Car Market Trends Analysis
with Car Dekho Data -

VENKATA SATYA SRIKAR
PARVATHA

PROBLEM STATEMENT

Analyze the Car Dekho dataset to identify key trends and insights in the car market. The analysis should focus on understanding the factors influencing car prices, identifying popular car features, and examining market dynamics over time.

This information will help stakeholders, including buyers, sellers, and manufacturers, make informed decisions. Key tasks include data cleaning, exploratory data analysis (EDA), data visualization, and deriving actionable insights from the data.

Project Description

The automotive industry is rapidly evolving, influenced by shifting consumer preferences, technological advancements, and economic factors. This project focuses on analyzing the Car Dekho dataset, which encompasses detailed information about various car models, their specifications, prices, and other relevant attributes. By examining this dataset, we aim to extract meaningful insights that can help stakeholders, including manufacturers, dealers, and consumers, understand the current trends and dynamics in the car market.

The analysis begins with data cleaning and preprocessing to ensure the dataset is accurate and consistent. This involves handling missing values, correcting data types, and removing any duplicates. Following this, exploratory data analysis (EDA) will be conducted to uncover patterns and relationships within the data. Key features such as brand, year, mileage, and fuel type will be examined to determine their impact on car prices. Additionally, the project will analyze market trends over time, identifying popular car models and features, and studying depreciation trends to understand how car values change with age and usage.

The insights derived from this analysis will be presented using various visualization techniques, making it easier to communicate the findings effectively. The final outcome of this project will include actionable recommendations for stakeholders, such as identifying the best time to buy or sell a car, understanding which features add the most value, and anticipating future market trends. These insights will enable better decision-making, ultimately benefiting all parties involved in the automotive industry.

Objectives

The primary objectives of this project are:

Data Cleaning and Preprocessing: Ensure the dataset is clean and consistent by handling missing values, correcting data types, and removing duplicates.

Exploratory Data Analysis (EDA): Conduct thorough EDA to understand the distribution of variables, identify patterns, and uncover initial insights.

Feature Analysis: Examine the relationship between car prices and various features such as brand, year, mileage, fuel type, and more. Determine which features have the most significant impact on car prices.

Market Trends and Dynamics: Analyze trends over time to understand how the popularity of different car models and features has evolved. Identify the most and least popular car features.

Depreciation Analysis: Investigate how car values depreciate over time, considering factors like age, mileage, and condition.

Data Visualization: Use matplotlib and other visualization tools to create clear and informative visual representations of the data, highlighting key insights.

Conclusions and Recommendations: Summarize the findings from the analysis and provide actionable insights for various stakeholders, including car buyers, sellers, and manufacturers.

Methodology

- **Data Cleaning**: Handle missing values, correct data types, and remove duplicates.
- **Exploratory Data Analysis**: Use pandas and numpy for data manipulation and summary statistics. Utilize matplotlib for visualizations.
- **Feature Engineering**: Create new features that might be useful for analysis or predictive modeling.
- **Visualization**: Develop various plots (scatter plots, bar charts, histograms, etc.) to visualize data distributions and relationships.
- **Analysis and Insights**: Conduct in-depth analysis to understand the factors affecting car prices and market trends.

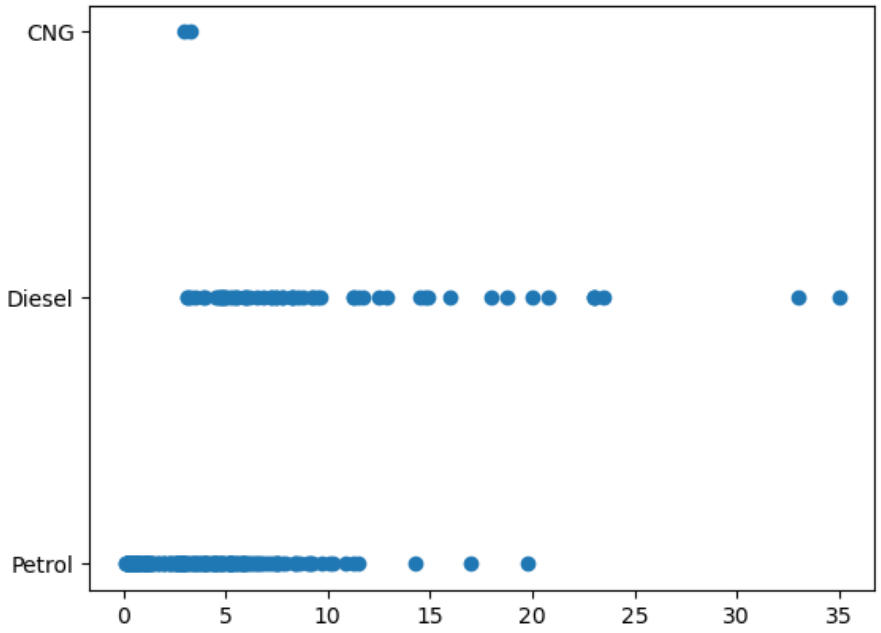
WHO ARE THE END USERS?

