

TEAM G

IN ACTION

EVOASTRA VENTURES
MINI PROJECT



AGENDA

In this mini project, our agenda is to scrape detailed car listings from cars24.com to build a comprehensive dataset. We will focus on gathering information such as car name, model, kilometers driven, fuel type, transmission type, discount price, actual price, location, and discount offered. Following the data collection, we will proceed with data cleaning and preprocessing, visualize key insights, and prepare a presentation to highlight our findings. The goal is to unlock valuable insights that will drive strategic decisions and enhance our understanding of the market.

1 **ABOUT COMPANY**

2 **PROJECT**

3 **OUR TEAM**

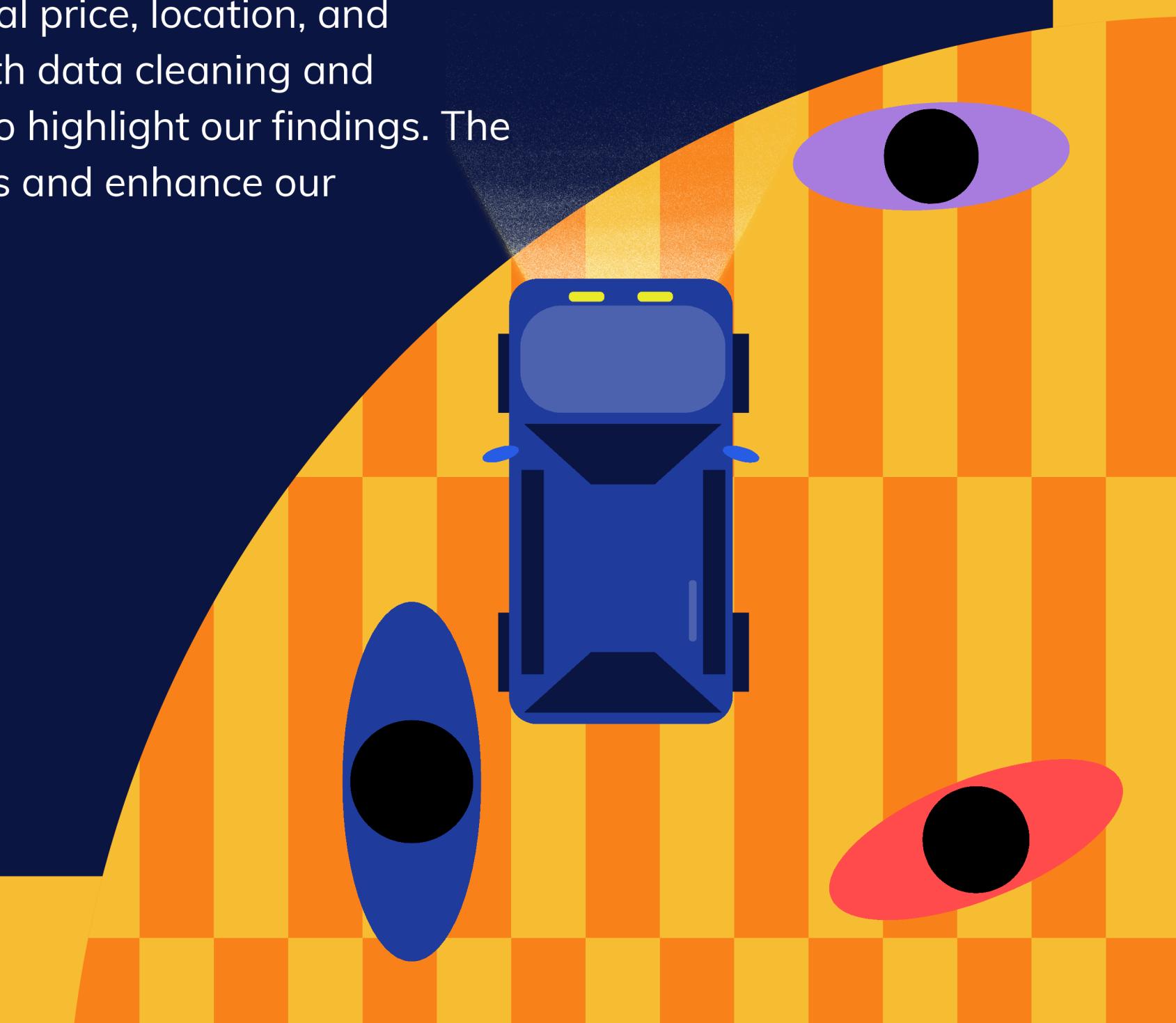
4 **WEB SCRAPPING**

5 **DATA WRANGLING**

6 **DATA VISUALIZATION**

7 **INSIGHTS**

8 **CONCLUSION**





EVOASTRA VENTURES

- EVOASTRA VENTURES IS A COMPANY SPECIALIZING IN DATA SCIENCE AND AI SERVICES, OFFERING CUSTOMIZED SOLUTIONS, AI STRATEGY CONSULTING, AND MACHINE LEARNING MODEL DEVELOPMENT.
- FOUNDED BY ANIKET MANWATKAR AND SYEDA MARIYA QUADRI IN 2023.
- EVOASTRA HAS COMPLETED OVER 356 PROJECTS AND SERVES MORE THAN 16 BRANDS AND ORGANIZATIONS.



* PROJECT

- CARS24 is a leading AutoTech company streamlining and revolutionizing the sale, purchase, and financing of pre-owned cars in India, Australia, Thailand, and UAE founded in 2015.
- Leveraging a Smart AI Pricing Engine, and 140 quality checks, selling and buying pre-owned vehicles is seamless and transparent with CARS24.

OBJECTIVE

To scrape detailed car listings from Cars24.com to build a comprehensive dataset for analysis

DETAILS TO SCRAPE

- Car Make
- Model
- Kilometers driven
- Fuel type
- Transmission type
- Price
- Location



OUR TEAM

TEAM LEAD

SUYASH CHADHA



SATWIK UPPADA



CO LEAD

ANKESH KUMAWAT



SANSKAR KESARI



MEMBERS

Sohan Kiran Patil

Uday Mahadev Sonar Pattar

Abhishek Kumar

Anant Anugrah Tudu

Sai eswar

Stella Maria Jose

Shivansh Jain

HomyaYaduma

Mariyambi nazar

Sanskrti Bordiya

Jai Vaishnav

Sanskars Kesari

Muhammed Kaif

* WEB SCRAPPING

Objective : To scrape detailed car listings from cars24.com, capturing various attributes including car name, year of registration, kilometers driven, fuel type, transmission type, discount price, actual price, discount offered, model, and test drive location.

Tools and Technologies

- Python
- Selenium
- BeautifulSoup
- Edge WebDriver
- Pandas
- Jupyter notebook

Scrapped Details

- Car brand (Maruti)
- Model (Alto)
- Registered year (2016)
- Km driven (97,698KM)
- Fuel type (petrol)
- Transmission type (Manual)
- Add-ons (low run cost)
- Discount price (2.31 lakh)
- Actual price (2.60 lakh)
- Discount offered (28.76k off)
- Test drive location (chhatarpur, Delhi)



* PROCESS OF SCRAPPING

INITIAL SCRAPPING ATTEMPT

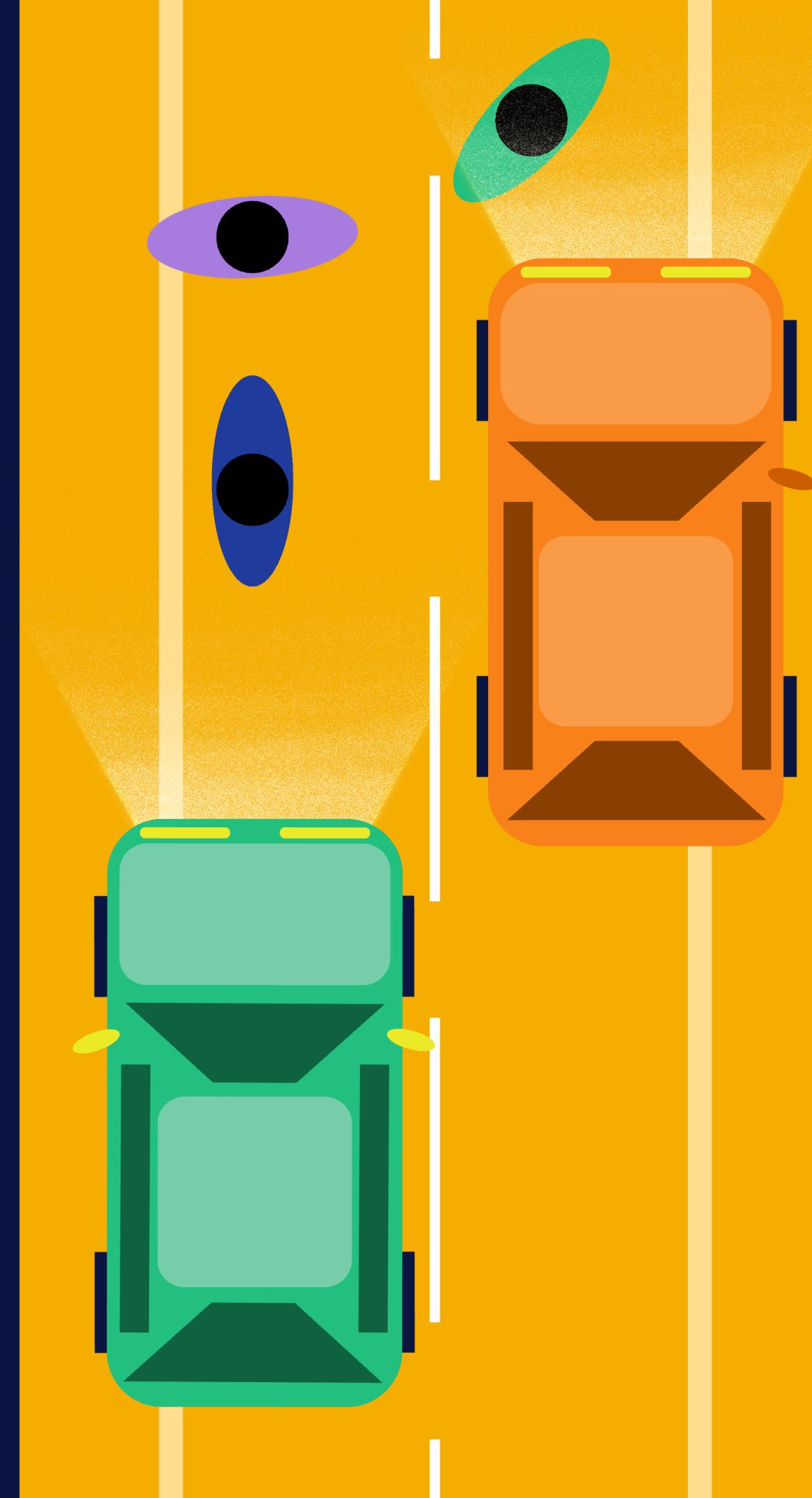
Issue: Only 20 entries were retrieved as the site loads additional content dynamically using JavaScript.

