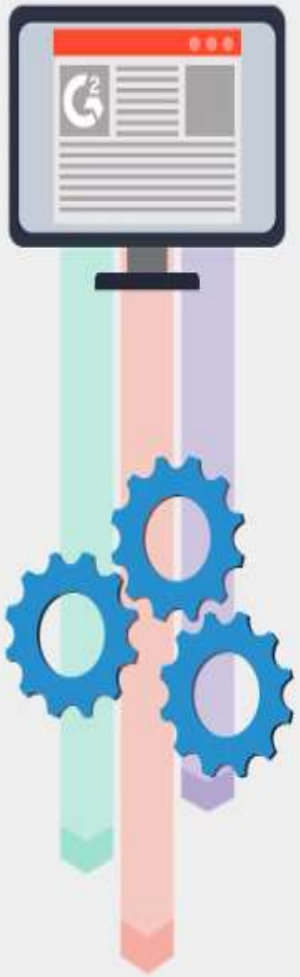
Driving Data Insights:

Scraping Tata Cars in Mumbai

Content

HTML website



Web scraping

Extracted data

Objectives

Our objectives are to collect and scrape detailed data on Tata cars listed for sale in Mumbai from Cars24.com, ensuring data accuracy through cleaning and processing and providing valuable information for potential buyers and sellers.

Data Collection

Website Used

Data was extracted from Cars24.com, a prominent platform for used car listings.

Key Data Points

Focused on extracting car model, price, year of manufacture, mileage, fuel type, transmission, and location.

Python Libraries Utilized

Employed Beautiful Soup for parsing HTML and Selenium for handling dynamic content.

Automated Scraping Process

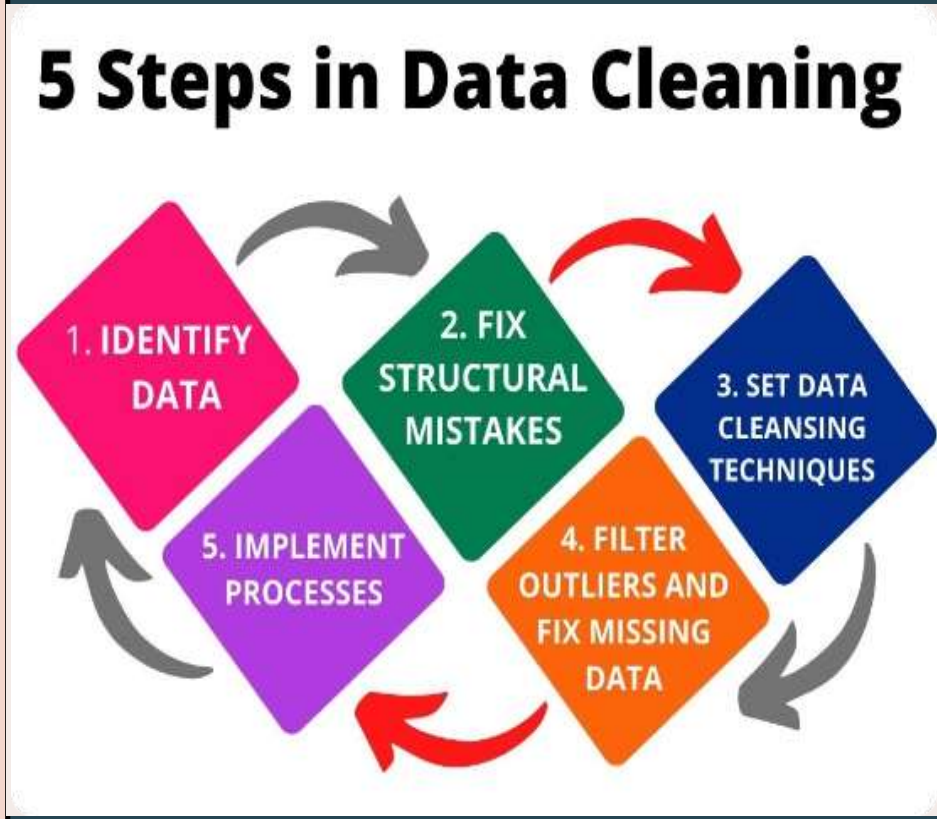
Developed Python scripts to automate the data extraction process, ensuring efficiency and thoroughness.