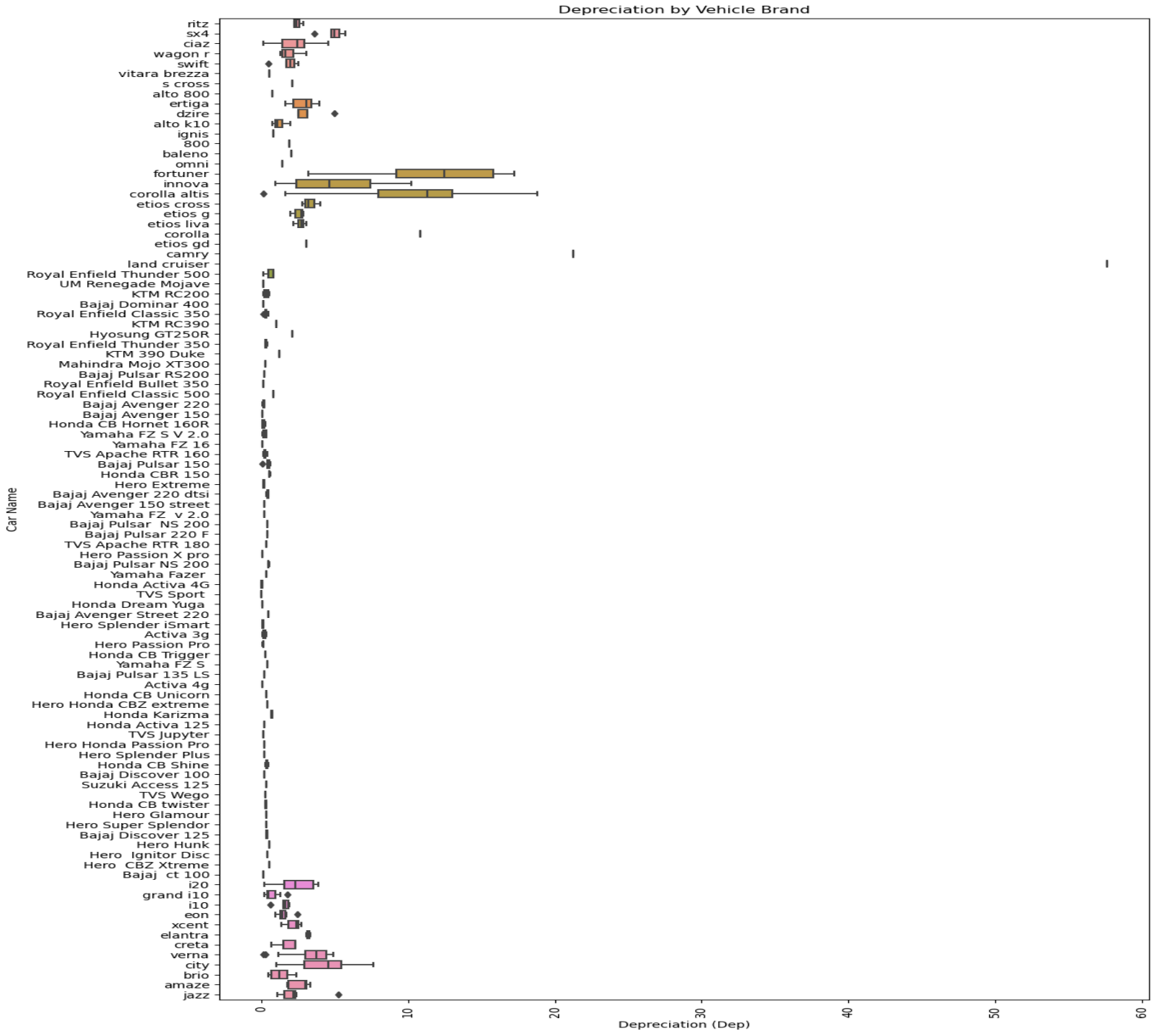
•**Car Buyers**: Individuals looking to purchase a car can use the insights to make informed decisions about which models offer the best value, understand depreciation trends, and identify the most desirable features.
•**Car Sellers and Dealers**: Dealerships and individual sellers can leverage the analysis to price their vehicles competitively, understand market demand, and identify which features and models are most popular among buyers.
•**Car Manufacturers**: Manufacturers can use the data to gain insights into consumer preferences, identify trends in the market, and make strategic decisions regarding production and marketing.
•**Market Analysts and Researchers**: Analysts and researchers studying the automotive market can use the findings to understand broader trends, forecast future developments, and provide advice to stakeholders in the industry.
•**Financial Institutions**: Banks and lending institutions can benefit from understanding car depreciation trends and market values, aiding in the valuation of car loans and leases.
•**Policy Makers**: Government agencies and policy makers can use the insights to inform decisions related to transportation policies, environmental regulations, and economic planning.

