

WEB SCRAPING Cars24

TEAM - A

Presented by
Team lead - Namrata

Co-Team lead - Divya charitha

Evoastra Ventures -
[Home](#) | [EVOASTRA](#)



Discussion Overview

1. Fundamentals of web scraping
2. Significance of data extraction
3. Key objective of project
4. Tools and techniques
5. Challenges and resolving
6. Data collection and insights



OUR TEAM- A MEMBERS

- NAMRATA TODKAR (TEAM LEAD)
- DIVYA CHARITHA AMBALLA (CO TEAM LEAD)
- D LOKESH REDDY
- SARNALINA DATTA
- HINA SHAH
- MAYANK
- VISHNU VARDAN RAJU CHAMRATI
- RONAK SURESHBHAI SUTARIYA
- SHASHANK
- PRAKURTI



INTRODUCTION TO WEB SCRAPING

- Automated extraction of data from websites using scripts or software.
- Essential for making
 - Market analysis
 - Price comparison
 - Research in various industries
 - Data driven solutions



Overview of Cars24 Website

- Cars24 facilitates buying and selling of used cars.
- Sale , purchase and financing pre -owned cars
- Additional services such as warranties, doorstep delivery, making the process seamless and convenient.

Objective :

- To efficiently extract and analyze car listings
- pricing details ,model details
- Location based web scraping techniques



Web Scrapping Tools and Technologies

Programming language :

Python : Simple and powerful for Web Scrapping.

Libraries and Frame work :

BeautifulSoup : Easy HTML/XML parsing

Scrapy : High-Level Web Scrapping framework.

Development Environment :

Jupyter Notebook: Interactive testing and debugging.