Diagnosis: Disabling JavaScript revealed that only a limited set of data is visible without JavaScript, confirming dynamic content loading.

SWITCH TO SELENIUM

Reason: To handle dynamic content loading and JavaScript-driven updates.

Approach: Implemented scrolling and refreshing to load all entries.



*CHALLENGES AND SOLUTIONS

Dynamic Content Loading

Challenge: The website uses JavaScript to dynamically load additional data.

Solution: Used Selenium to automate scrolling and reloading of the page to ensure all content is loaded and accessible.

Handling Multiple Data Points

Challenge: Different pieces of information were located in various parts of the HTML

Solution: Implemented a series of data extraction steps, each focusing on different attributes of the car listings.

Synchronization Issues

Challenge: Ensuring that data extraction steps are synchronized with page loading and content visibility.

Solution: Added delays and multiple page refreshes to make sure all content was loaded before extracting data.



* DATA CLEANING

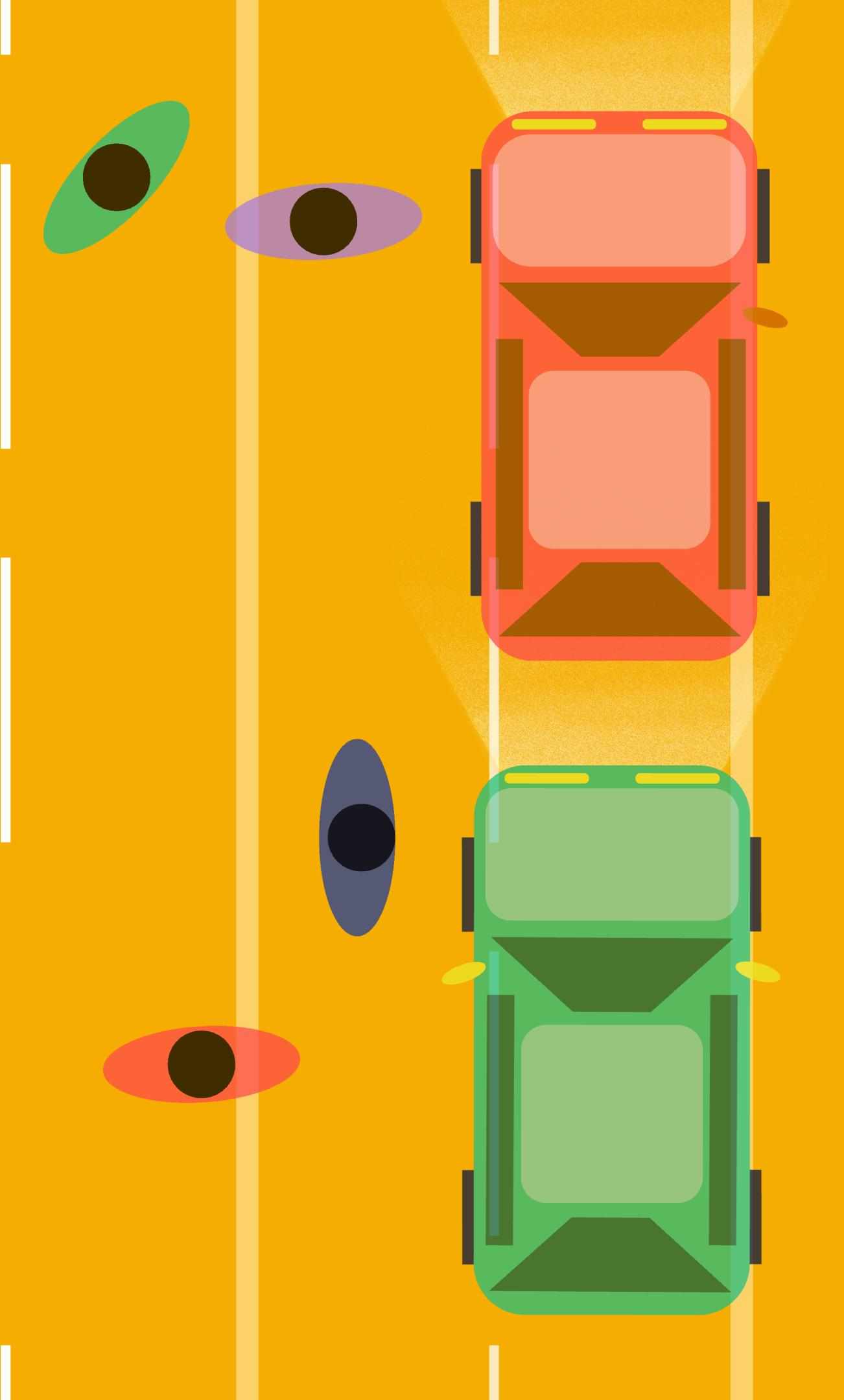
Objective : The goal of this project was to clean a dataset containing information about car sales. The dataset included columns such as "Km Driven," "Offer Price," "Original Price," "Discount Offered," and "Add ONs." The cleaning process involved handling missing values, converting data types, and standardizing the formats of various fields.

Technologies and Tools Used

- Python
- Pandas
- Numpy

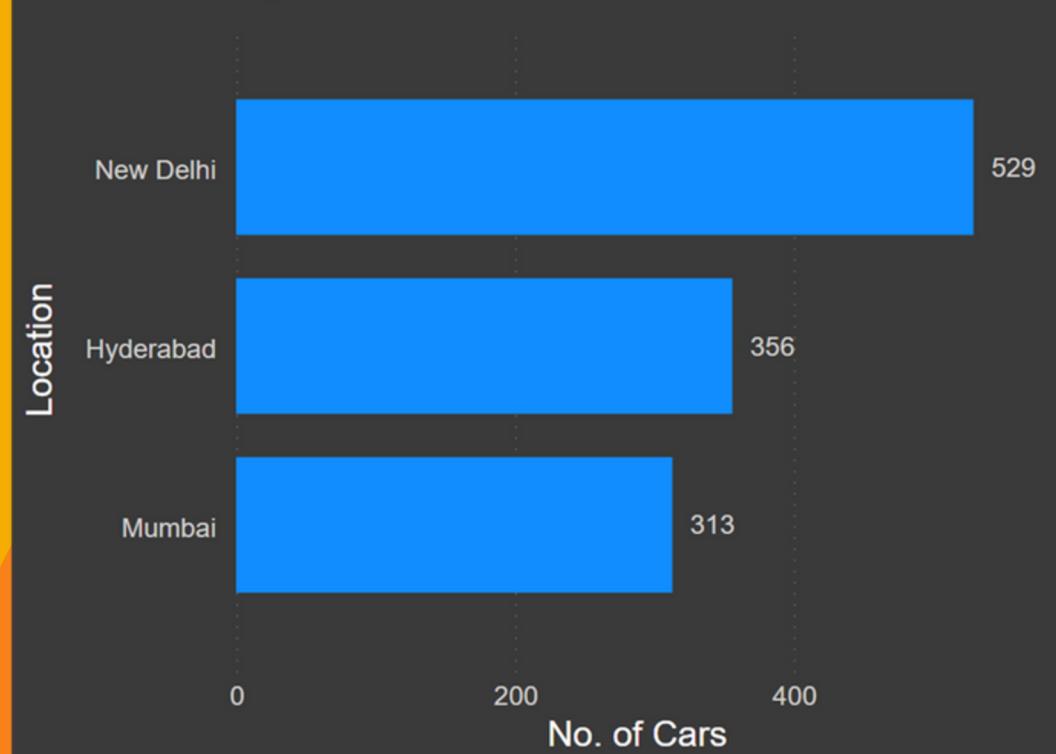
Steps Involved in Data Cleaning

1. Loading the Data
2. Cleaning Functions
3. Applying Cleaning Functions
4. Handling Missing Values
5. Saving the Cleaned Data