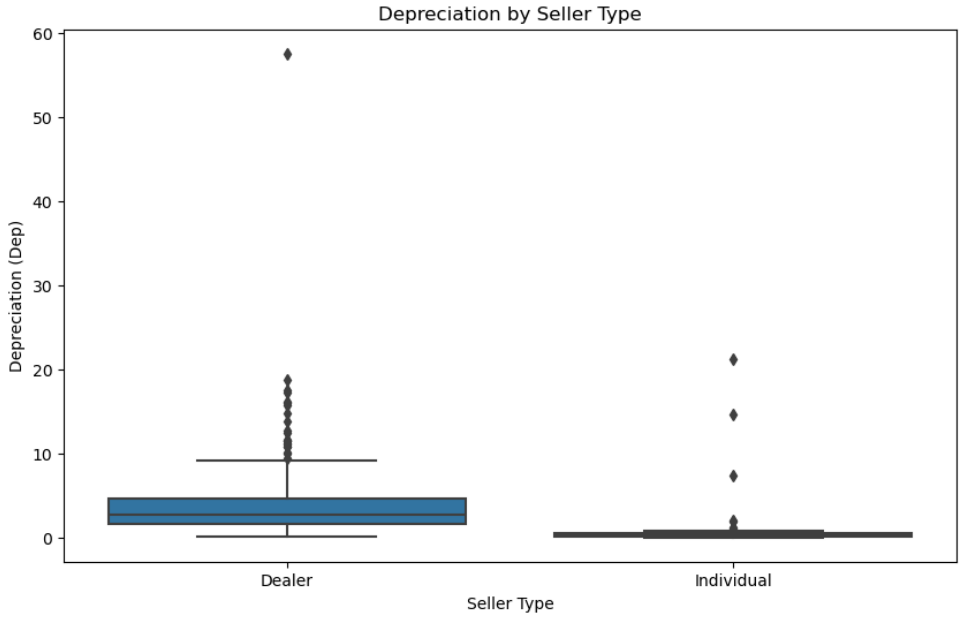
Technology Used

•**Python**: The primary programming language used for data analysis and manipulation.
•**Pandas**: A powerful data manipulation and analysis library in Python used for data cleaning, preprocessing, and exploratory data analysis (EDA).
•**NumPy**: A library for numerical operations in Python, providing support for arrays, matrices, and a collection of mathematical functions.
•**Matplotlib**: A plotting library in Python used for creating static, interactive, and animated visualizations to represent the data insights.
•**Jupyter Notebook**: An open-source web application that allows the creation and sharing of documents containing live code, equations, visualizations, and narrative text.
•**Seaborn**: A Python visualization library based on matplotlib that provides a high-level interface for drawing attractive and informative statistical graphics.
•**Scikit-learn**: A machine learning library in Python used for implementing machine learning algorithms, which may be utilized for advanced analysis such as predictive modeling.
•**Google Colab**: A cloud service that supports Jupyter notebooks and provides free access to computational resources, often used for coding and running Python notebooks.

