

# Task: E-commerce Product Pricing Analysis

## 1. Data Collection Process

- Web scraping strategy using Selenium
- Use Myntra men t-shirt for web scraping.
- Collect data including product name,brand,category,price,discounted price,rating,number of review and url
- Get output of 300 products containing all data with csv file.

## 2. Data Cleaning & Preparation

- Remove duplicates.
- Handle missing values
- Convert prices and ratings to numeric formats.
- Systematic cleaning process for duplicates, missing values, and standardization
- Standardized brand names.
- Get the final dataset with a cleaned csv file.

## 3. Data Analysis

- Descriptive statistics for prices and ratings
- Brand analysis with top 5 brands by product count
- Discount analysis identifying highest and lowest discount brands
- Category-specific insights across electronics, fashion, and home appliances

## 4. Data Visualization Results

- Get Bar chart – Average discount percentage by brand.
- Using Matplotlib created bar chart

## 5. Web Scraping Challenges

- Scraping issue : Using playwright but did not get the rating and number or review
- Cannot get the 300 product easily

- Data cleaning cannot get Nan at the space of empty file.

