

**MINIPROJECT REPORT**

**2022-23**

**PROJECT NAME :** Web Scraping Using Python

**NAME :** Vansh Rastogi

**UNIVERSITY ROLL NO :** 2017118

**STUDENT ID :** 20021070

**CLASS ROLL NO :** 23

**COURSE :** B.Tech

**BRANCH :** Computer Science And Engineering

**SEMESTER :** 4th

**TEACHER NAME:** Mrs. Garima Sharma

**INTRODUCTION**

**PROBLEM STATEMENT**

To build a Web Scrapper using python to extract the details of the products.

**MOTIVATION FOR DOING THE PROJECT**

The ‘Web Scraping Using Python’

The project has been developed in order to find the details of particular product on a website or particular data from a website.

The motivation of this project is to provide information about how can we use the Selenium library to automate the tasks which are pretty common while accessing a website.

Actually the reason behind developing this project was that actually lots of time is wasted in searching the product we want. So, this model finds out the list of products we want matching our requirements.

Since, everyone in this world today is busy somewhere or the other and people don’t have time to surf over the internet and different websites to search what they need. So, a web scraper will be a quite useful tool for the people as it give you a list of items of information you want just on a click .

**OBJECTIVE**

The objective of the project ‘Web Scraping Using Python’ is to search for the things a user wants and creating a list of those things which matches the requirements of a user.

**TOOLS USED**

I have used the Pycharm for the demonstrating my Web Scraper model. Pycharm is a dedicated Python Integrated Development Environment(IDE) providing a wide range of essential tools for python developers. I have also used the latest version of Chrome Driver to access the website.

**THE PROJECT**

The project ‘Web Scraping Using Python’ aims at developing a model which can efficiently scrape data from a website based on the requirements of the user. I have used csv file for storing the scraped data for demonstrating how web scraper works.

**METHODOLOGY**

Actually this project is divided into stages.

1. **Finding the URL:** In this project, I have scraped data from Amazon to extract price, rating, details of the products. The URL for the website is <https://www.amazon.in/>.
2. **Inspecting the Page:** The data is usually nested in tags. So, we need to inspect the page to see, under which tag the data we want to scrape is nested. To inspect a page, we have to right click on the element and click on “Inspect”, it opens a page on the right side showing the html code.
3. **Find the data you want to Scrape:** So we need to inspect the name, price and, rating of a product in order to check the tags associated with it. Like for name, span tag is associated .
4. **Writing the code:** So actually to build a web scraper I have used the selenium library for automating the tasks and latest version of chrome driver that will open the webpage when we run the code.

**WORKING**

So firstly we have to import webdriver from selenium and after that create a class containing a constructor for initializing the values. We have to give the path of location where chrome driver is installed and also specify the url of the website we want to Scrape.