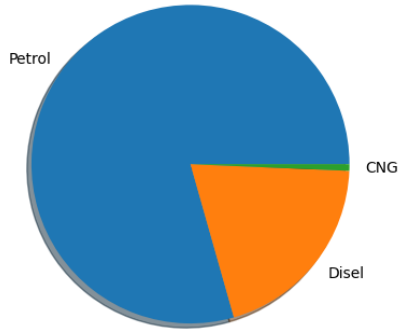
EDA RESULTS


CNG
Diesel
Petrol
0
5
10
15
20
25
30
35

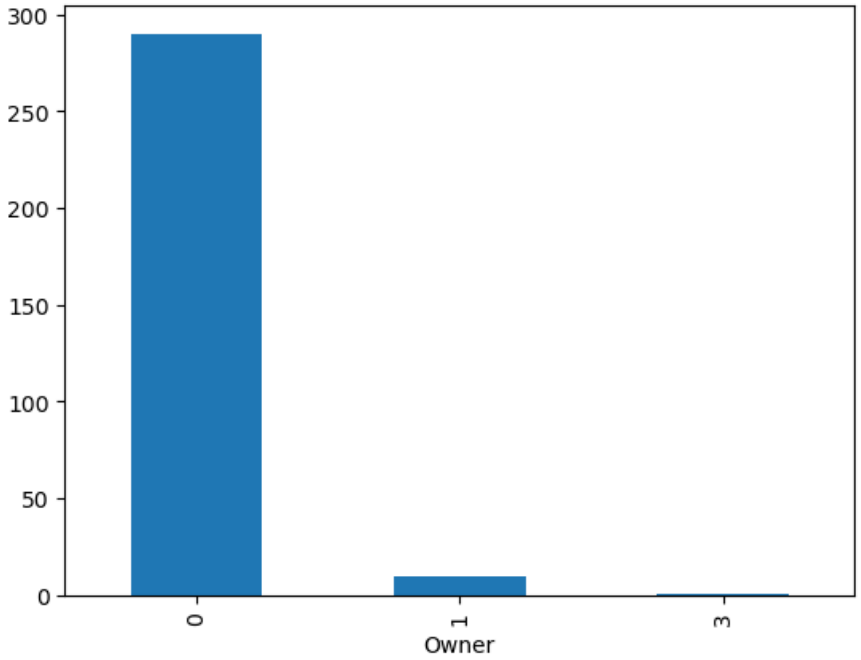

Depreciation by Vehicle Brand
Car Name
ritz
sx4
ciaz
wagon r
swift
vitara brezza
s cross
alto 800
ertiga
dzire
alto k10
ignis
800
baleno
omni
fortuner
innova
corolla altis
etios cross
etios g
etios liva
corolla
etios gd
camry
land cruiser
Royal Enfield Thunder 500
UM Renegade Mojave
KTM RC200
Bajaj Dominar 400
Royal Enfield Classic 350
KTM RC390
Hyosung GT250R
Royal Enfield Thunder 350
KTM 390 Duke
Mahindra Mojo XT300
Bajaj Pulsar RS200
Royal Enfield Bullet 350
Royal Enfield Classic 500
Bajaj Avenger 220
Bajaj Avenger 150
Honda CB Hornet 160R
Yamaha FZ S V 2.0
Yamaha FZ 16
TVS Apache RTR 160
Bajaj Pulsar 150
Honda CBR 150
Hero Extreme
Bajaj Avenger 220 dtsi
Bajaj Avenger 150 street
Yamaha FZ v 2.0
Bajaj Pulsar NS 200
Bajaj Pulsar 220 F
TVS Apache RTR 180
Hero Passion X pro
Bajaj Pulsar NS 200
Yamaha Fazer
Honda Activa 4G
TVS Sport
Honda Dream Yuga
Bajaj Avenger Street 220
Hero Splender iSmart
Activa 3g
Hero Passion Pro
Honda CB Trigger
Yamaha FZ S
Bajaj Pulsar 135 LS
Activa 4g
Honda CB Unicorn
Hero Honda CBZ extreme
Honda Karizma
Honda Activa 125
TVS Jupyter
Hero Honda Passion Pro
Hero Splender Plus
Honda CB Shine
Bajaj Discover 100
Suzuki Access 125
TVS Wego
Honda CB twister
Hero Glamour
Hero Super Splendor
Bajaj Discover 125
Hero Hunk
Hero Ignitor Disc
Hero CBZ Xtreme
Bajaj ct 100
i20
grand i10
i10
eon
xcent
elantra
creta
verna
city
brio
amaze
jazz
0
10
20
30
40
50
60
Depreciation (Dep)

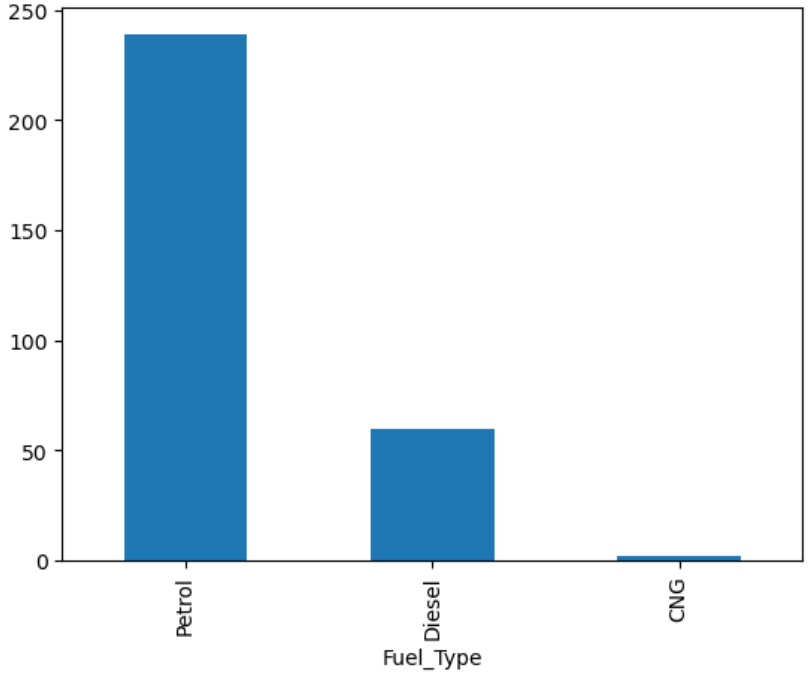

Depreciation by Seller Type
Depreciation (Dep)
60
50
40
30
20
10
0
Dealer
Individual
Seller Type