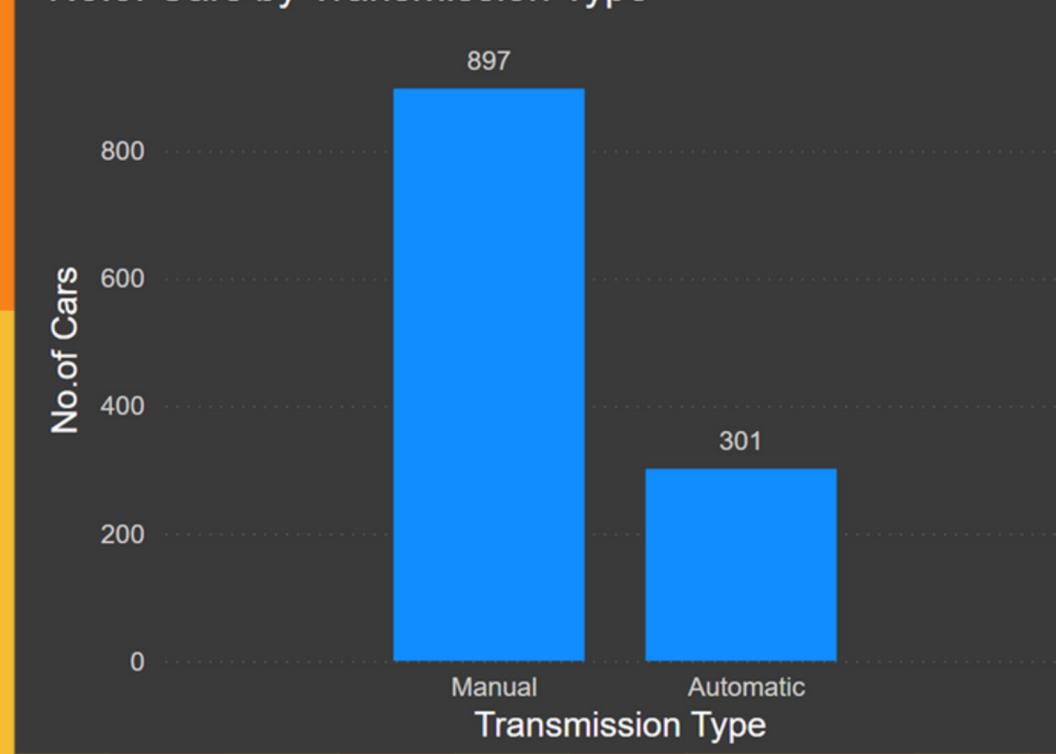


DATA VISUALIZATION

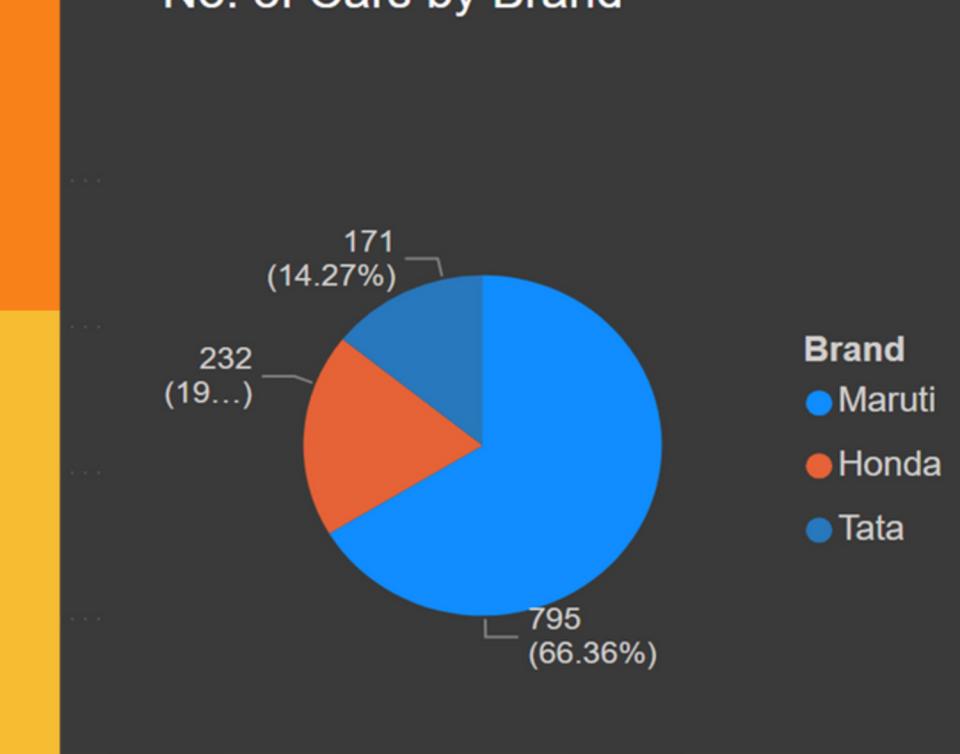
No. of Cars by Location



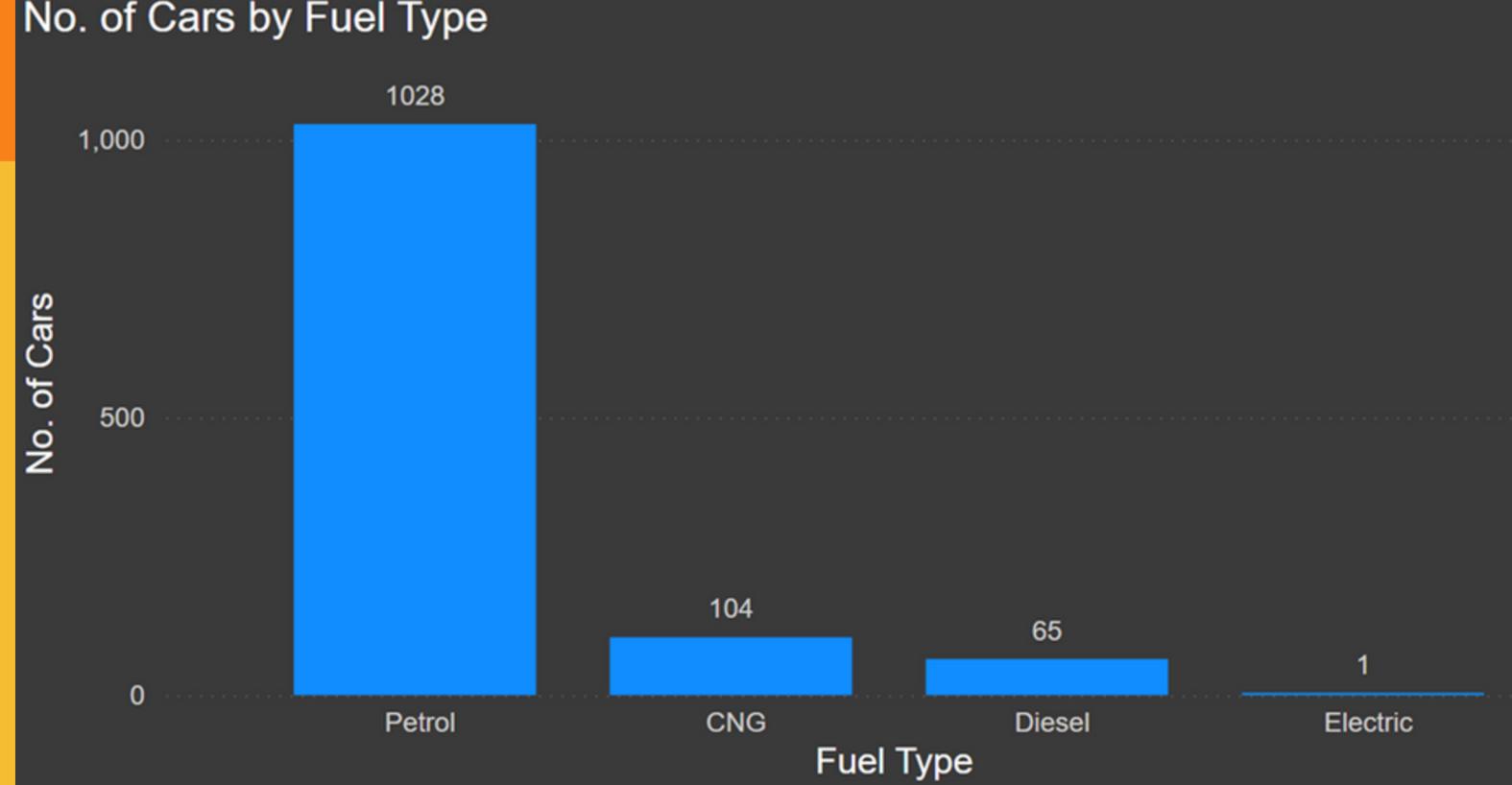
No. of Cars by Transmission Type