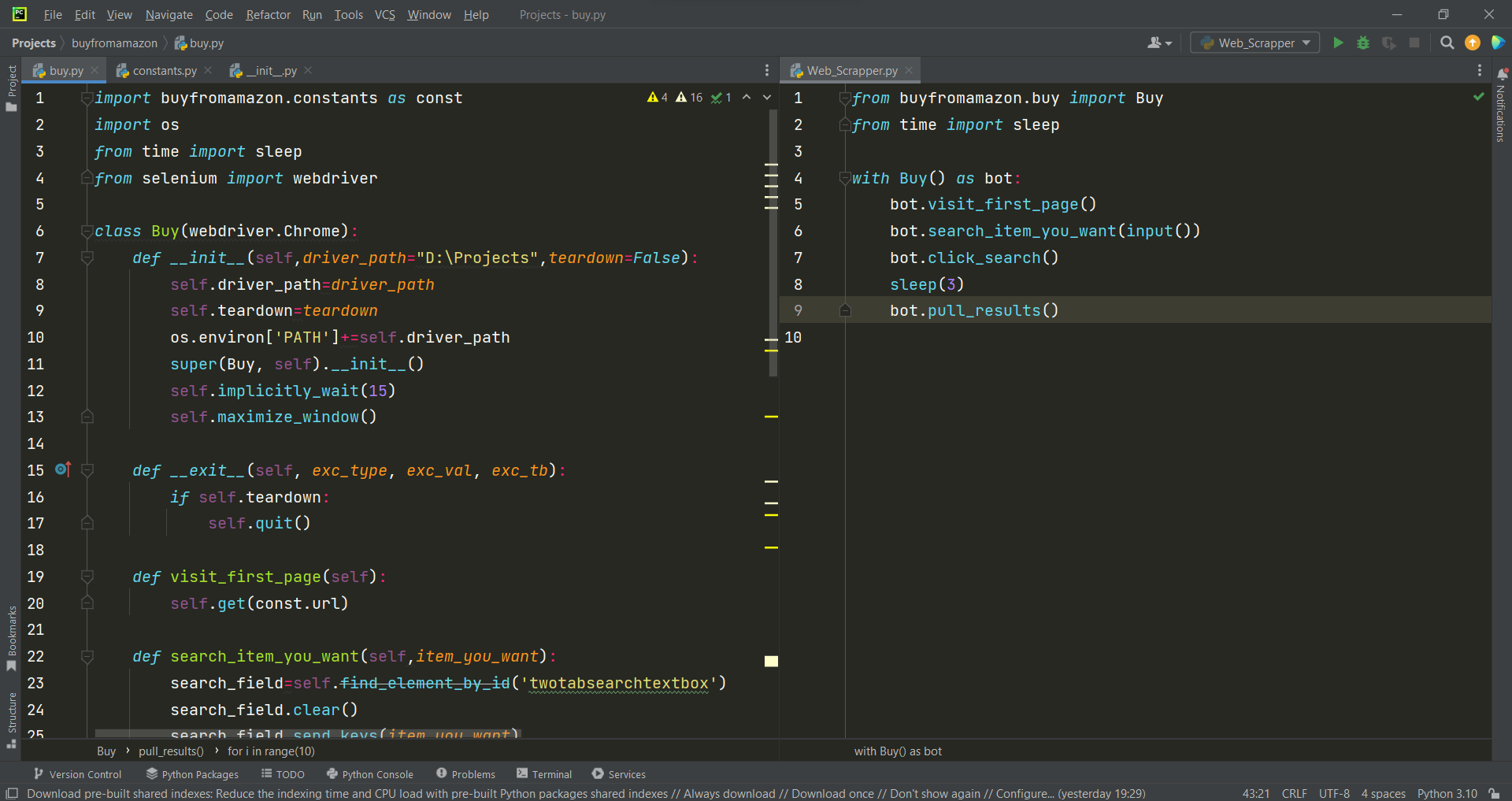
After doing all this, once we reached the home page of the website, we created a function to search product. To find the tag id for the search bar to search the product. After finding the tag we will used the find\_element\_by\_id and in its argument pass the id and user click() to get the search bar clicked. When we call the function, in function argument call input() that takes input from user of the product user want to search.

