Google Data Analytics Capstone Project Presentation

Project Title: Analyzing Automotive Data for Market Insights

Author: Lars Petschke
Date: 2025-06-04

Ask

Business Task:
"What factors influence the price and popularity of cars, and how are engine performance and fuel efficiency related to these factors?"

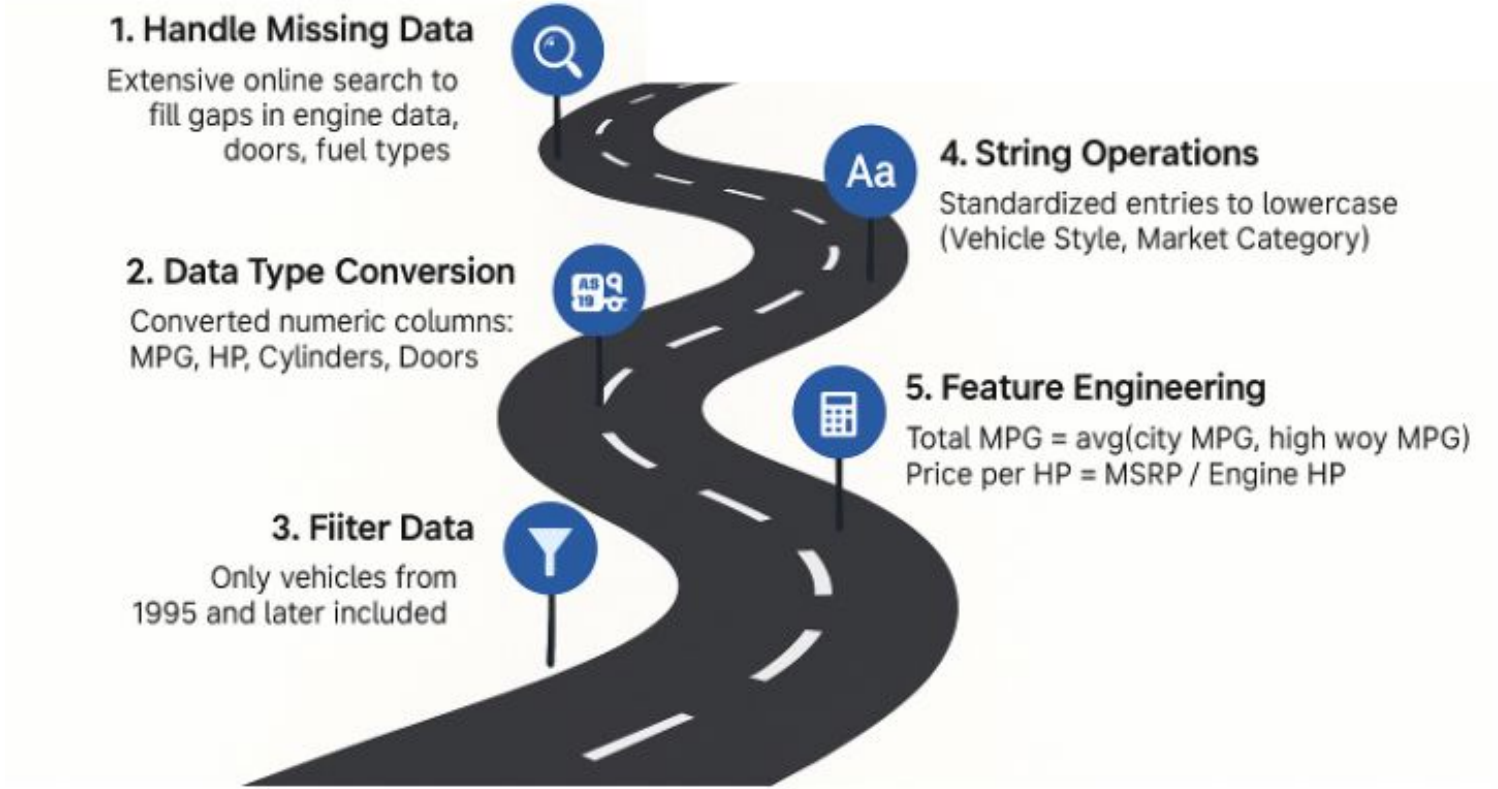
Project Objective:
To uncover pricing patterns, market trends, and performance insights in automotive data to support business decisions in marketing and positioning.

