Capstone Project:Price Prediction

By: Rutik Kadam

Introduction to Vehicle Valuation

We obtained a comprehensive dataset from Kaggle, which includes various features and attributes relevant to our research. This dataset contains numerous entries, allowing for extensive analysis and insights into the subject matter. Our goal is to leverage this data to uncover patterns and trends that can inform our findings.

selling_price	km_driven	fuel	seller_type	transmission
60000	70000	Petrol	Individual	Manual
135000	50000	Petrol	Individual	Manual
600000	100000	Diesel	Individual	Manual
250000	46000	Petrol	Individual	Manual
450000	141000	Diesel	Individual	Manual
140000	125000	Petrol	Individual	Manual
550000	25000	Petrol	Individual	Manual
240000	60000	Petrol	Individual	Manual
850000	25000	Petrol	Individual	Manual
365000	78000	CNG	Individual	Manual
260000	35000	Petrol	Individual	Manual
250000	100000	Petrol	Individual	Manual
1650000	25000	Petrol	Dealer	Automatic
60000	70000	Petrol	Individual	Manual
135000	50000	Petrol	Individual	Manual
600000	100000	Diesel	Individual	Manual
250000	46000	Petrol	Individual	Manual
450000	141000	Diesel	Individual	Manual
140000	125000	Petrol	Individual	Manual
550000	25000	Petrol	Individual	Manual
240000	60000	Petrol	Individual	Manual
850000	25000	Petrol	Individual	Manual
365000	78000	CNG	Individual	Manual
260000	35000	Petrol	Individual	Manual
250000	100000	Petrol	Individual	Manual

Understanding Vehicle Pricing Factors

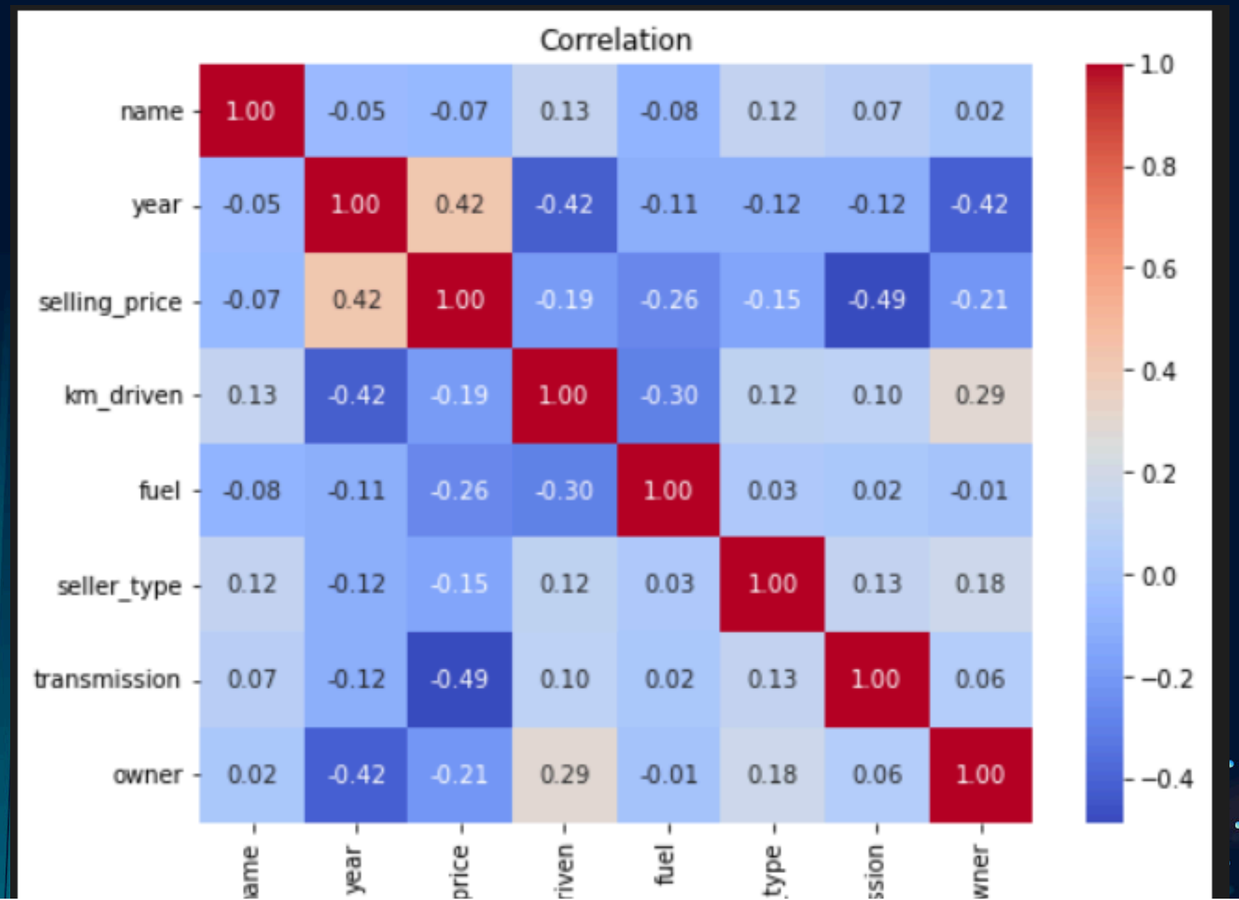
Several **key factors** influence vehicle pricing, including **Type,**, **year**, and **condition, Driven, Transmission**. Our research highlights how these elements interact and affect overall market value. By analyzing these factors, we can develop a more precise valuation model that reflects true market conditions.

Data Filtering by EDA

Data cleaning is essential for ensuring the accuracy and reliability of datasets. Various methods to clean data include removing duplicates to avoid redundancy, handling null values by either replacing them with meaningful data or deleting them, and standardizing formats to maintain consistency. Additionally, validating data through checks can enhance overall quality.

Model Development and Techniques

Exploratory Data Analysis (EDA) charts are essential tools for gaining deeper insights into a dataset. Various types of EDA charts, such as histograms, box plots, and scatter plots, allow analysts to visualize patterns, trends, and anomalies, facilitating a comprehensive understanding of the data's structure and relationships.



Selecting Best Model

The results of our analysis demonstrated a significant improvement in **valuation accuracy** compared to traditional methods. Our model achieved a high **R-squared value**, indicating a strong correlation between predicted and actual prices. These findings underscore the potential of data-driven approaches in enhancing vehicle valuation.

```
Evalutaing the Model Linear Regression

MAE 215522.4645487041
MSE 197896678171.88495
RMSE 444855.7948053335
R2 score 0.3856684691225903


Training Score 0.44058121312139 11
Testing Score 0.3856684691225903
```

```
Evalutaing the Model: KNN

MAE 239459.16861350942
MSE 276148439805.9148
RMSE 525498.2776431477
R2 score 0.1427511744754 8002


Training Score 0.3134740452692356
Testing Score 0.1427511744754 8002
```

Car Price Prediction App

Predict the selling price of a used car based on its characteristics

Year of Manufacture

2015

Kilometers Driven

50000

Fuel Type CNG, Diesel,Electric,LPG,Petrol

0

Seller Type Dealer, Individual ,Trust Dealer

0

Transmission Automatic,Manual

0

Owner 1st Owner, 4th & above owner, 2nd owner, Test Drive, 3rd Owner

0

Predict Price

The predicted selling price of the car is: ₹ 737636.27

Thanks!

GIT:https://github.com/Mercs-r/Car_value_prediction.git

Steamlib:https://carvalueprediction-clpfjenjv8wbj8szqm3uxm.streamlit.app/