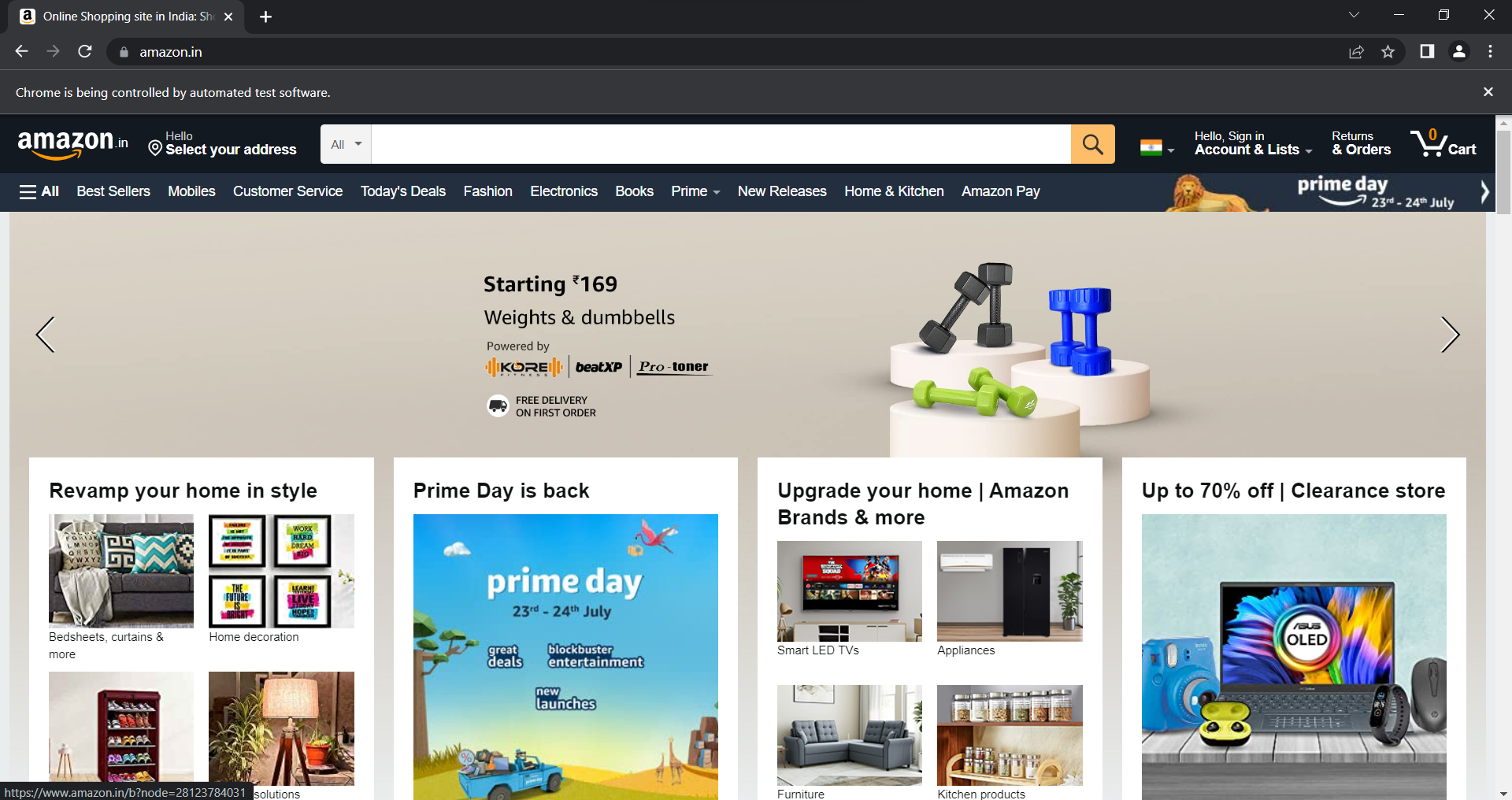
After that we have create another function to pull results from website. So we will