Data Accuracy

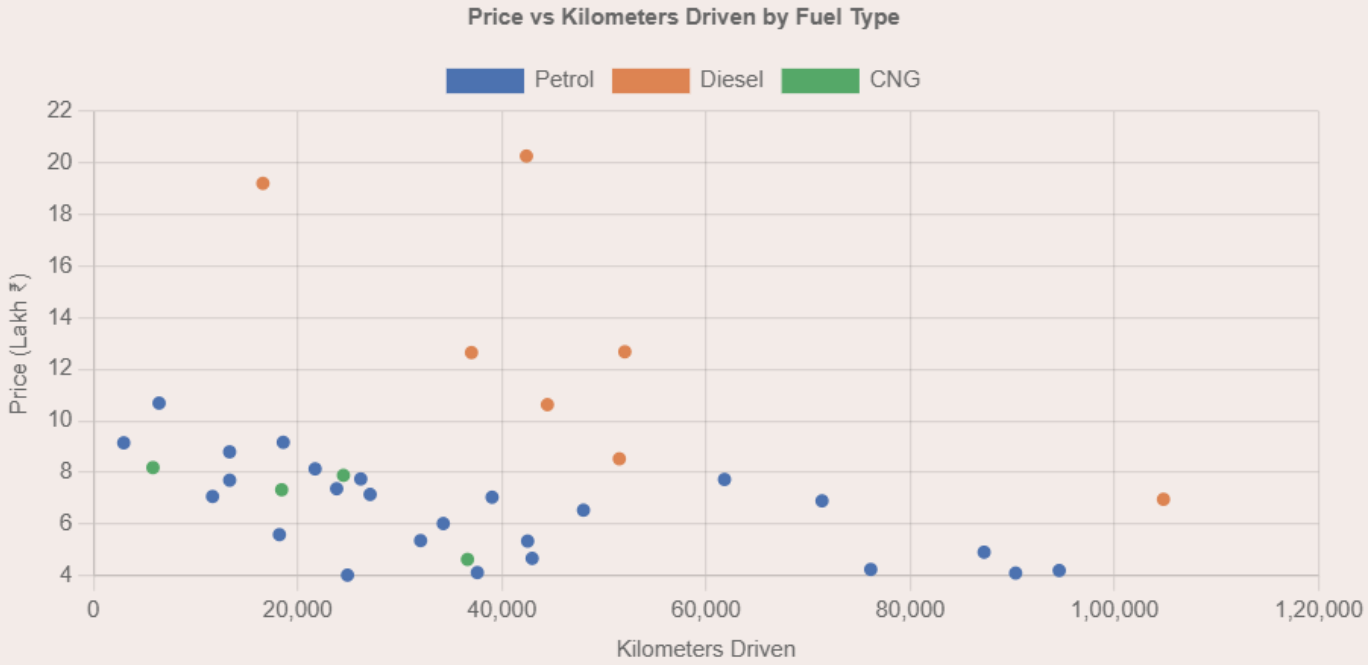
Implemented measures to verify the accuracy and completeness of the scraped data, enhancing dataset reliability