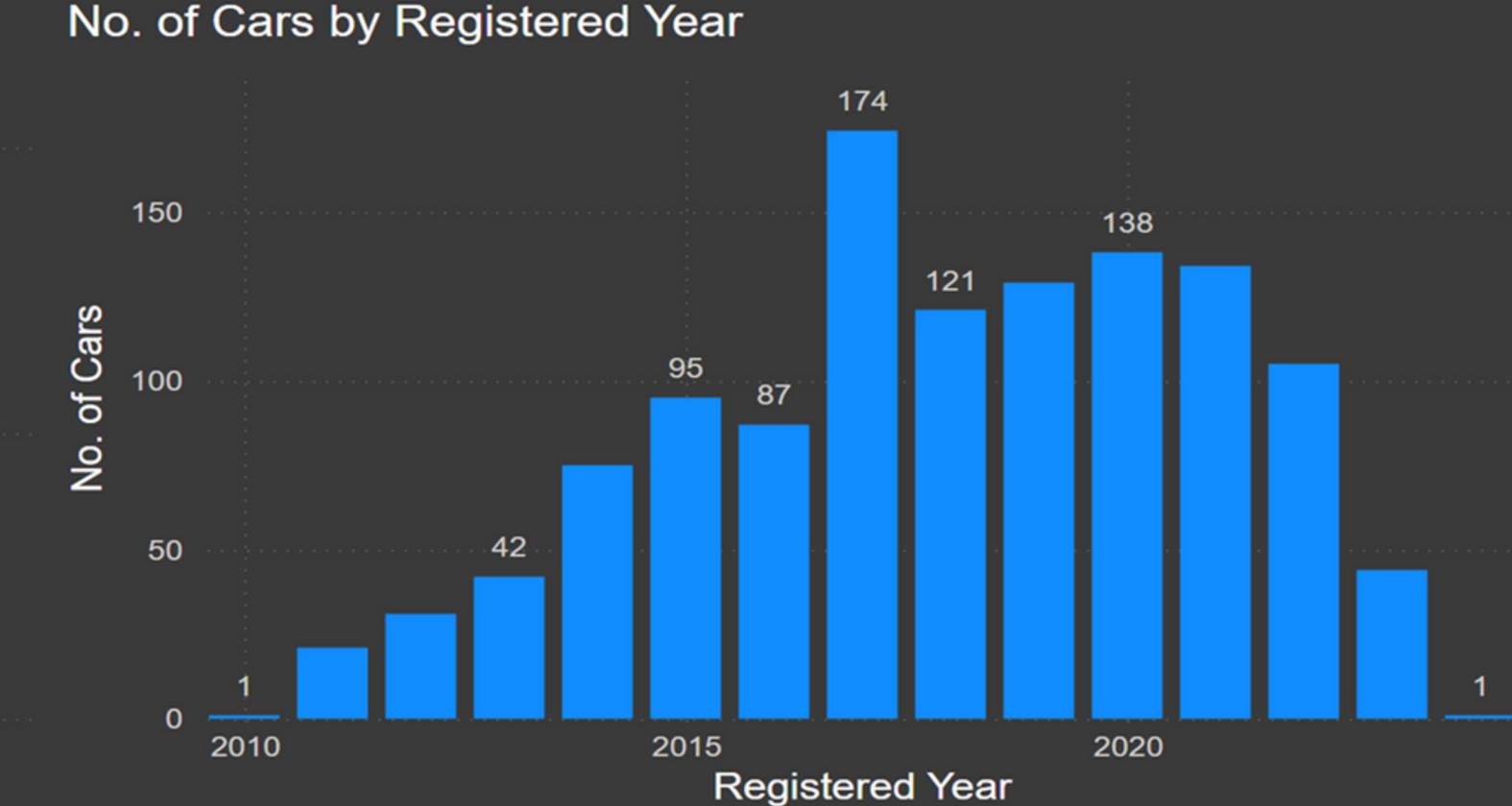
No. of Cars by Brand

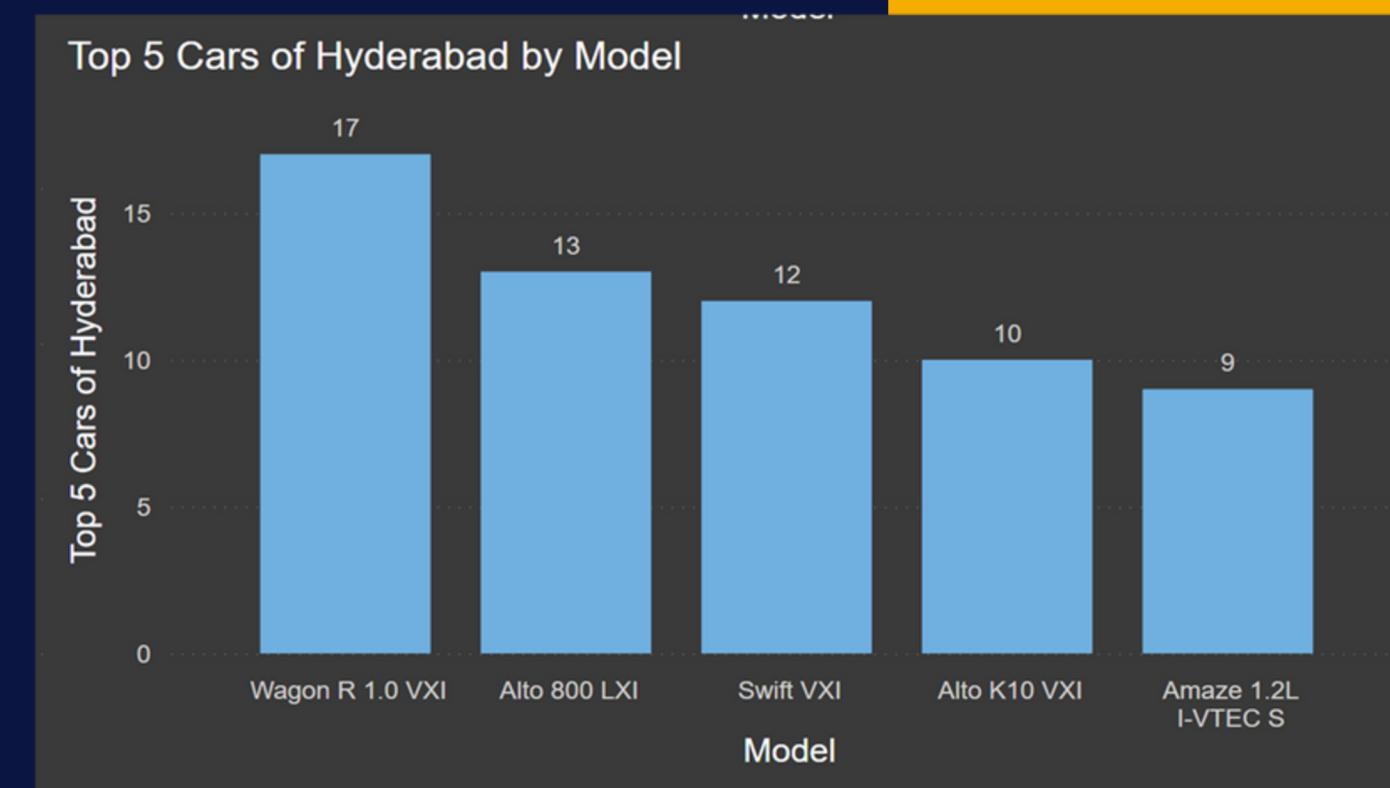
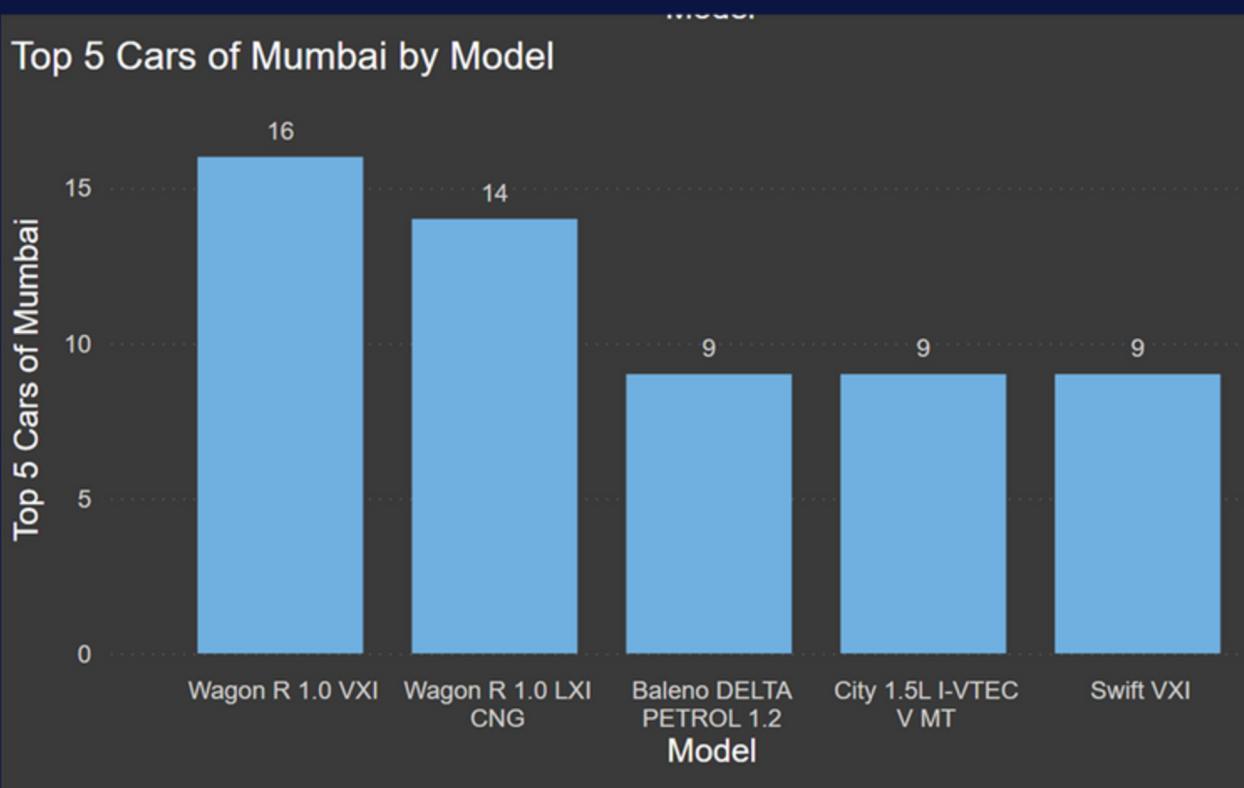
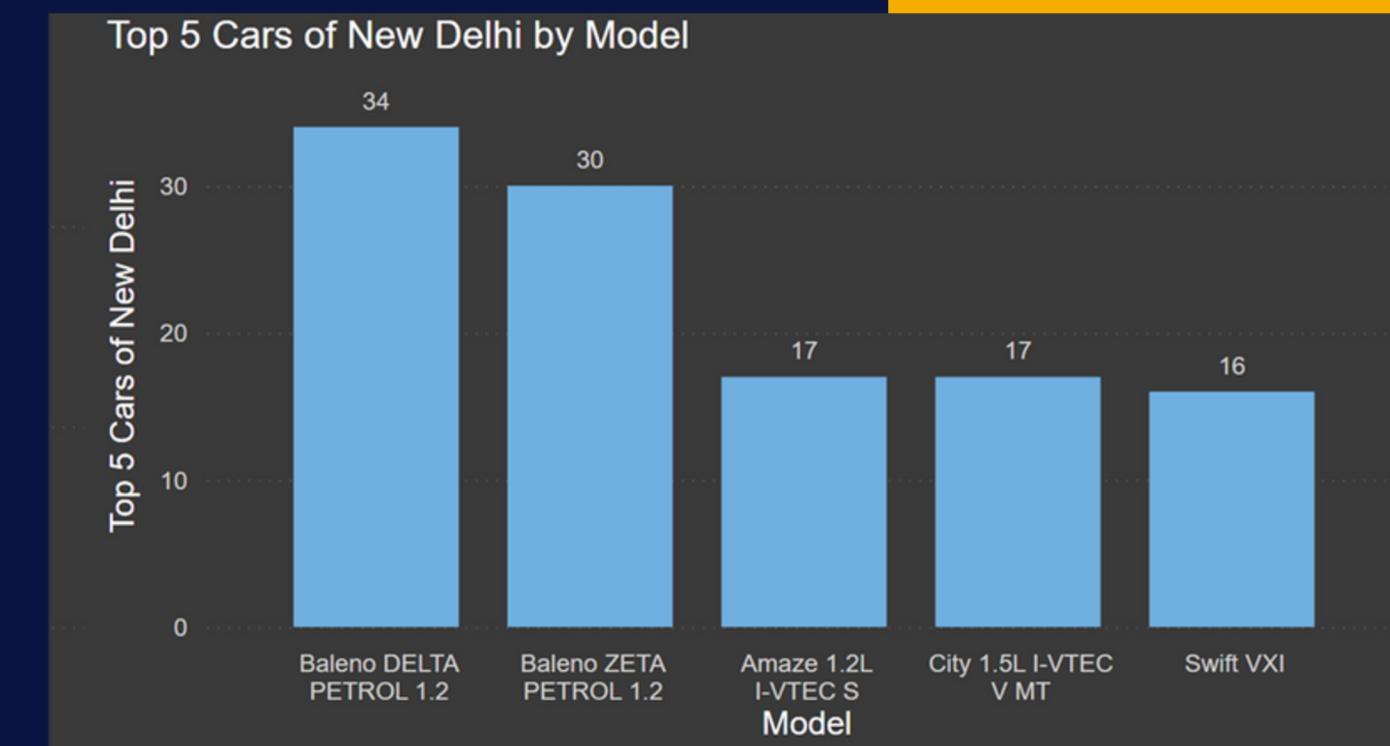