EDA RESULTS

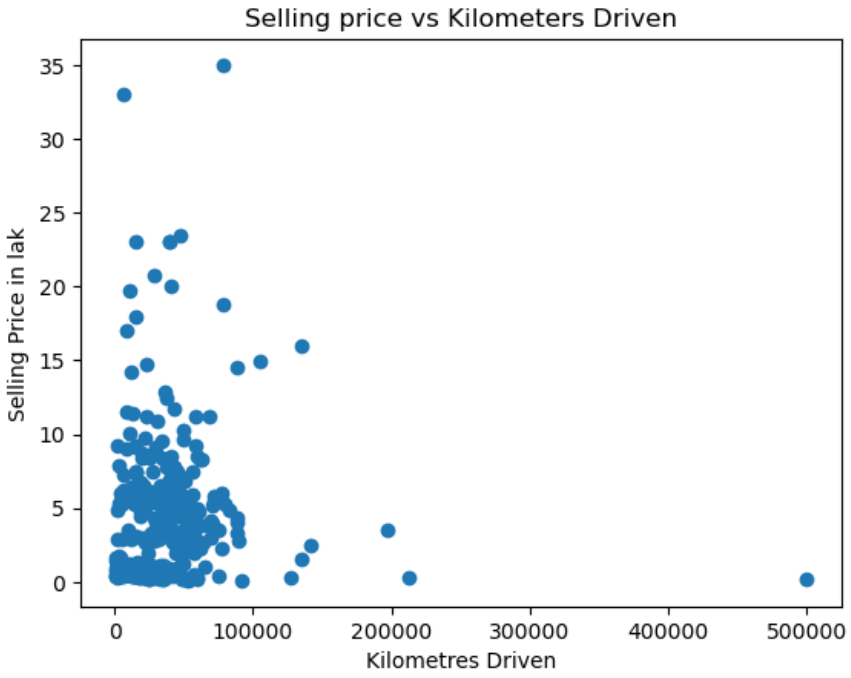

Petrol
CNG
Disel

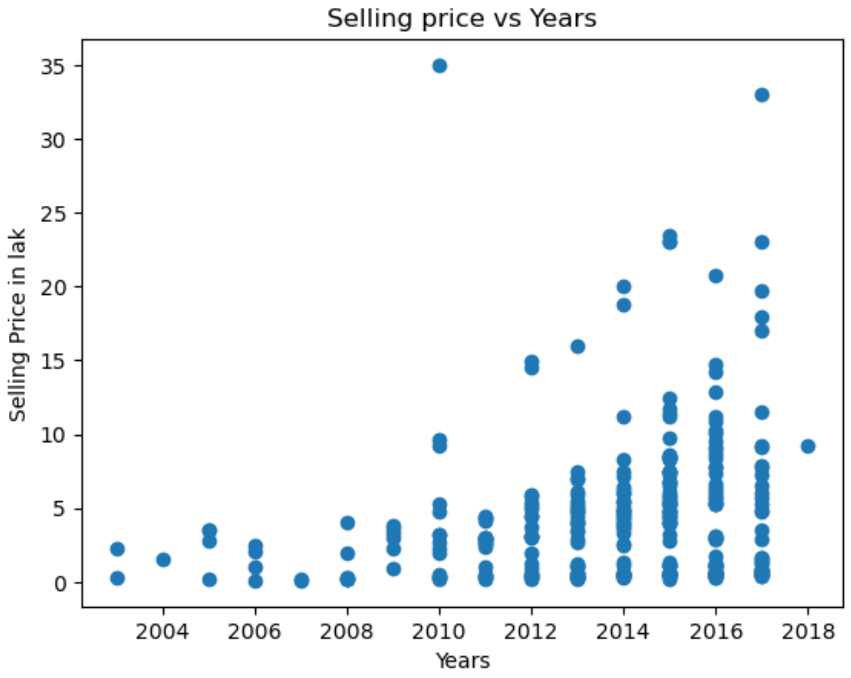
EDA RESULTS


300
250
200
150
100
50
0
0
1
3
Owner

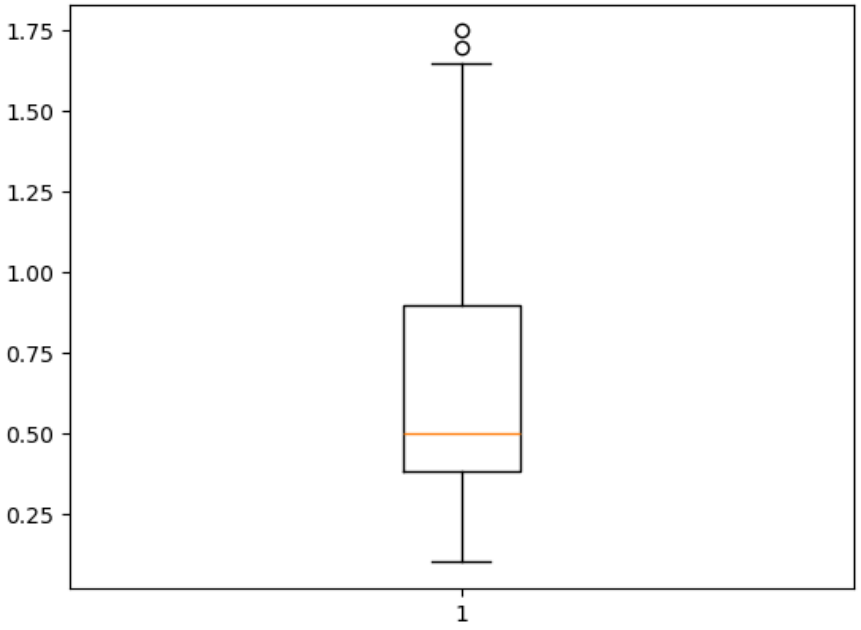

250
200
150
100