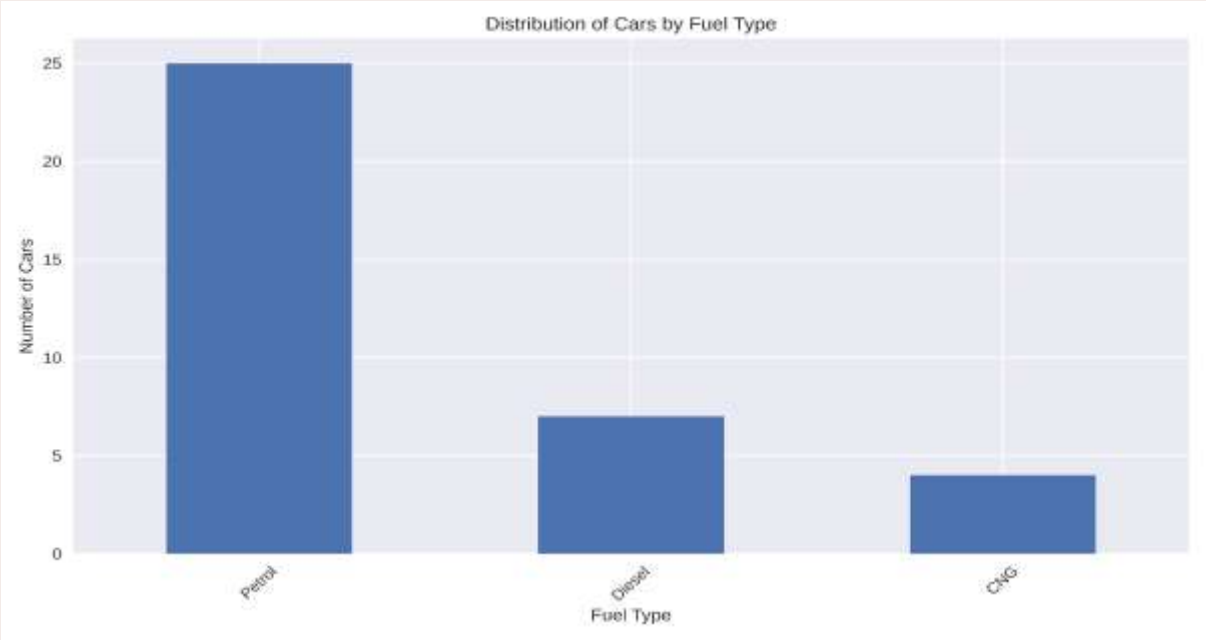
Data Cleaning

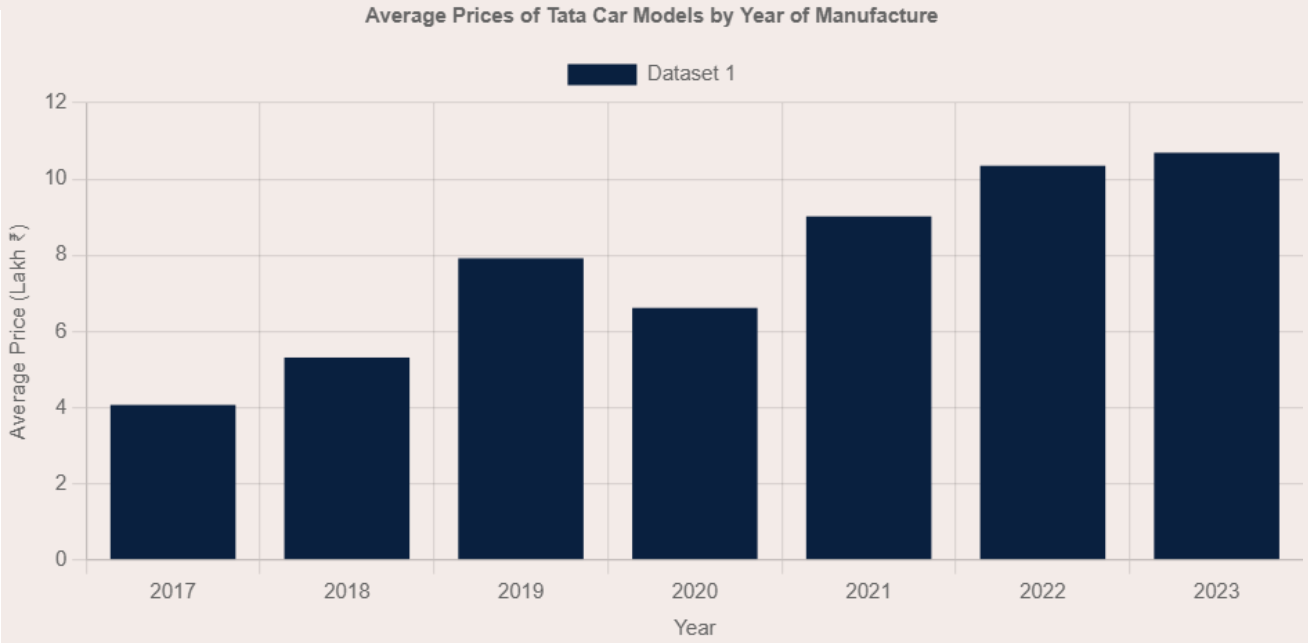
We used data cleaning to enhance our dataset by removing duplicate entries, filling or excluding missing values, standardizing data formats, correcting data types, and identifying outliers. These steps ensured our data was accurate, consistent, and reliable for analysis.



Data Analysis







Insights and Findings

Popular Models

Tata Tiago and Tata Tigor

Average Pricing

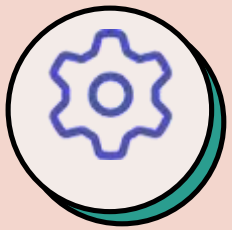
Manual - ₹7.34 lakhs , Automatic - ₹8.66 lakhs

Mileage Ranges

Analysis of the data shows a general trend of lower prices for cars with higher mileage

Fuel type preference

Petrol cars dominate the listings, with 25 petrol cars compared to 7 diesel cars and 4 CNG cars