No. of Cars by Fuel Type



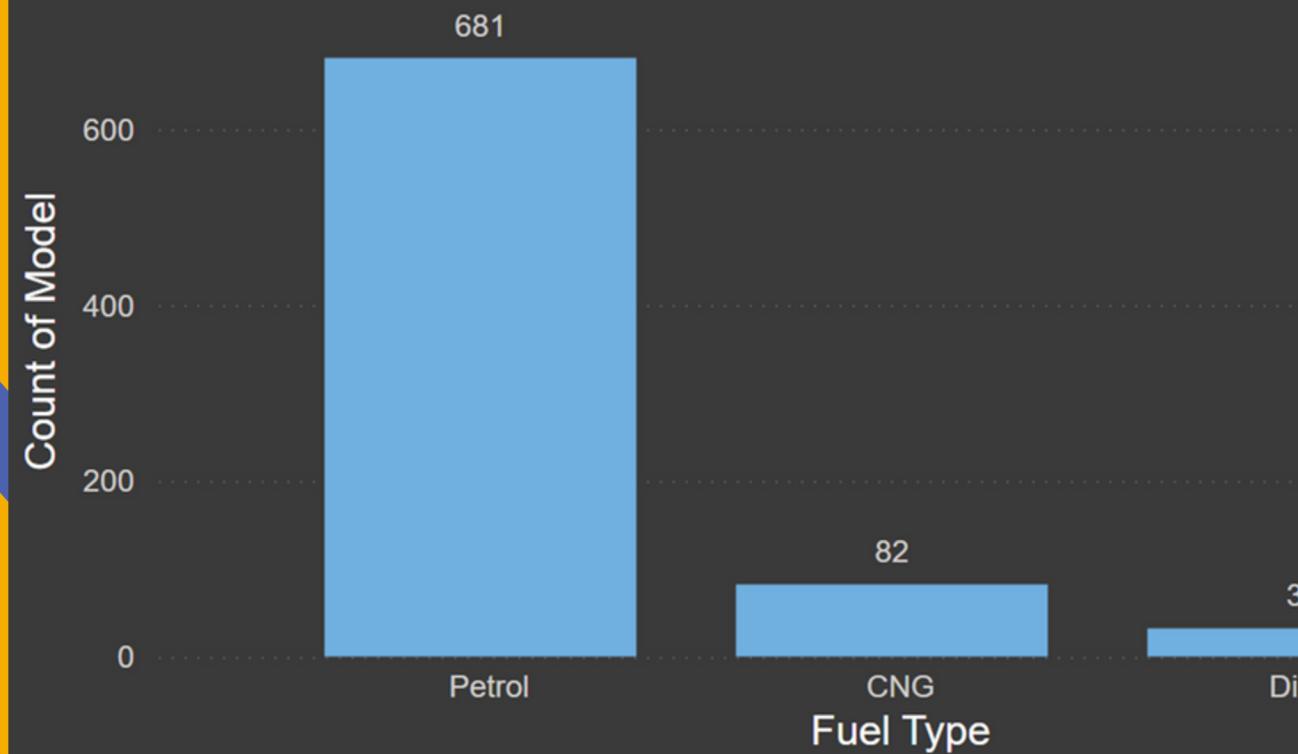
No. of Cars by Registered Year





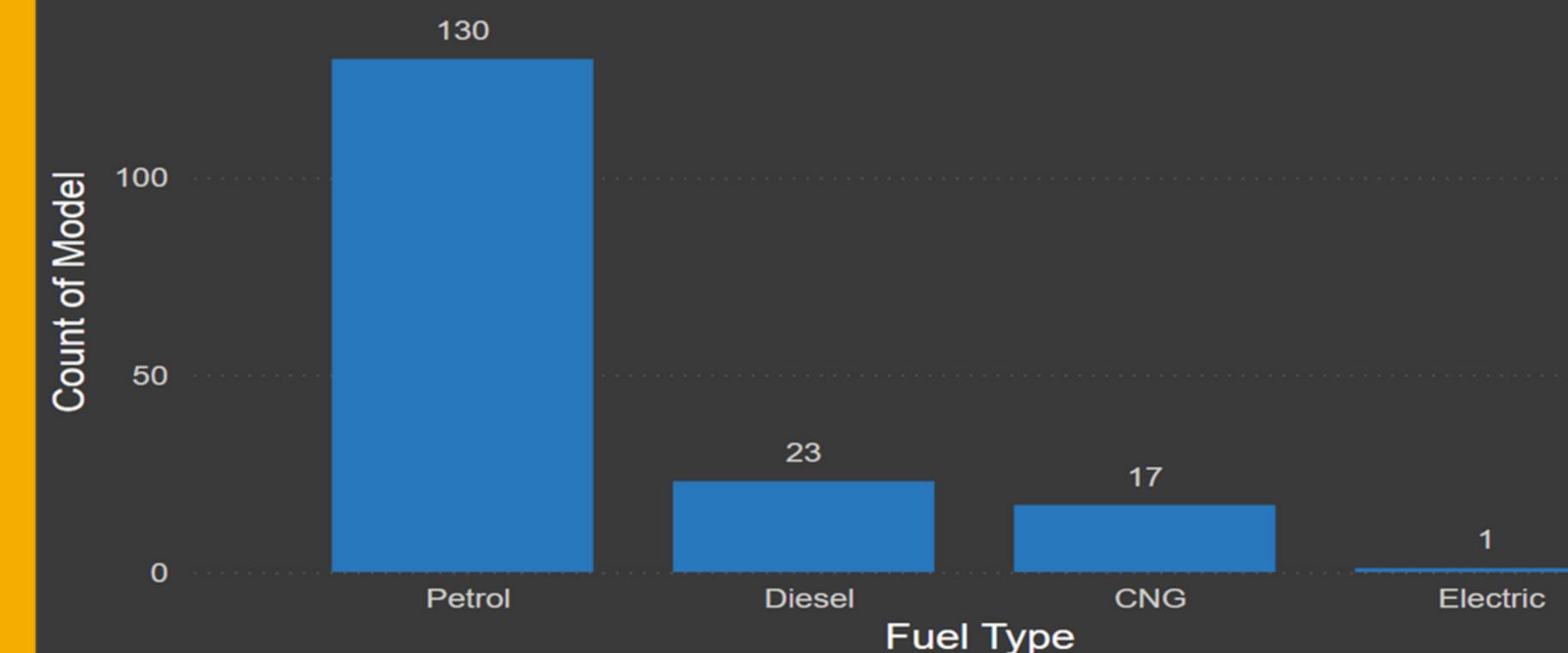
Count of Model by Fuel Type and Brand

Brand ● Maruti



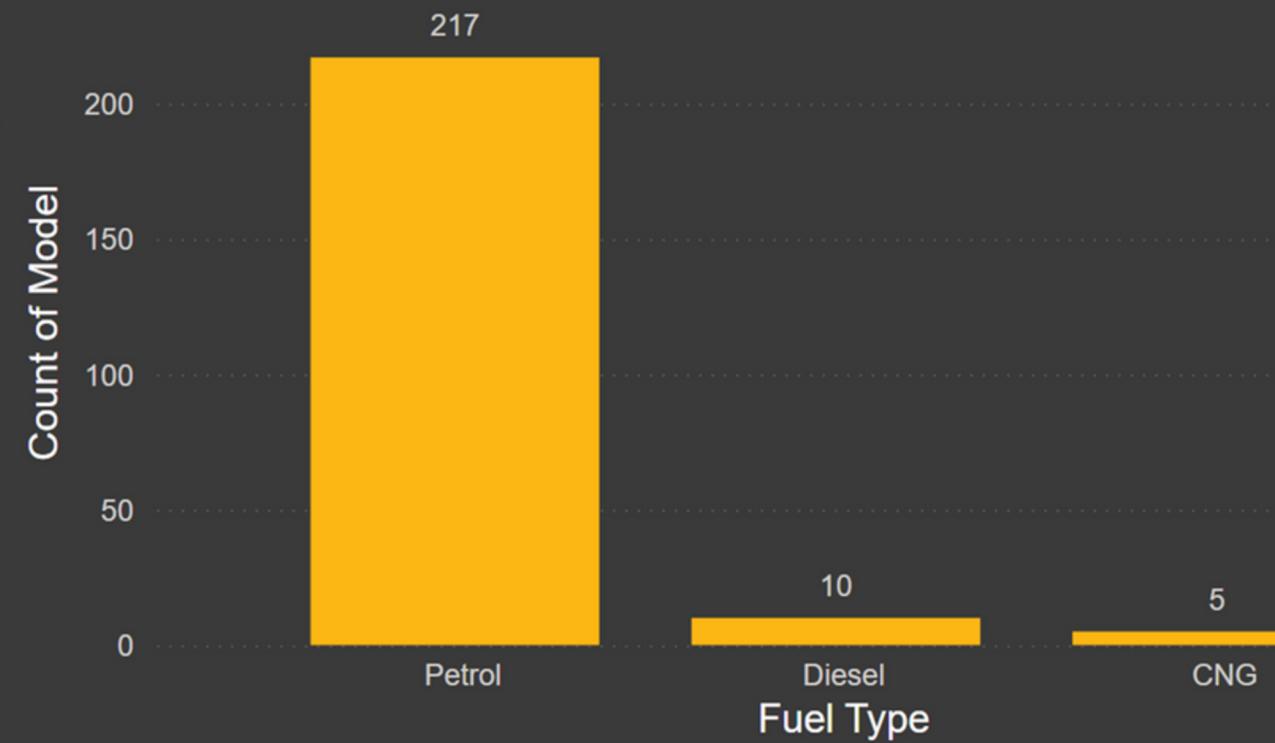
Count of Model by Fuel Type and Brand