50
0
Petrol
Diesel
CNG
Fuel_Type

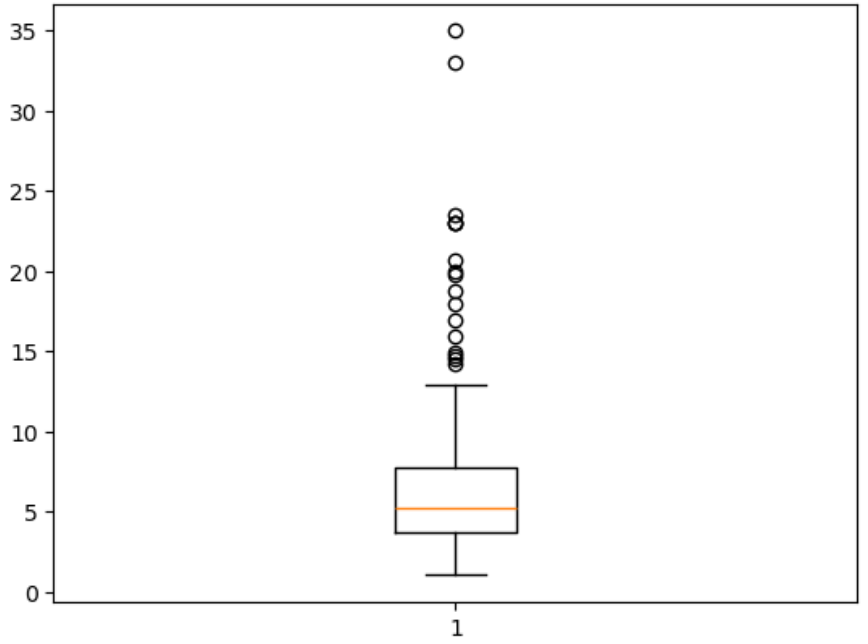
EDA RESULTS


Selling price vs Kilometers Driven
Selling Price in lak
Kilometres Driven
0
5
10
15
20
25
30
35
0
100000
200000
300000
400000
500000