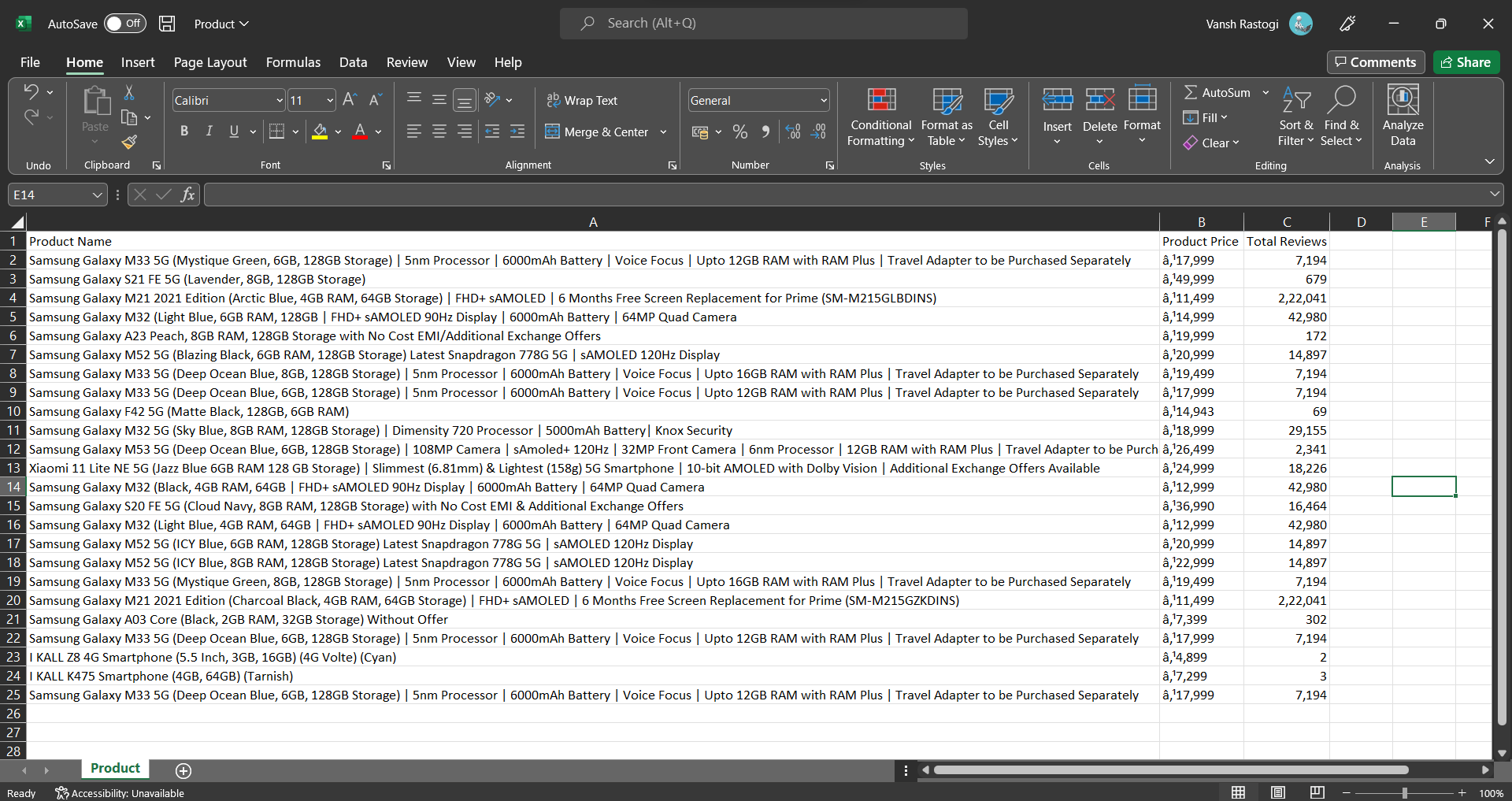
find the tags associated with name of the product, price and, rating and store these details in different lists.

After that we will create a csv file that will store the data scraped from the website.

**SCREENSHOTS**







**CONCLUSION**

So after creating above project, we have seen how to automate tasks on a browser and scrape results from a website using html tags and xpaths.

**REFERENCES**

1. [**https://www.edureka.co/blog/web-scraping-with-python/**](https://www.edureka.co/blog/web-scraping-with-python/)
2. [**https://selenium-python.readthedocs.io/**](https://selenium-python.readthedocs.io/)
3. [**https://www.youtube.com/**](https://www.youtube.com/)