Brand ● Tata



Count of Model by Fuel Type and Brand

Brand ● Honda



Sum of Km Driven by Fuel Type

50M

40M

30M

20M

10M

0M

0M

Petrol

0M

CNG

0M

Diesel

0M

Electric

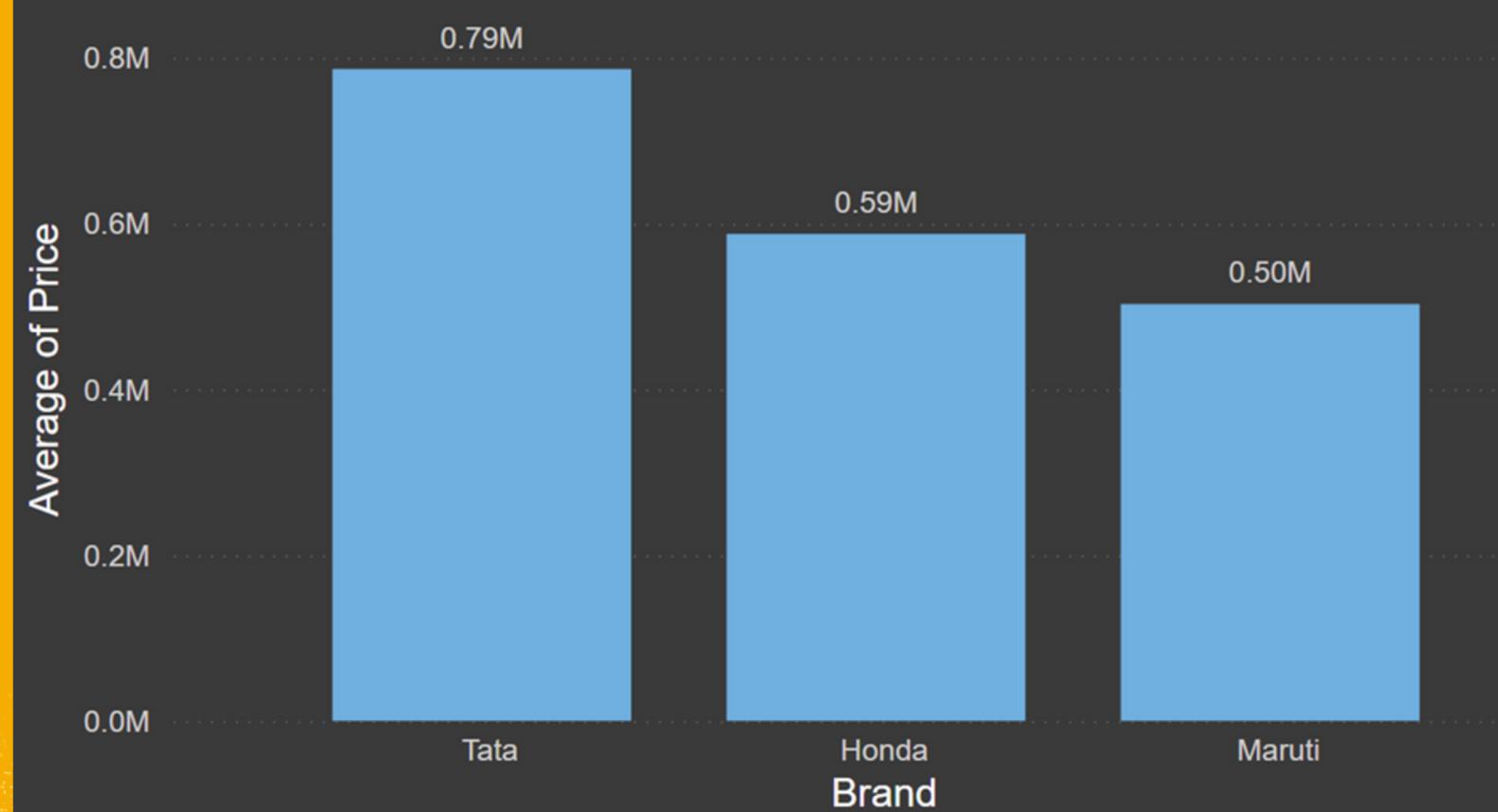
0M

Fuel Type

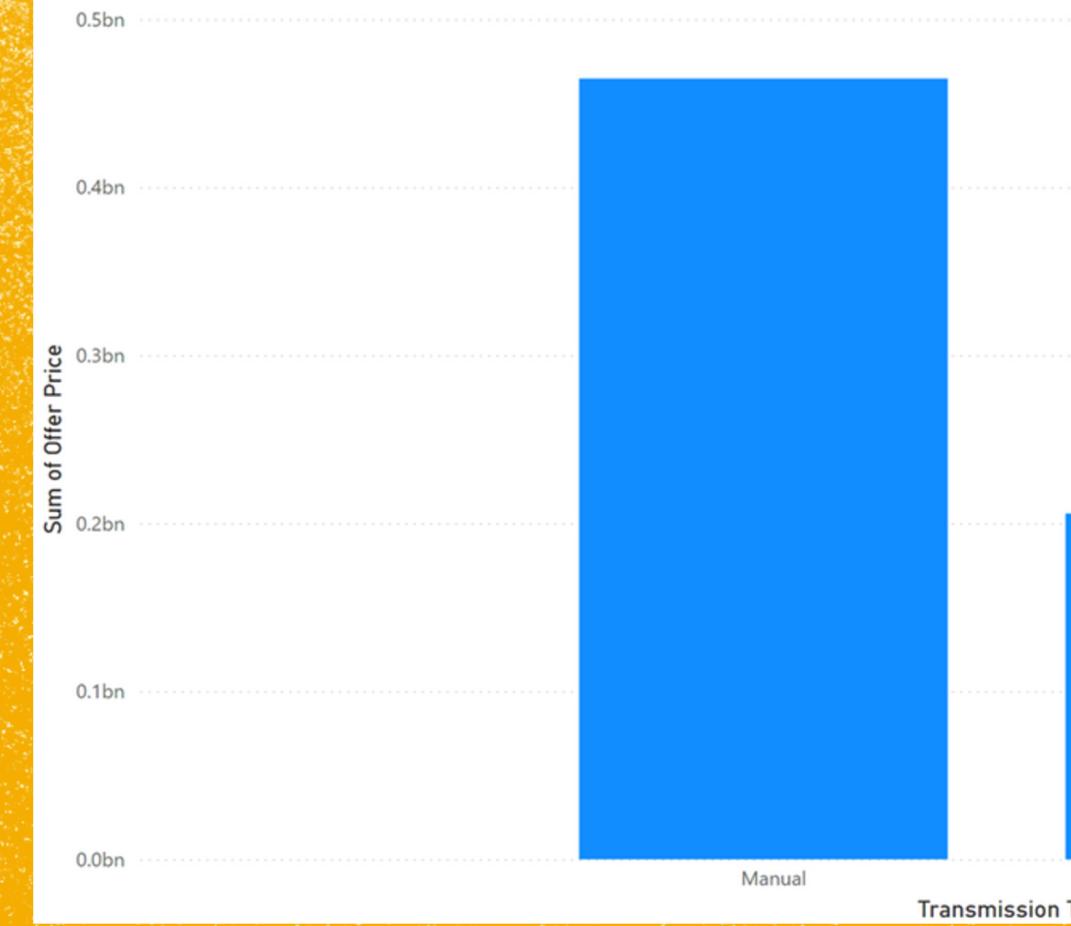