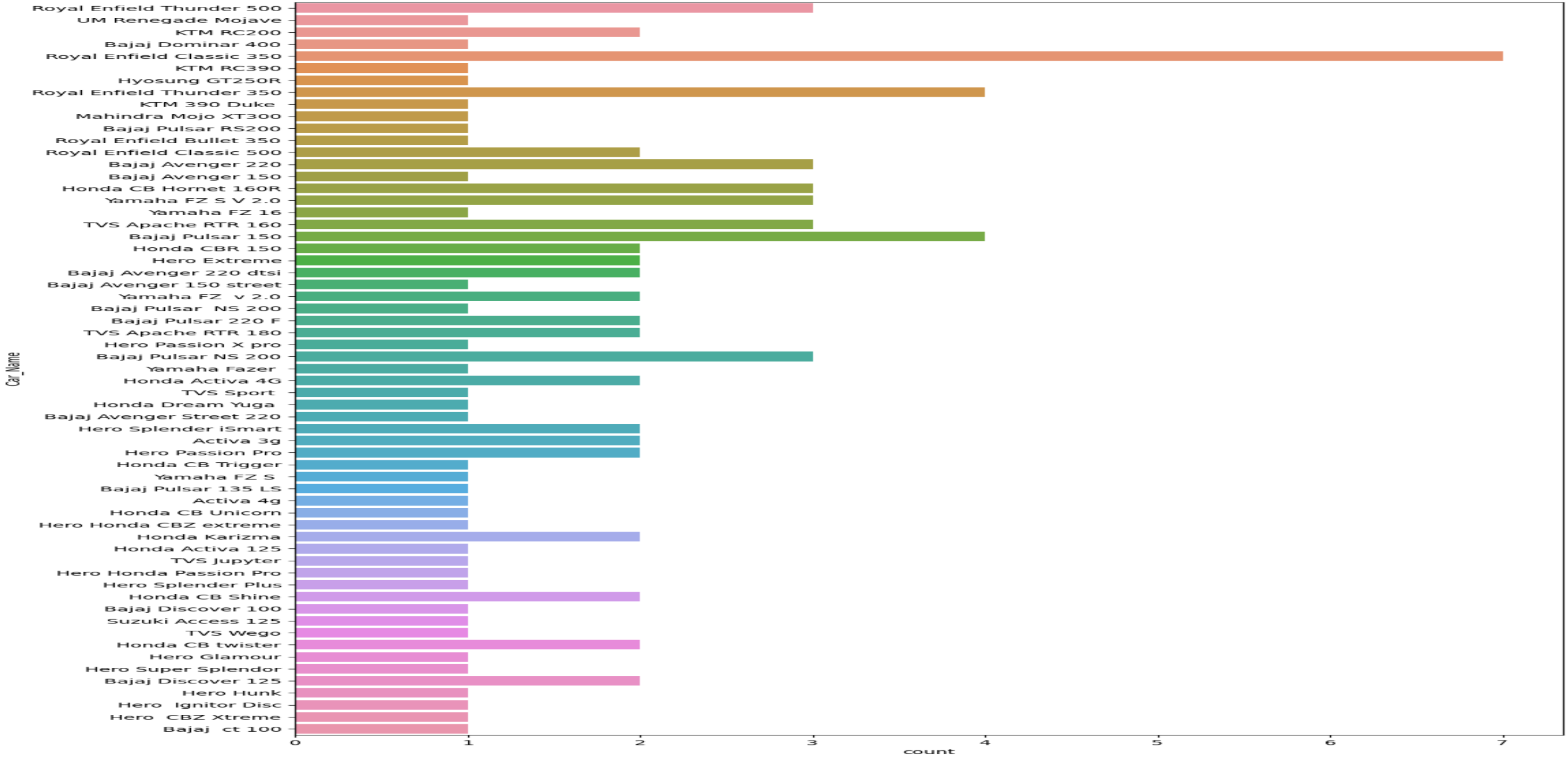

Selling price vs Years
Selling Price in lak
Years
0
5
10
15
20
25
30
35
2004
2006
2008
2010
2012
2014
2016
2018