METHODOLOGY

Saving Data
into file

Research and
Planning

Data
Extraction

Data
Cleaning

Initializing Selenium
WebDriver and
BeautifulSoup

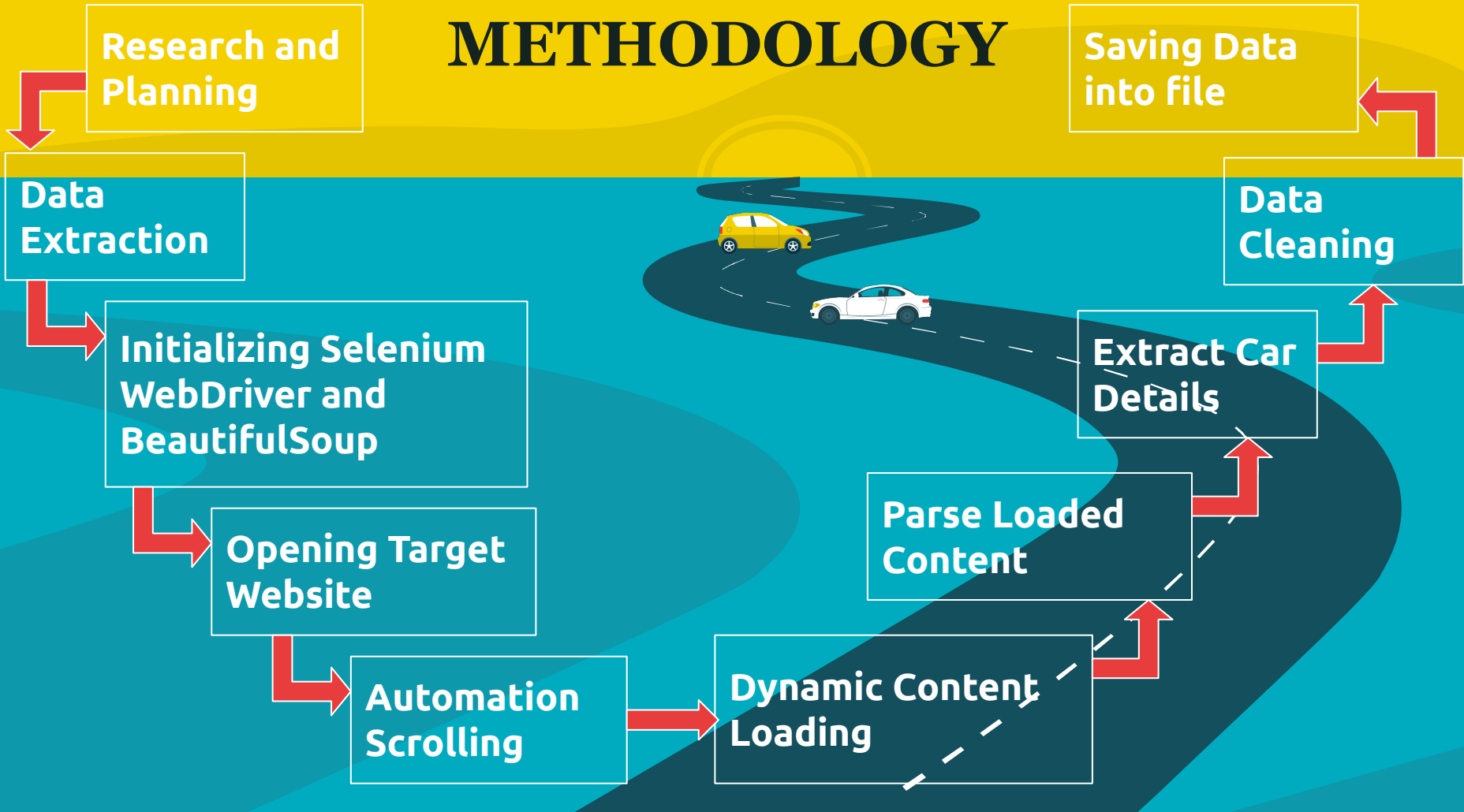
Extract Car
Details

Opening Target
Website

Parse Loaded
Content

Automation
Scrolling

Dynamic Content
Loading



Web Scraping Process Overview

Step-by-Step Process :

- Identifying the data to scrape
- Accessing the webpage to inspect it
- Extracting relevant data
- Storing the data in a structured format (CSV, JSON, etc.)
- Handling anti-scraping measures (e.g., CAPTCHAs, IP blocking)

Extracting Data from Cars24

Target Data

- Car make
- Model
- Year
- Price
- Location
- Fuel type
- Kilometer Driven
- Transmission
- Wheel type
- EMI per month

Using BeautifulSoup

- Parsing HTML content
- Extracting specific elements using tags, classes, and IDs

Data Storage

Options for Storing Data

- CSV files ideal for small tabular data
- Simple and lightweight, supported by most data analysis tools.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		YEAR	MAKE	MODEL	FUEL_TYPE	PRICE	KM_DRIVE	TRANSMIS	WHEEL_TY	EMI_PER_	LOCATION				
2	0	2015	Hyundai	Creta S	Diesel	8.5	61,718	Manual	Alloy whee	18,481	Garuda Mall, Magrath Road, Bengaluru				
3	1	2013	Honda	City 1.5L	Petrol	4.64	84,818	Manual	Alloy whee	12,219	Whitefield Main Road, Boruka Tech Park, Benga...				
4	2	2015	Renault	Duster 110	Diesel	5.74	94,510	Manual	Reg. servic	12,768	Bellahalli, Bengaluru				
5	3	2021	Tata	PUNCH AC	Petrol	8.49	8,945	Automatic	Top Mode	16,160	Garuda Mall, Magrath Road, Bengaluru				
6	4	2016	Ford	Ecosport T	Petrol	6.16	71,617	Manual	Spl. reg. no	12,043	Garuda Mall, Magrath Road, Bengaluru				
7	5	2017	Renault	Duster 85	Diesel	7.75	71,918	Manual	Reg. servic	15,151	Garuda Mall, Magrath Road, Bengaluru				
8	6	2015	Mahindra	XUV500 W	Diesel	7.85	98,338	Manual	Alloy whee	17,462	Bellahalli, Bengaluru				
9	7	2019	Toyota	Glanza G	Petrol	6.81	53,435	Manual	Alloy whee	13,314	Bellahalli, Bengaluru				
10	8	2020	Tata	ALTROZ XZ	Petrol	7.35	11,336	Manual	Low run co	14,370	Bellahalli, Bengaluru				
11	9	2017	Datsun	Redi Go	Petrol	2.62	43,814	Manual	Alloy whee	5,122	Bellahalli, Bengaluru				
12	10	2018	Maruti	Baleno RS	Petrol	6.54	61,005	Manual	Alloy whee	12,786	Whitefield Main Road, Boruka Tech Park, Benga...				



