

# Car Price Analysis Project

This report analyses the relationships between performance, efficiency, brand reputation, and regional market segmentation in determining vehicle prices. It integrates Python (VS Code) and Tableau visualisations to build a coherent and data-driven narrative of the global car market.

## Hypothesis 1: Price Segmentation through Performance Indicators

### Context

The analysis begins with a correlation heatmap built in Python (Seaborn), visualising numeric relationships among key variables such as price, engine size, horsepower, and mileage. The heatmap identifies strong positive correlations between engine size, horsepower, and price, forming the analytical foundation for price segmentation.

### Hypothesis

Car prices follow a structured segmentation driven by engine capacity and performance metrics. Vehicles with larger engines and higher horsepower tend to belong to higher value segments, while smaller engine vehicles form the economy class.

### Charts Used

Correlation Heatmap (Seaborn): numeric variable relationships.

Scatter Plot (Engine Size vs Price): performance-based price segmentation.

### Interpretation

The heatmap visually confirms a strong relationship between engine size and price, suggesting performance as a key price driver. The scatter plot expands on this finding: as engine size increases, vehicle prices rise proportionally. The trend line ( $R^2 \approx 0.45$ ,  $p < 0.0001$ ) confirms the relationship is statistically significant. Higher priced cars cluster at larger engine sizes, while smaller engines dominate the low price range typical of economy cars.

### Conclusion

The hypothesis is supported. Car prices reflect engine capacity and performance, establishing the foundation for three market tiers: Economy, Mid-range, and Luxury. The correlation heatmap and scatter analysis together validate the performance-based segmentation of the car market.

## Hypothesis 2: Performance, Efficiency, and Value Perception

### Context

Building upon the price segmentation, this section evaluates the dual relationship between horsepower, fuel efficiency, and vehicle price. The advanced scatter plots were created in Tableau, including an additional comparison of highway MPG versus price to deepen the analysis. Two supporting scatter plots were also created in VS Code by a colleague to validate the same patterns from the Python environment.

### **Hypothesis**

Higher horsepower leads to higher vehicle prices but lower fuel efficiency. Diesel vehicles maintain better mileage and higher price points compared to gas vehicles at similar performance levels.

### **Charts Used**

Advanced Tableau Scatters (author): Horsepower vs City MPG; Horsepower vs Price; Highway MPG vs Price (additional advanced visual). Supporting VS Code Scatters (colleague): Two validation scatter plots reproducing the horsepower-MPG and horsepower-price relationships.

### **Interpretation**

Both the Tableau and VS Code visuals reveal an inverse relationship between horsepower and fuel efficiency: as engine power increases, city mileage declines. Diesel vehicles maintain higher MPG than gas cars at comparable horsepower levels. The horsepower-price relationship shows a positive correlation: vehicles with higher horsepower are more expensive. The highway MPG vs price comparison further reinforces this trend: fuel efficient vehicles tend to occupy lower price tiers, while high performance models dominate the upper range.

### **Conclusion**

The hypothesis is supported. A clear trade-off exists between performance and efficiency: cars with greater power command higher prices but lower fuel economy. Diesel vehicles perform best in balancing these attributes, demonstrating efficiency and premium appeal simultaneously.

## **Hypothesis 3: Brand Level Pricing and Market Positioning**

### **Context**

After exploring the influence of technical attributes, this section examines how brand identity impacts pricing. The analysis begins with a bar chart in VS Code showing the average price per brand, followed by two Tableau visuals: median price per brand and a boxplot showing brand-level price distribution.

### **Hypothesis**

Car brands position themselves across the market spectrum according to perceived quality, design, and prestige. Luxury brands (Jaguar, Porsche, BMW) command higher median prices, while economy brands (Toyota, Nissan, Honda) maintain affordability through price consistency.

## **Interpretation**

Luxury-oriented European brands such as Jaguar, Buick, and Porsche record the highest average and median prices. Mid-tier brands BMW, Volvo, Audi, Peugeot, and Mercury occupy the 15,000 to 25,000 range, while Toyota, Nissan, and Honda cluster between 10,000 and 15,000. The boxplot highlights wider price variation in luxury brands, while economy brands maintain tighter, consistent price distributions.

## **Additional Insight – Design & Pricing Patterns**

Complementary analysis of car body types further supports the brand-level segmentation. Convertible and hardtop models exhibit the highest average prices (£22,000+), reflecting their association with premium and performance-focused brands. Family-oriented designs such as hatchbacks and sedans occupy the mid to lower tiers, aligning with mainstream manufacturers like Toyota, Nissan, and Honda. This confirms that beyond brand identity, design and body style also play a defining role in perceived market value.

## **Conclusion**

The hypothesis is strongly supported. Brand prestige significantly influences pricing beyond performance metrics. European and American luxury brands dominate the premium tier, while Asian and US manufacturers lead the economy class. The combined visuals illustrate a brand-driven hierarchy that complements the earlier performance-based segmentation.

# **Hypothesis 4: Regional Distribution and Market Segmentation**

## **Context**

This final section builds on the Price Distribution vs Segments Bar Chart created in VS Code, which highlights how pricing tiers vary across regions. The Tableau visuals then expand this view: mapping average prices and segment dominance by country and region.

## **Hypothesis**

The global car market exhibits clear regional segmentation. European brands dominate the mid-class and luxury tiers, while Asian manufacturers lead in the economy segment.

## **Interpretation**

The VS Code bar chart establishes baseline differences in price segments among regions. The Tableau map visualises distinct clusters: Europe shows high concentrations of luxury cars, Asia dominates economy models, and North America presents a balanced distribution. The Tableau bar chart quantifies these distinctions, confirming that Europe leads the mid to premium market share, Asia dominates the economy class, and North America remains diversified.

## **Conclusion**

The hypothesis is supported. Regional segmentation aligns with continental strengths: Europe (Premium and Mid-Class), Asia (Economy and Efficiency-focused), and North America (Balanced Range). By linking VS Code and Tableau visuals, the findings provide a unified understanding of how geography and brand origin shape market positioning.

## **Overall Summary**

The project presents a connected analysis of vehicle pricing using Python and Tableau. It begins with correlation-based segmentation and progresses through performance-efficiency relationships, brand perception, and regional dynamics. The findings confirm that performance and engine size drive baseline pricing; fuel efficiency inversely affects perceived value; brand reputation increases price tolerance; and regional origin reinforces market segmentation. This integrated approach combines analytics and business interpretation to offer actionable insights into how the automotive industry positions and prices its products globally.