EDA RESULTS

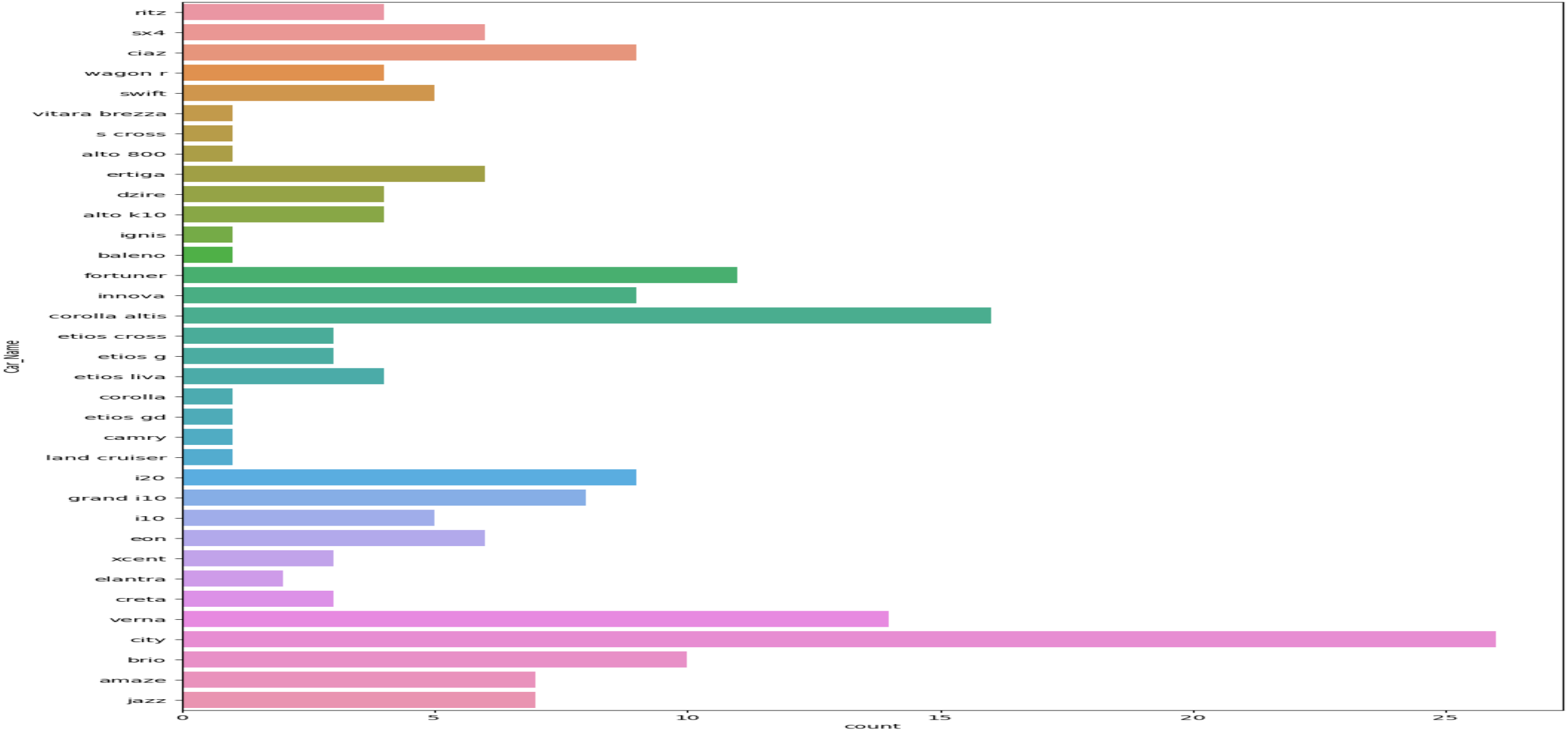




EDA RESULTS


Car_Name
Royal Enfield Thunder 500
UM Renegade Mojave
KTM RC200
Bajaj Dominar 400
Royal Enfield Classic 350
KTM RC390
Hyosung GT250R
Royal Enfield Thunder 350
KTM 390 Duke
Mahindra Mojo XT300
Bajaj Pulsar RS200
Royal Enfield Bullet 350
Royal Enfield Classic 500
Bajaj Avenger 220
Bajaj Avenger 150
Honda CB Hornet 160R
Yamaha FZ S V 2.0
Yamaha FZ 16
TVS Apache RTR 160
Bajaj Pulsar 150
Honda CBR 150
Hero Extreme
Bajaj Avenger 220 dtsi
Bajaj Avenger 150 street
Yamaha FZ v 2.0
Bajaj Pulsar NS 200
Bajaj Pulsar 220 F
TVS Apache RTR 180
Hero Passion X pro
Bajaj Pulsar NS 200
Yamaha Fazer
Honda Activa 4G
TVS Sport
Honda Dream Yuga
Bajaj Avenger Street 220
Hero Splender iSmart
Activa 3g
Hero Passion Pro
Honda CB Trigger
Yamaha FZ S
Bajaj Pulsar 135 LS
Activa 4g
Honda CB Unicorn
Hero Honda CBZ extreme
Honda Karizma
Honda Activa 125
TVS Jupyter
Hero Honda Passion Pro
Hero Splender Plus
Honda CB Shine
Bajaj Discover 100
Suzuki Access 125
TVS Wego
Honda CB twister
Hero Glamour
Hero Super Splendor
Bajaj Discover 125
Hero Hunk
Hero Ignitor Disc
Hero CBZ Xtreme
Bajaj ct 100
0
1
2
3
4
5
6
7
count

EDA RESULTS


ritz
sx4
ciaz
wagon r
swift
vitara brezza
s cross
alto 800
ertiga
dzire
alto k10
ignis
baleno
fortuner
innova
corolla altis
etios cross
etios g
etios liva
corolla
etios gd
camry
land cruiser
i20
grand i10
i10
eon
xcent
elantra
creta
verna
city
brio
amaze
jazz
Car_Name
0
5
10
15
20
25
count

Code Execution & Data Sets

Git Hub -

VenkataSatyaSrikarParvatha/VOIS (github.com)

- → Visit the link by - ctrl + click
- → It contains a reference data set taken from car deko
- → A Jupyter Notebook Executed python code for the reference of execution

References

Medium - Data Analysis on vehicle Dataset from Car-Dekho using Python Libraries.

Data Analysis on vehicle Dataset from Car-Dekho using Python Libraries. | by Vikrant Balwant Nikumbhe | Geek Culture | Medium

Vehicle dataset from cardekho - Kaggle Data Set.

Vehicle dataset from cardekho (kaggle.com)

Conclusion

This project will provide a detailed analysis of the Car Dekho dataset, offering valuable insights into the car market. By understanding the factors that influence car prices and identifying market trends, stakeholders can make more informed decisions, leading to better outcomes in the automotive industry.

Thank you