Prepare

Dataset Summary:

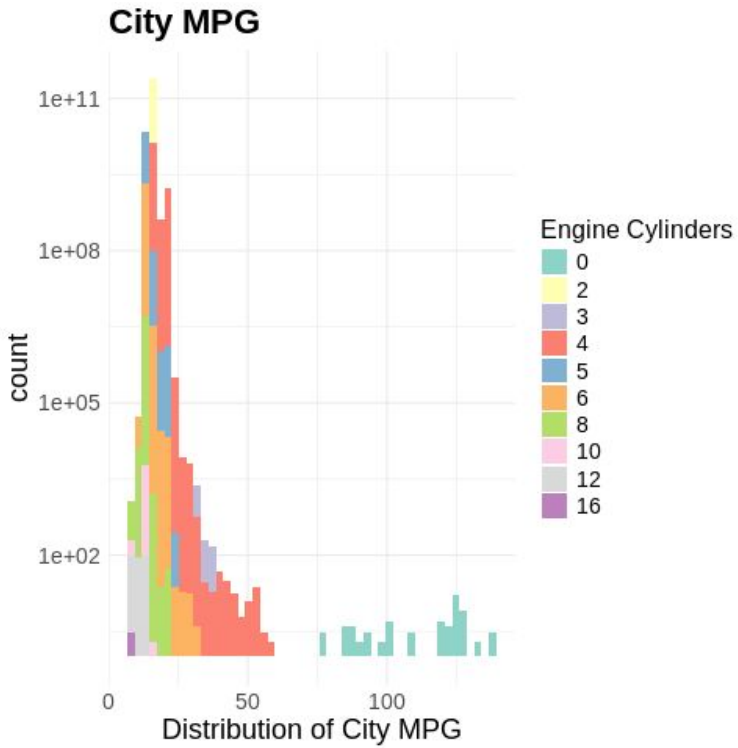
- Includes attributes: Make, Model, Year, Engine Type, Fuel Economy, Transmission, Drivetrain, Popularity, MSRP
- Focus: Cars from 1995 onward
- Manual internet research used to fill in missing values
- Duplicates were removed, data consistency ensured

Process


1. Handle Missing Data
Extensive online search to fill gaps in engine data, doors, fuel types
2. Data Type Conversion
Converted numeric columns: MPG, HP, Cylinders, Doors
3. Fiiter Data
Only vehicles from 1995 and later included
4. String Operations
Standardized entries to lowercase (Vehicle Style, Market Category)
5. Feature Engineering
Total MPG = avg(city MPG, high woy MPG)
Price per HP = MSRP / Engine HP

Analyze

1. City MPG - Distribution by Engine Cylinders

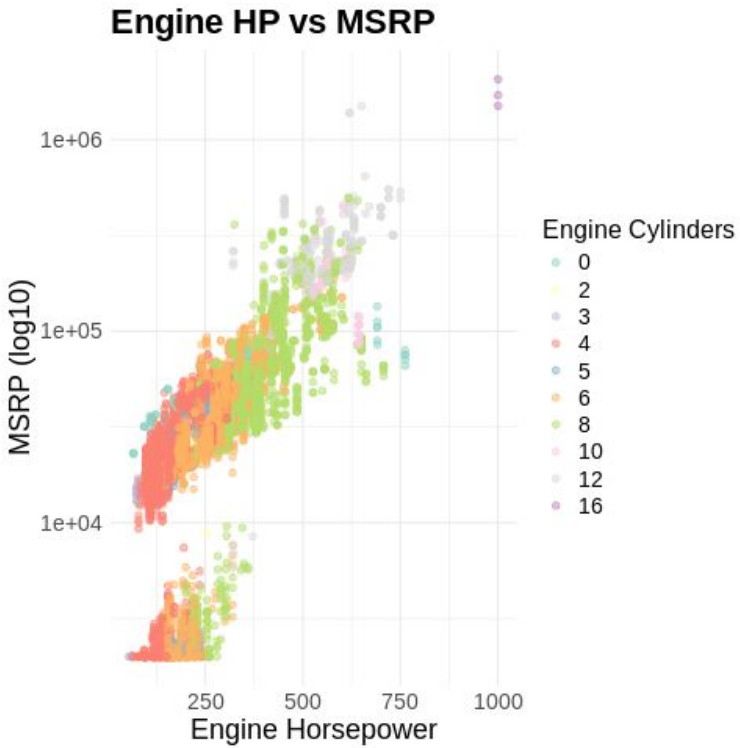


Vehicles with fewer cylinders (0–4)
achieve significantly **higher city MPG**
→ more fuel-efficient.

Engines with more cylinders (6–16)
consume **more fuel**
→ lower city MPG.

Outliers with extremely high MPG
mostly represent **non-combustion vehicles**
(e.g., electric cars).