0M

Fuel Type

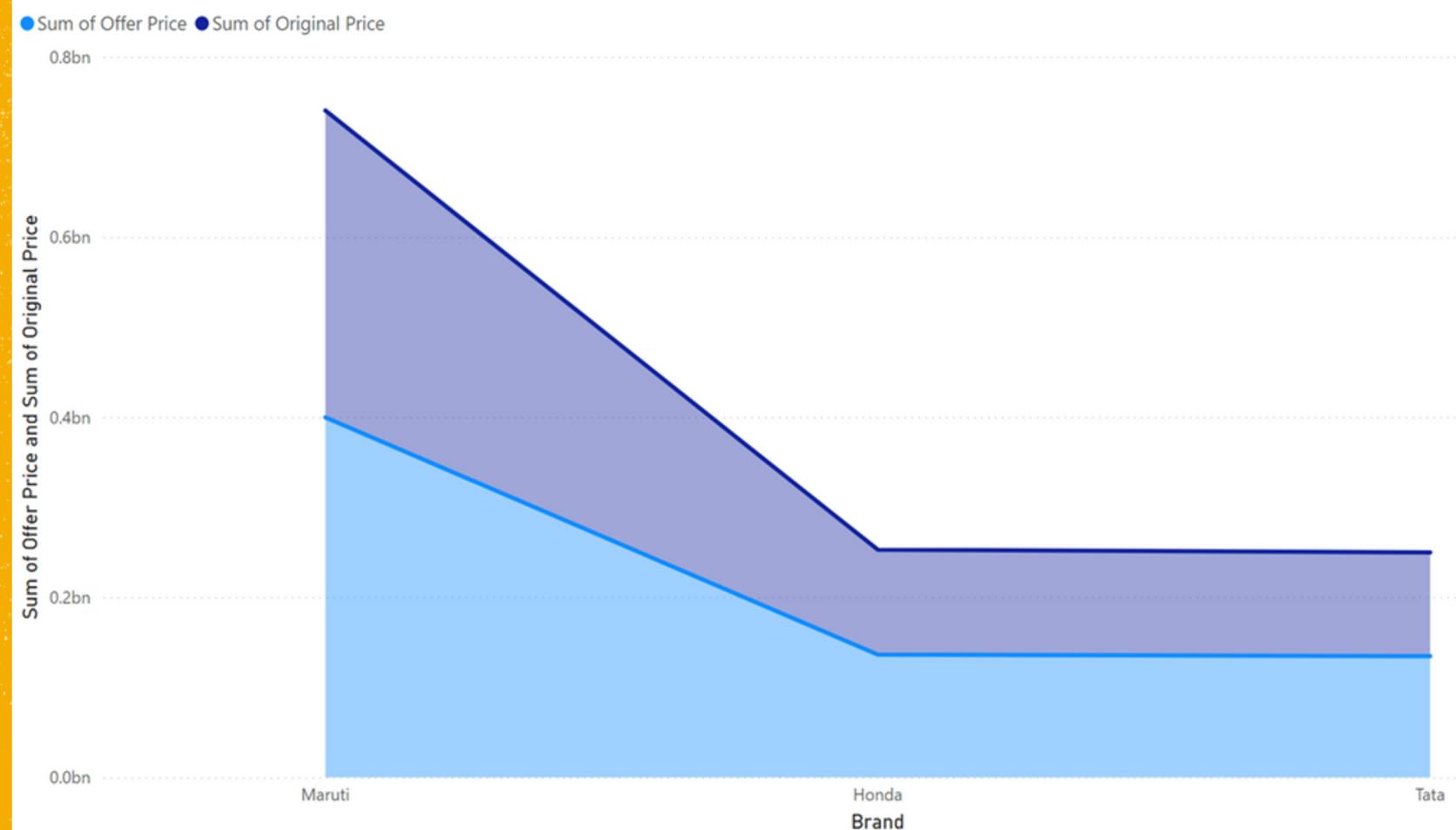
Average of Price by Brand



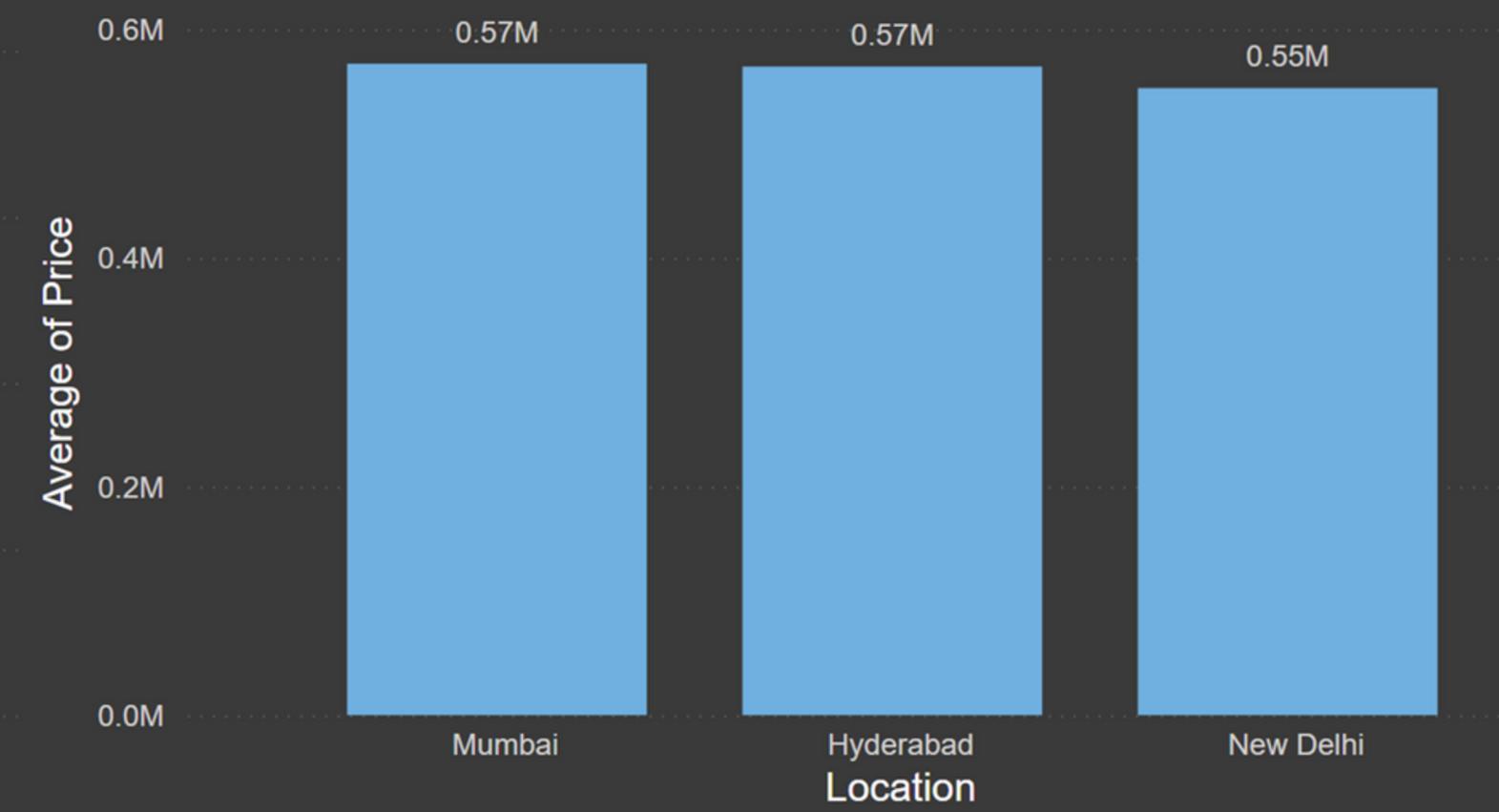
Sum of Offer Price by Transmission Type



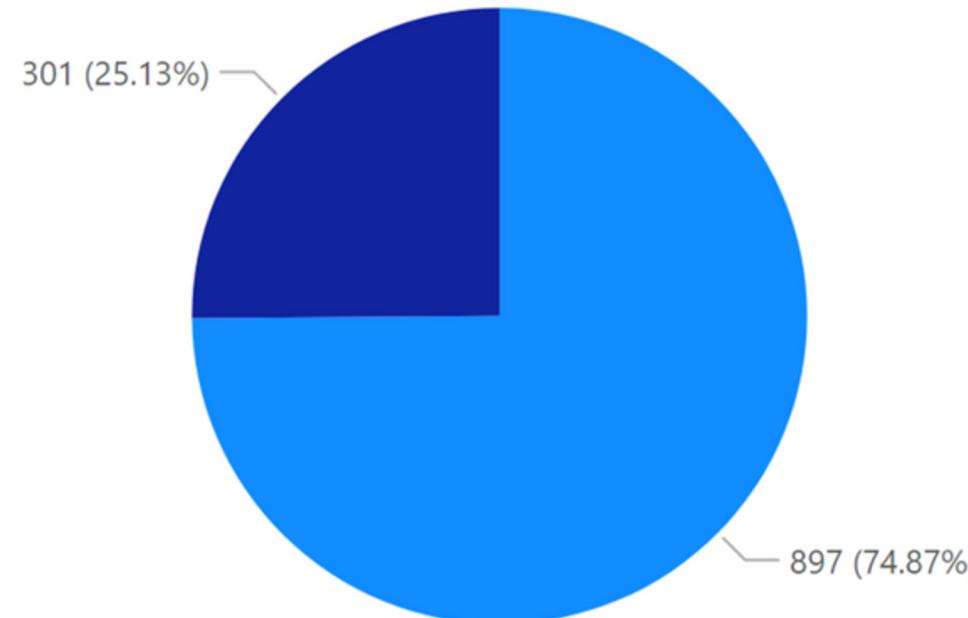
Sum of Offer Price and Sum of Original Price by Brand



