TEAM HONDA

# MINI PROJECT WEB SCRAPPING PRESENTATION

BY TEAM HONDA.

TEAM HONDA

# INDEX

TEAM HONDA

# SUMMARY

Welcome to the Mini project of Data Scraping.
This mini-project focuses on web scraping Honda car listings from the Acko Drive platform. The goal is to automate the extraction of detailed vehicle data such as car name, fuel type, transmission, color availability, and price. By using Python and libraries like BeautifulSoup, the data is collected, cleaned, and stored in a structured format (CSV) for easy analysis. This project enables comprehensive insights into the available Honda models, supporting trend evaluation, price comparison, and informed decision-making. It showcases the power of web scraping in collecting real-world data for business and research applications.

TEAM HONDA

# OBJECTIVE

To scrape detailed Honda car listings from Acko drive and create a structured dataset for analyzing pricing, features, and market trends

**Details to Scrape:**

1. Car name
2. Fuel tank
3. Transmission
4. Availability
5. Price

TEAM HONDA

# SCOPE OF WORK

**•Target Platform:**

The data extraction process focuses on the website **Acko drive.**

**•Data Category:**

The project specifically targets listings of **pre-owned Honda vehicles**.

**•Data Attributes to Collect:**

The following key parameters will be retrieved from each listing:

**•Car Model Name**

**•Listed Price**

**•Manufacturing Year**

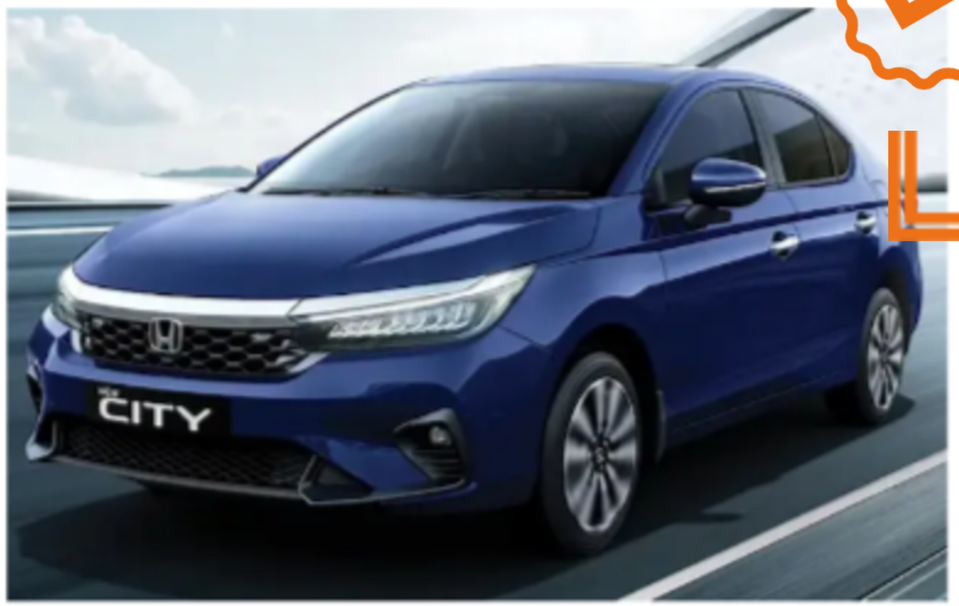
**•Total Kilometers Driven**

**•City or Location of the Vehicle**

**•Fuel Variant** (e.g., Petrol, Diesel, CNG)

**•Transmission Type** (e.g., Manual, Automatic)

**•Number of Previous Owners**




On Road Price range

₹ 13.89 L - ₹ 23.79 L Delhi

Key Specifications

Hybrid, Petrol

1498 cc

Automatic, Manual

2025 Model

5 Seater

Sedan

## Honda City Price in Delhi

Honda City on-road price (including RTO & Insurance) in Delhi ranges from 13.89 lakhs to 23.79 lakhs and it comes with 19 different variants.

TEAM HONDA

# SAMPLE EXTRACTED DATA

| | Car Name | Fuel Type | Transmission | Availability | Price |
|---|---|---|---|---|---|
| 1 | HondaCity | Petrol | Manual | Sedan | ₹15.39 L |
| 2 | HondaElevate | Petrol | Manual | SUV | ₹14.54 L |
| 3 | HondaAmaze | Petrol | Manual | Sedan | ₹9.22 L |
| 4 | HondaCity (2020-2023) | Petrol | Automatic | Sedan | ₹15.67 L |
| 5 | HondaJazz (2020-2023) | Petrol | Automatic | Hatchback | ₹11.11 L |
| 6 | ondaAmaze (2021-2024) | Petrol | Automatic | Sedan | ₹9.98 L |
| 7 | HondaWR-V (2020-2023) | Petrol | Manual | SUV | ₹10.44 L |

TEAM HONDA

# APPLICATION

## Market Research & Analysis:

This data can be used by car dealerships, auto analysts, or enthusiasts to compare prices, track availability, and analyze trends across different models and variants of Honda cars.

## Competitor Comparison:

By expanding this project, one can gather data for multiple car brands and compare features and pricing across manufacturers.

## Customer Decision Support:

Scraped data can be turned into an interactive dashboard to help potential buyers compare Honda cars and make informed decisions.

## Machine Learning Input:

Structured car data can be used to train models that predict price ranges based on features or detect outliers in car listings.

TEAM HONDA

# METHODOLOGY

**1.Connect to the Website:**

The script initiates an HTTP GET request to theURL:https://ackodrive.com/collection/honda-cars/ to retrieve the webpage content

**2.HTML Parsing:**

The fetched HTML content is processed using the BeautifulSoup library with the 'html.parser' parser to enable easy navigation and data extraction.

**3.Identify Car Information Blocks:**

The script employs the .find_all() method to locate and collect all  elements that encapsulate individual car listings.

**4.Extract Key Attributes:**

For each car listing identified, the script extracts specific details such as the model, price, fuel type, transmission, and other relevant features.

TEAM HONDA

# CHALLENGES FACED

**1.CAPTCHA Verification Barriers:**
Automated scripts are often blocked by CAPTCHA systems, which are designed to prevent non-human access. This interrupts the scraping process and requires human intervention or advanced solving techniques.

**2.Dynamically Loaded Content:**
A significant portion of the website’s data loads dynamically using JavaScript after the initial HTML is rendered. This makes it difficult for standard HTML parsers like BeautifulSoup to access the required elements directly.

**3.Frequent HTML Structure Changes:**
The website occasionally updates its design or layout, causing changes in the HTML structure. This results in previously working selectors or scraping logic becoming invalid, requiring constant script maintenance.

# CONCLUSION

This project showcases how web scraping can be used to automatically collect real-time data from websites, enabling efficient and scalable data analysis. By targeting platforms like Acko drive, it becomes possible to gather detailed information on used Honda cars—including pricing, specifications, and market trends—without manual effort.

Web scraping proves especially valuable in industries such as automotive sales, e-commerce, and market research, where timely access to structured data is essential. It allows businesses and analysts to monitor market fluctuations, compare offerings, and make informed decisions based on current data.

In summary, this project demonstrates the power of web scraping as a practical tool for turning online content into actionable insights.

# THANK YOU

By Team Honda--
Aparna Tiwari
Sufiyan Zamindar
Vignesh Thingalaya
Sonal Vishwakarma