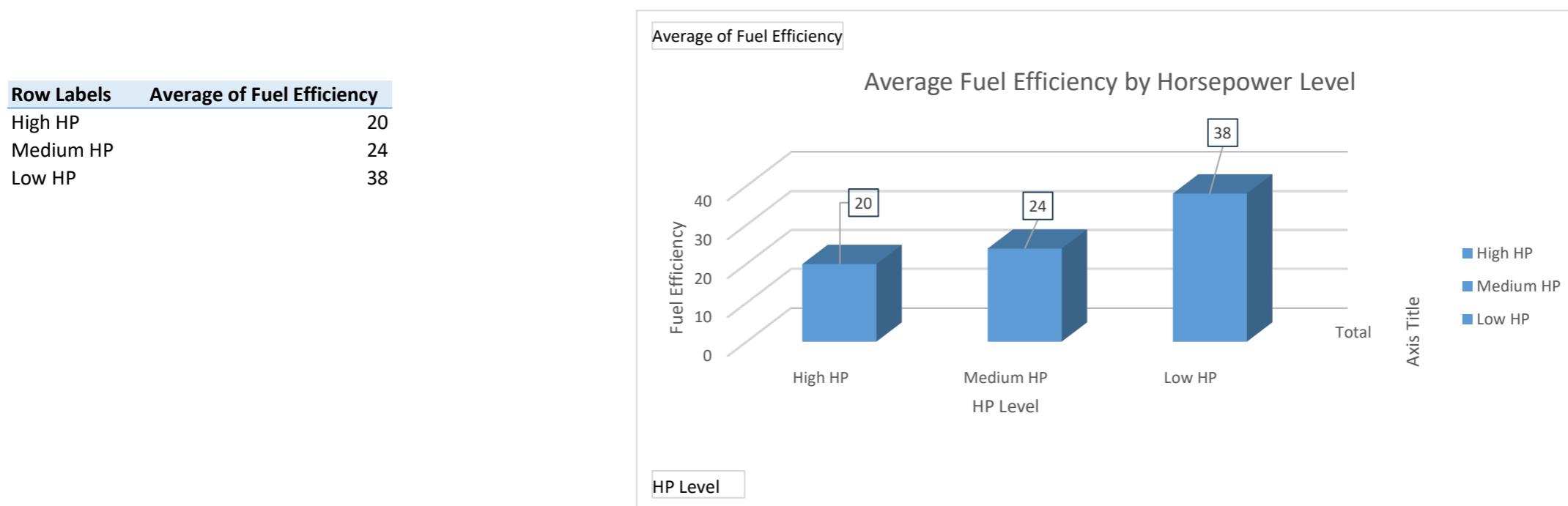
Cars such as SUVs and utility vehicles tend to serve broader markets, often balancing price with functionality and affordability.

Key Insight (1 phrase conclusion)

Vehicle type influences pricing structure, with passenger vehicles showing marginally higher average market value.

Goal 3: Fuel Efficiency by HP Level

Prepared by Abdellah Boudlal



Goal

Analyze the relationship between horsepower level and fuel efficiency.

Chart

Column chart displaying the average fuel efficiency (MPG) by horsepower level.

Story / Analysis

This visualization examines how fuel efficiency varies across different horsepower levels.

The analysis shows that:

Low HP vehicles have the highest fuel efficiency (38 MPG).

Medium HP vehicles average around 24 MPG.