Average of Price by Location



Count of Brand by Transmission Type

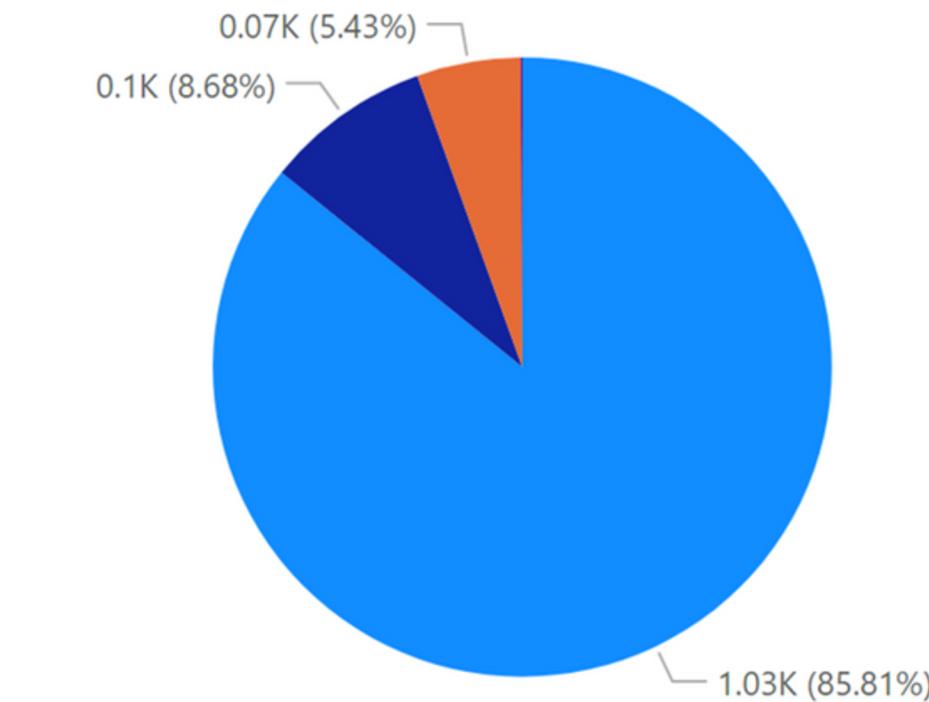


Transmission Type

● Manual

● Automatic

Count of Model by Fuel Type



Fuel Type

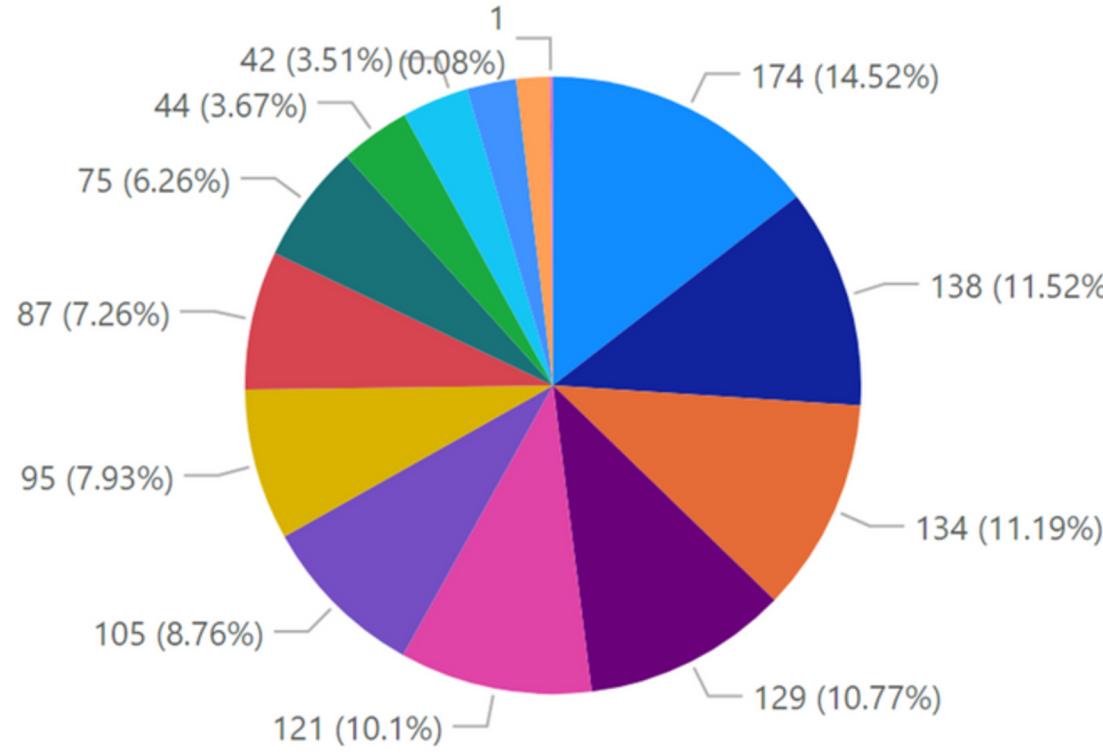
● Petrol

● CNG

● Diesel

● Electric

Count of Brand by Registered Year



Registered Year

● 2017

● 2020

● 2021

● 2019

● 2018

● 2022

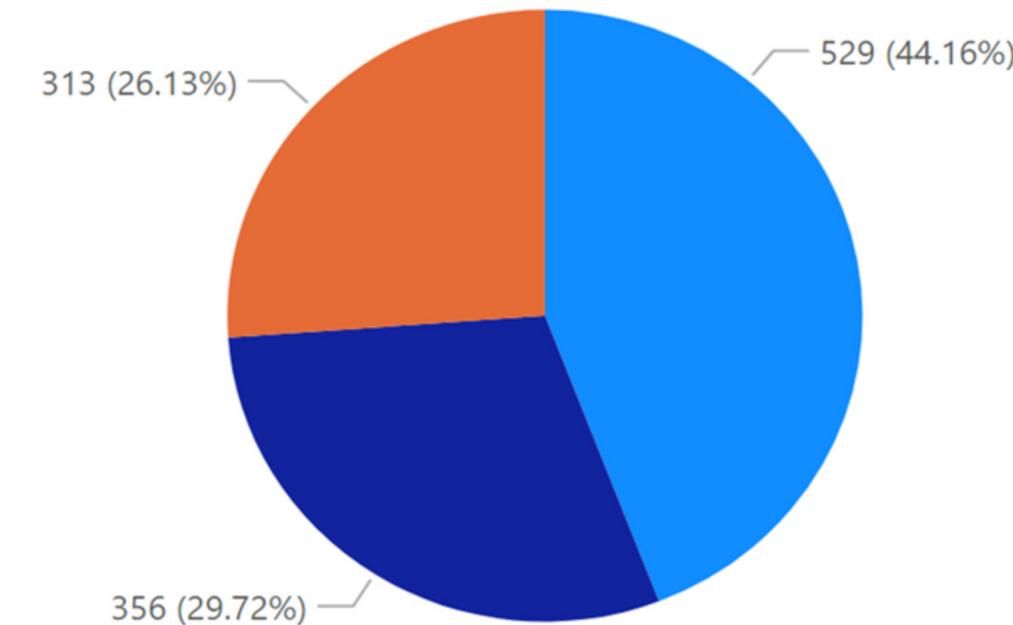
● 2015

● 2016

● 2014

● 2023

Count of Model by Location



Location

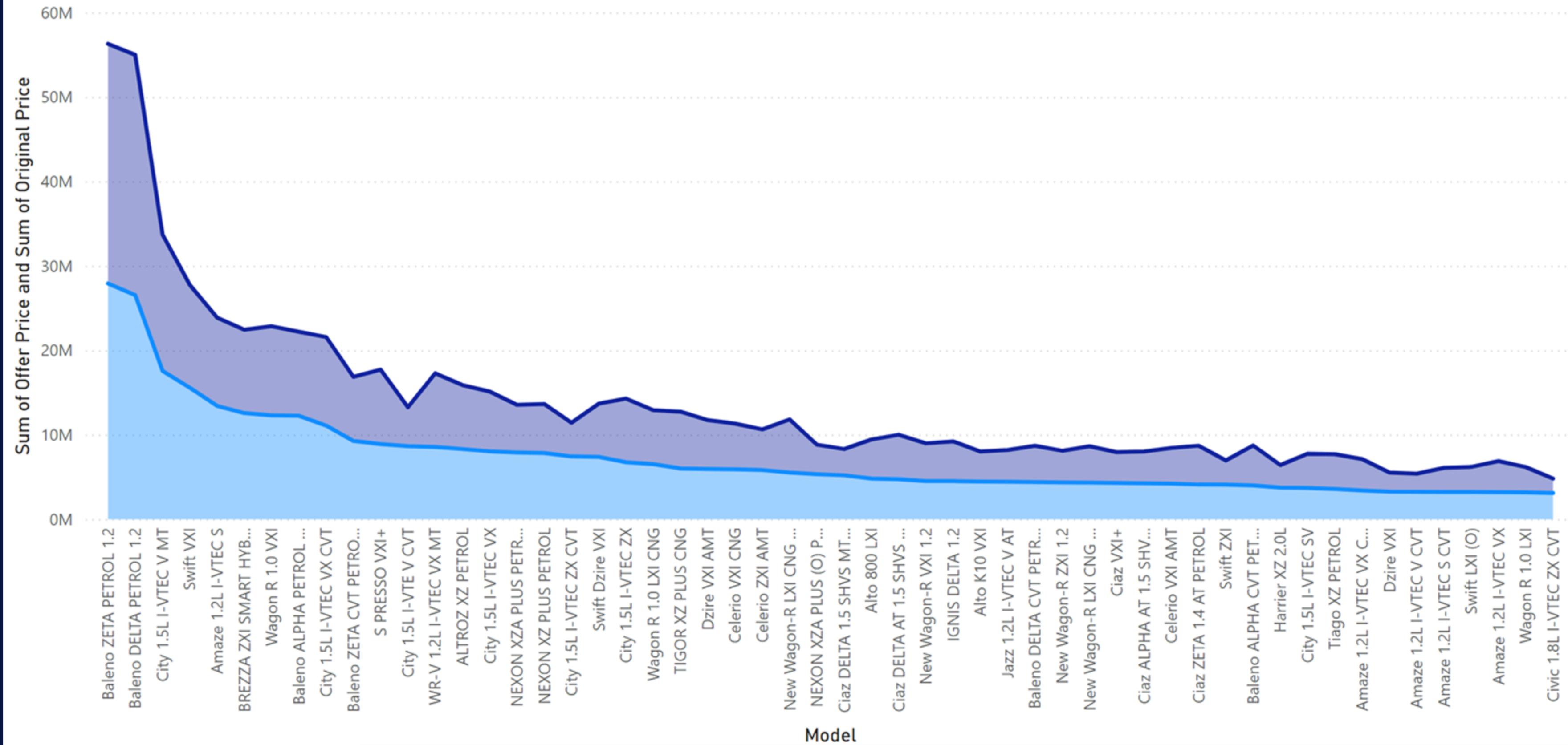
● New Delhi

● Hyderabad

● Mumbai

Sum of Offer Price and Sum of Original Price by Model

● Sum of Offer Price ● Sum of Original Price



*INSIGHTS

1. What is the average price of cars listed on Cars24.com?

Average Price: 502763.52

2. How does the average price vary by car brand?

MARUTI: 502763.522

HONDA: 587353.4483

TATA: 786345.0292

3. What is the distribution of cars by fuel type (e.g., petrol, diesel, electric)?

MARUTI: 795

CNG: 82

PETROL: 681

DIESEL: 32

HONDA: 232

CNG: 5

PETROL: 217

DIESEL: 10

TATA: 171

CNG: 17

PETROL: 130

DIESEL: 23

ELECTRIC: 01



* CONCLUSION

In this mini project, we successfully scraped and analyzed detailed car listings from cars24.com to build a comprehensive dataset. By extracting key data points such as car name, model, kilometers driven, fuel type, transmission type, discount price, actual price, and location, we were able to gain valuable insights into market trends and pricing strategies. The data cleaning and preprocessing steps ensured the accuracy and reliability of our findings. Through effective data visualization with Power BI, we presented actionable insights that can drive strategic decisions and enhance market understanding. This project demonstrates the power of data-driven analysis in uncovering trends and informing business strategies in the automotive sector.

