Used Car Price Analysis

By: Sk Vaheed

Date: April 13, 2025

# Abstract

This project focuses on analyzing the price trends of used cars listed on CarWale. Using web scraping, data was collected, cleaned, and visualized to derive insights into car pricing patterns, brand dominance, fuel preferences, and other relevant metrics.

# Problem Statement

• Understand how car prices vary across different brands and models.

• Identify budget-friendly cars offering good value.

• Discover patterns in car listings such as fuel types, mileage, and year of manufacture.

# Objective

The main goal of this project is to extract, clean, and analyze used car listing data to help buyers and sellers make informed decisions based on market insights.

# Tools & Technologies Used

• Python: Selenium, BeautifulSoup for web scraping; Pandas for data cleaning.

• Power BI: For visualization and dashboard creation.

• Regex: For data extraction and text manipulation.

• CSV: For data storage and transfer.

# Data Collection

The data was collected from CarWale’s used car listings specific to Hyderabad, using automated web scraping techniques.

The website requires scrolling to load full content, handled using JavaScript execution within Selenium.

Extracted fields include brand, model, year, fuel type, transmission, location, price, and EMI.

# Data Cleaning & Preparation

Data was processed using regular expressions to extract structured fields from raw HTML content.

Noise and inconsistencies were removed, and values were categorized where necessary.

Cleaned data was stored in CSV format for further analysis and dashboard creation.

# Data Analysis & Visualizations

A Power BI dashboard was created with key metrics and charts:

• Total cars listed: 729

• Average KM driven: 57.33K

• Most common fuel type: Petrol

• Top listed brands: Tata, Maruti, MG, Hyundai

• Price range distribution: Majority above ₹10L

• Average price highest for Diesel cars

# Challenges Faced

• Dynamic content loading requiring simulated scrolling.

• Data formatting issues like embedded symbols and inconsistent delimiters.

• Occasional bot protection during scraping.

# Conclusion

This project successfully demonstrates the use of web scraping, data cleaning, and analysis for real-world applications.

It provides valuable insights into used car markets and helps users find good quality cars at reasonable prices.

# References

• https://www.carwale.com

• Python documentation: selenium.dev, pandas.pydata.org

• Power BI resources: powerbi.microsoft.com