Transmission Preference

The data reveals a significant preference for manual transmission cars

Challenges

- **Dynamic Web Content**

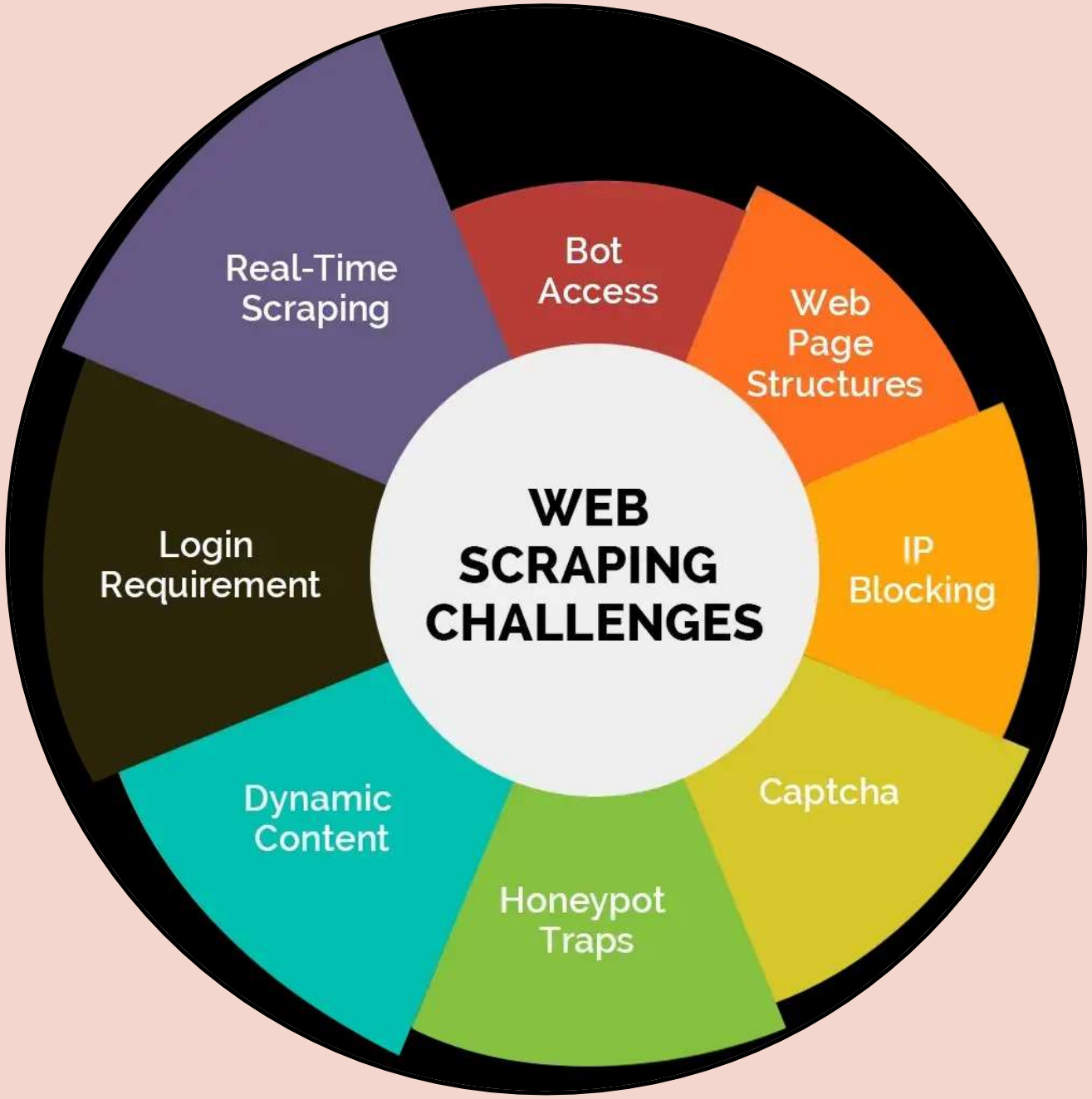
Adapting scraping techniques to handle dynamic web content

- **Data Accuracy**

Ensuring data accuracy and completeness

- **Missing Data**

Dealing with incomplete or missing data



Conclusion

- **Successful Data Collection:** Scraped detailed information on Tata cars in Mumbai from Cars24.com using Python libraries like Beautiful Soup and Selenium.

- **Key Insights:**
- Most popular models identified.
- Average pricing trends analyzed.
- Mileage ranges, fuel type, and transmission preferences determined.

- **Challenges Overcome:** Managed dynamic web content and ensured data accuracy.

- **Future Work:** Automate scraping for real-time data updates and continuous analysis.

- **Value:** Provides valuable insights for buyers and sellers in the Mumbai Tata car market.

Thank You

Thank you for your attention. We are now open to any questions or feedback you may have.