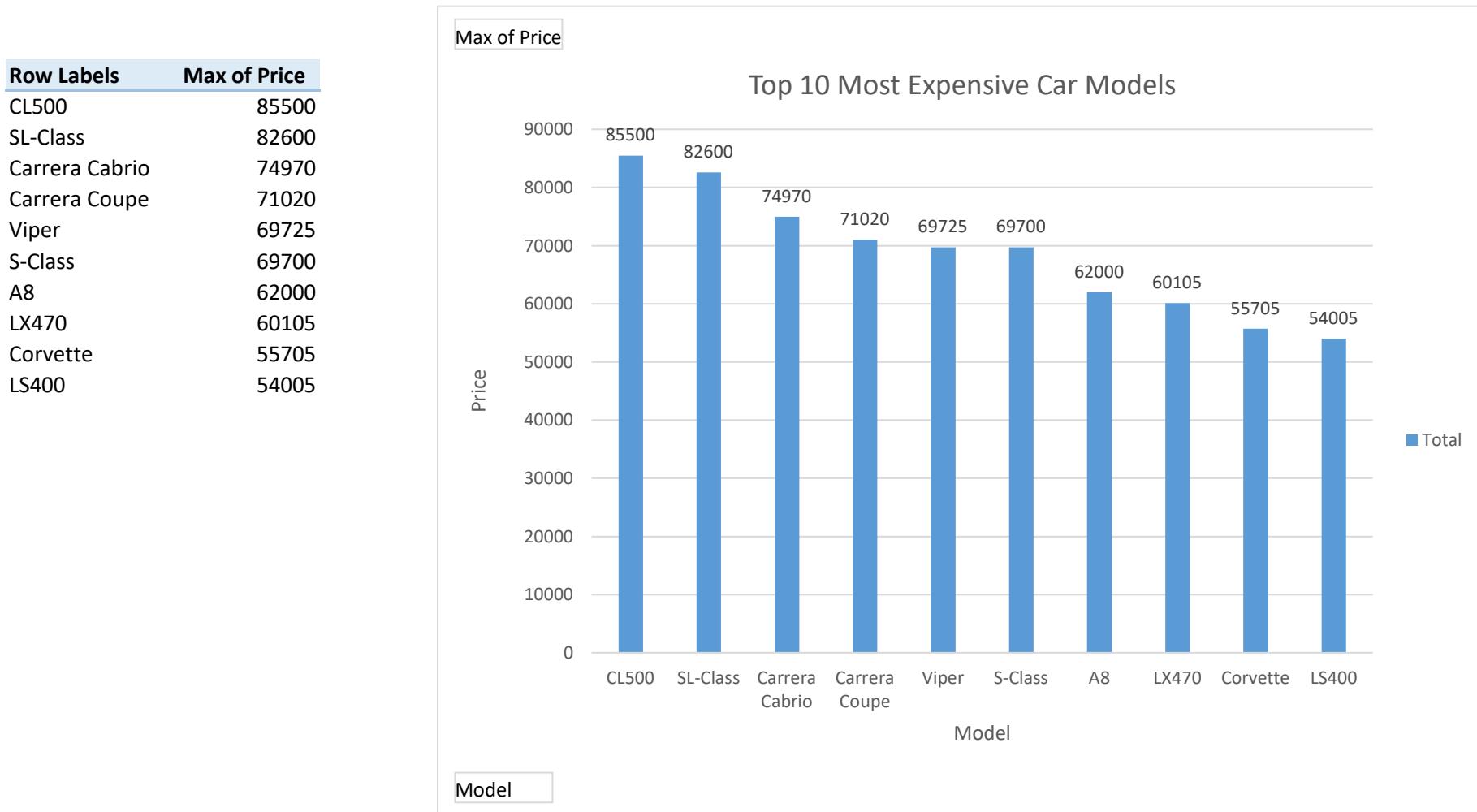
High HP vehicles have the lowest fuel efficiency (20 MPG).

This demonstrates a clear inverse relationship between horsepower and fuel efficiency. As engine power increases, fuel consumption also increases, resulting in lower MPG values.

This pattern reflects real-world automotive design, where high-performance vehicles prioritize speed and power over fuel economy.

Goal 4: Most Expensive Models

Prepared by Abdellah Boudlal



Goal

Identify the most expensive vehicle models in the dataset.

Chart

Column chart displaying the top 10 most expensive car models based on maximum price.

Story / Analysis

This visualization highlights the highest-priced vehicles within the dataset.

The analysis reveals that luxury and performance brands dominate the top price rankings.

The most expensive model identified is the **Mercedes-Benz CL500**, with a price exceeding \$85,000.

Other high-ranking models include:

Mercedes-Benz S-Class

Porsche Carrera Cabrio

Porsche Carrera Coupe

Dodge Viper

These vehicles are characterized by high horsepower engines, premium materials, and luxury features, which significantly increase their market price.

The chart demonstrates how luxury automotive brands position themselves in the high-end pricing segment.

Key Insight

Luxury and performance vehicles command significantly higher prices, reflecting brand value, engineering complexity, and premium market positioning.

Goal 5: Highest Resale Value

Prepared by Abdellah Boudlal



Goal

Identify which vehicle models retain the highest resale value.

Chart

Column chart displaying the highest resale value by vehicle model.

Story / Analysis

This visualization highlights the vehicle models that maintain the strongest resale value over time.

The analysis reveals that premium and luxury sports vehicles dominate the top resale rankings.

The model with the highest resale value is the **Porsche Carrera Cabrio**, exceeding \$67,000.