Data PreProcessing and Cleaning

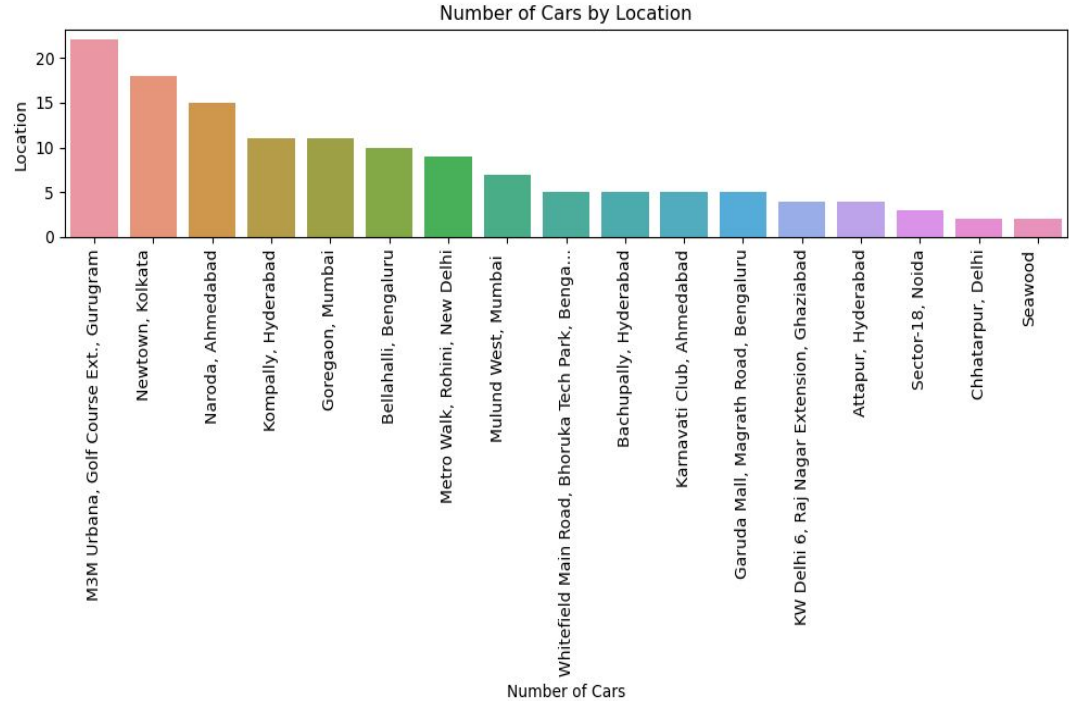
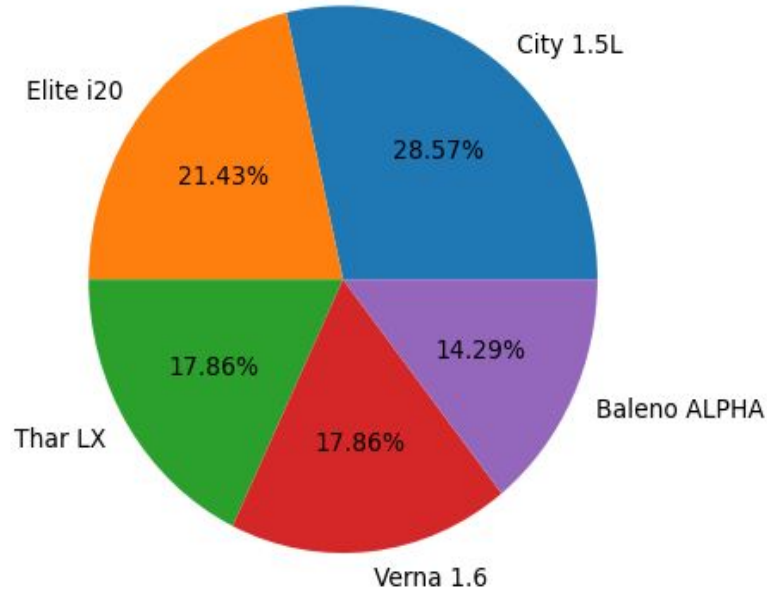
- **Importance of Clean Data**
 - Removing duplicates, null values, handling missing values.
- **Tools and Techniques**
 - Pandas for data manipulation