Analyze 2. Engine HP vs MSRP by Engine Cylinders

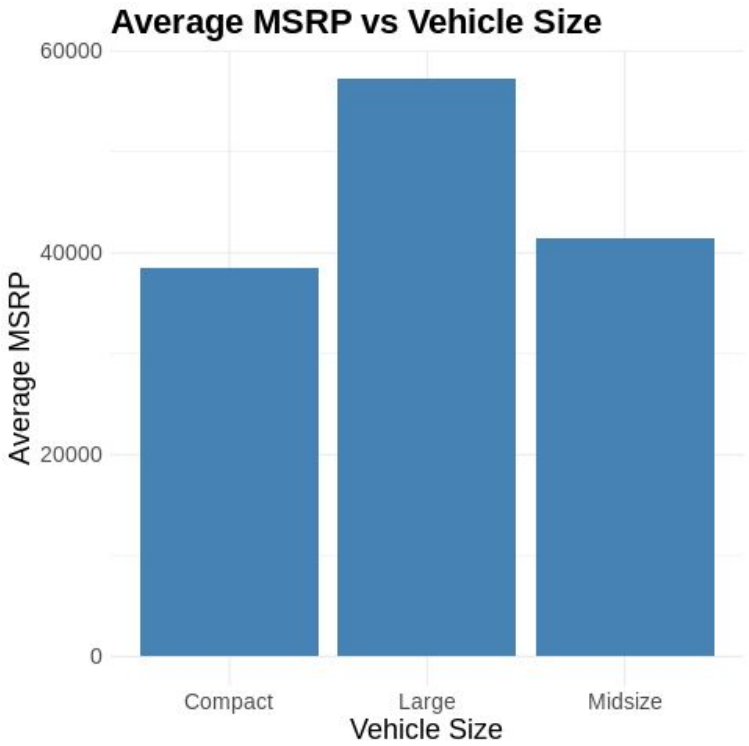


Higher horsepower strongly correlates with higher MSRP (logarithmic scale).

Vehicles with more cylinders (6–16) cluster in the high power and high price range.

Luxury and performance cars dominate the top-right corner with **>500 HP and high MSRP**.

Analyze 3. Average MSRP vs Vehicle Size



Large vehicles have the **highest average MSRP** (~$58,000).

Midsize cars are moderately priced (~$42,000).

Compact cars are the most affordable (~$38,000).

Analyze

4. MSRP Distribution by Driven Wheels

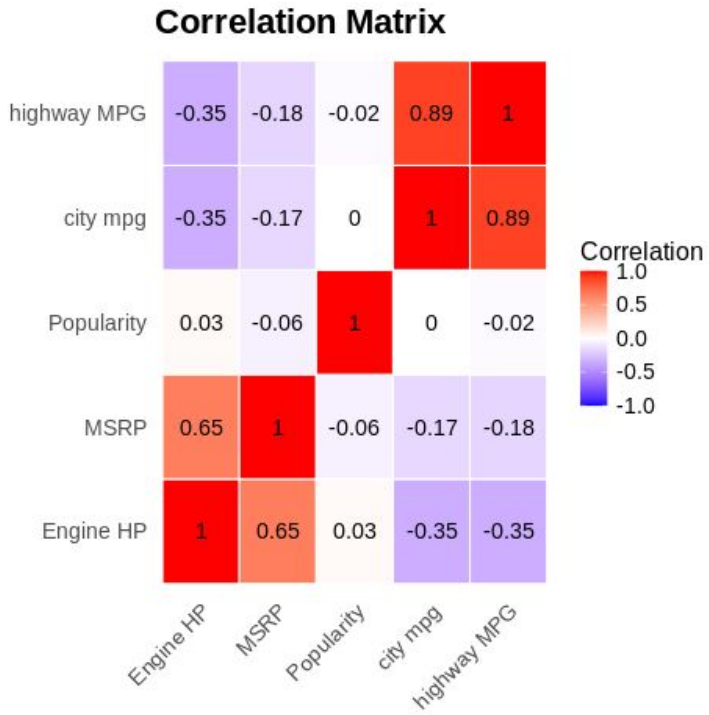
AWD & RWD vehicles have similar **median MSRP,** although RWD has a wider price range.

AWD & RWD models show several outliers to high price category.

FWD cars are generally the **least expensive**.

Analyze

5. MSRP Distribution by Driven Wheels



Positive: MSRP strongly correlates with **Engine Horsepower**.

Negative: MPG (City & Highway) negatively correlates with **Engine Horsepower**.

No Correlation: Popularity shows no meaningful correlation with any numeric feature.

High Internal Correlation: City MPG and **Highway MPG** are strongly positively correlated.

Share – Key Insights and Findings

- High horsepower and AWD linked to premium MSRP
- Compact/midsize vehicles deliver best fuel economy
- 'Direct Drive' transmissions provide top MPG values
- Popularity metric not predictive of price or performance

Act – Recommendations

- Market compact/midsize cars as economical and eco-friendly.
- Position AWD/high-HP cars in luxury/performance categories.
- Use 'Direct Drive' MPG advantage in product promotion.
- Focus on measurable specs over abstract popularity scores.

Thank You

Final Notes

- Tools: Python, R, Pandas, ggplot2, Google Colab
- Structure followed:

  Ask → Prepare → Process → Analyze → Share → Act

- Visualizations and notebook available for stakeholders
- Connect with me: LinkedIn | GitHub