Other top performers include:

Mercedes-Benz CL500

Porsche Carrera Coupe

Mercedes-Benz S-Class

Dodge Viper

These vehicles maintain strong resale value due to high brand prestige, superior engineering quality, and sustained market demand.

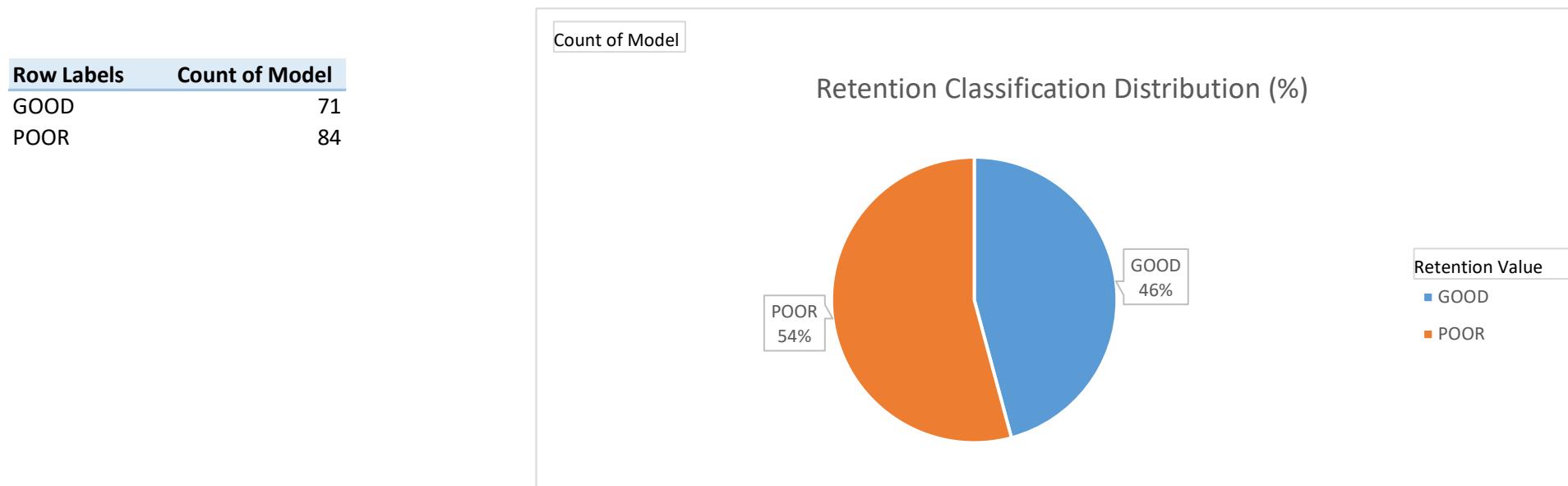
Luxury and performance vehicles depreciate at a slower rate compared to standard passenger vehicles.

Key Insight

Premium automotive brands demonstrate significantly higher resale retention, indicating stronger long-term value preservation.

GOAL 6 — Retention Distribution

Prepared by Abdellah Boudlal



Goal

Analyze the distribution of vehicle retention classifications.

Chart

Pie chart displaying the percentage distribution of vehicles classified as GOOD vs POOR retention.

Story / Analysis

This visualization illustrates the distribution of vehicle retention performance across the dataset.

The analysis shows that:

54% of vehicles fall under the POOR retention category.

46% of vehicles are classified as GOOD retention.

This indicates that the majority of vehicles depreciate at a faster rate, losing a significant portion of their resale value over time.

However, a substantial segment of vehicles maintains strong value retention, typically associated with premium brands and higher-quality engineering.

Key Insight

Vehicle depreciation is prevalent across the dataset, with more than half of models experiencing weaker resale retention.

Section 7 — Conclusion

This project explored automotive sales and vehicle performance data through multiple visualization perspectives.

Key findings include:

Certain manufacturers dominate total unit sales, reflecting strong brand market share.

Passenger vehicles show slightly higher average pricing compared to general vehicle categories.

Fuel efficiency decreases as horsepower increases, demonstrating the trade-off between performance and economy.

Luxury and performance models rank highest in both initial price and resale value.

Premium brands maintain stronger long-term value retention.

More than half of vehicles fall into the poor retention category, highlighting industry-wide depreciation trends.

Overall, the analysis demonstrates how data visualization can transform complex automotive datasets into actionable insights.

These findings could support manufacturers, dealerships, and investors in pricing strategy, product positioning, and long-term asset valuation decisions.