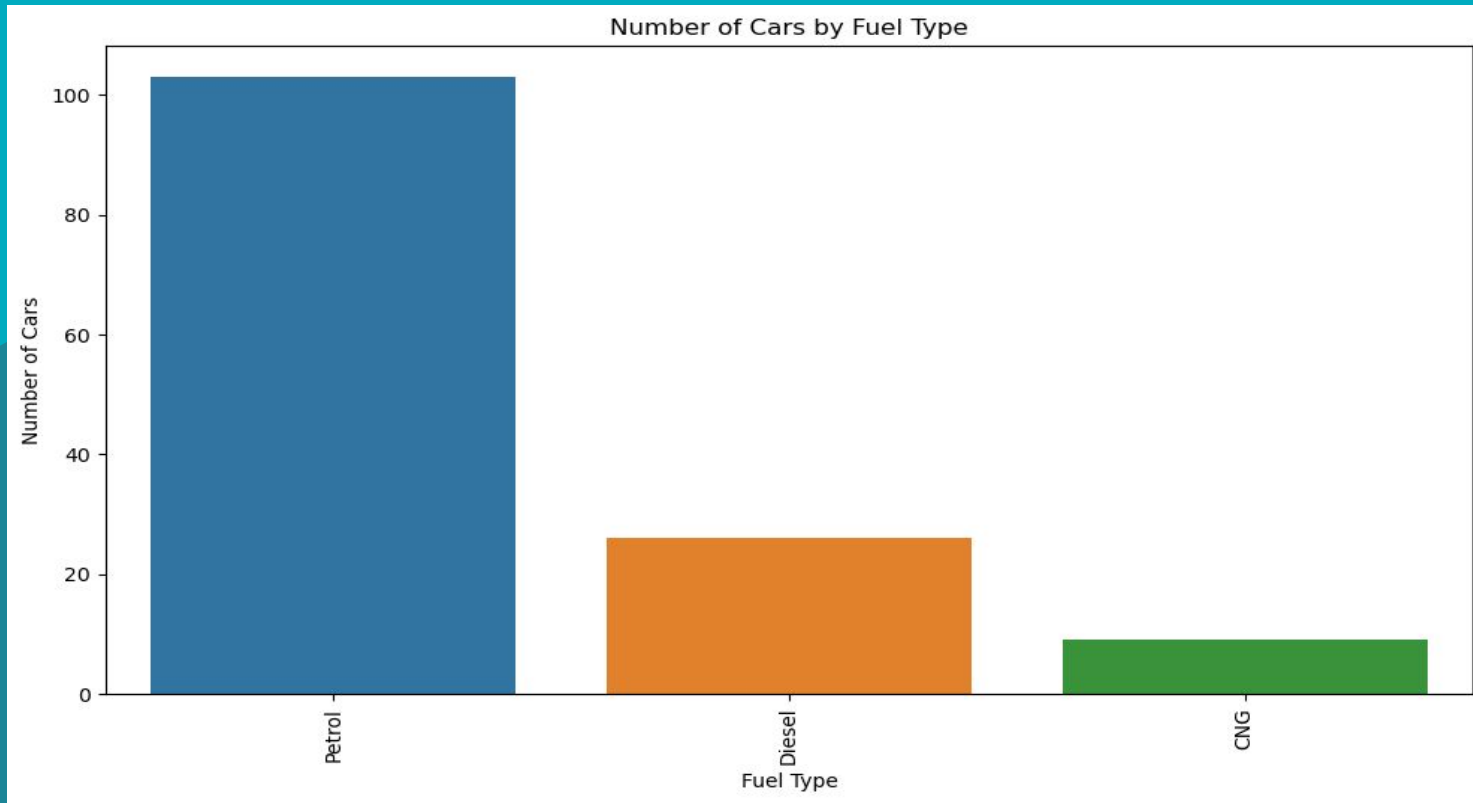
Data Analysis and Visualization

- **Basic Analysis**
 - Price trends
 - popular models
 - Usage based on fuel type, transmission , location
- **Visualization Tools**
 - Matplotlib
 - Seaborn for data visualization

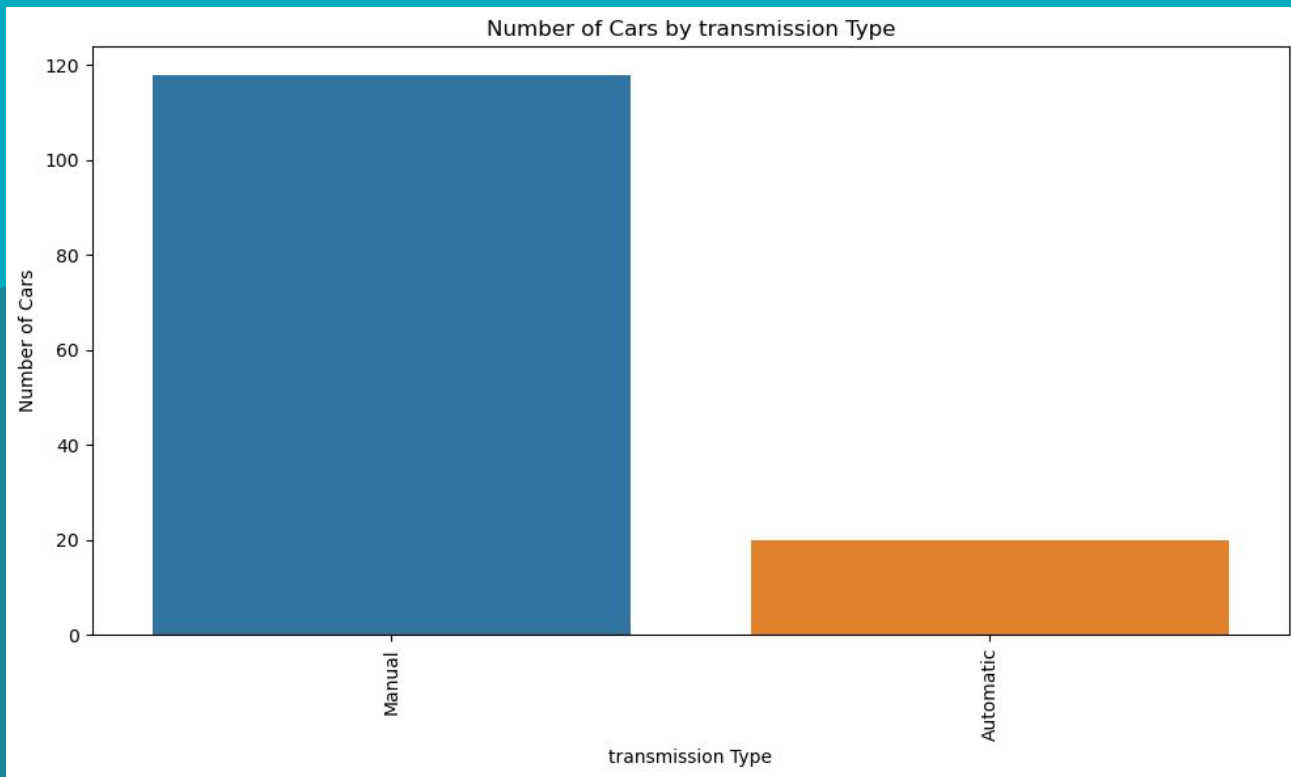
Models and Number of cars by location



Number of cars by fuel type



Number of cars by transmission type



Conclusion

- **Summary of Key Points**

- Extracting required data and analysing trends helps in making data driven solutions
- Setting up business in new location becomes easy by studying customer trends.
- Ethical considerations is important
- Unethical scraping practices can harm websites, violate privacy, and lead to legal consequences.

Thank You

Thank you for your attention

Namratha Todkar - Team Lead

Divya charitha Amballa - co TeamLead

Aniket Manwatkar - CEO EVOASTRA