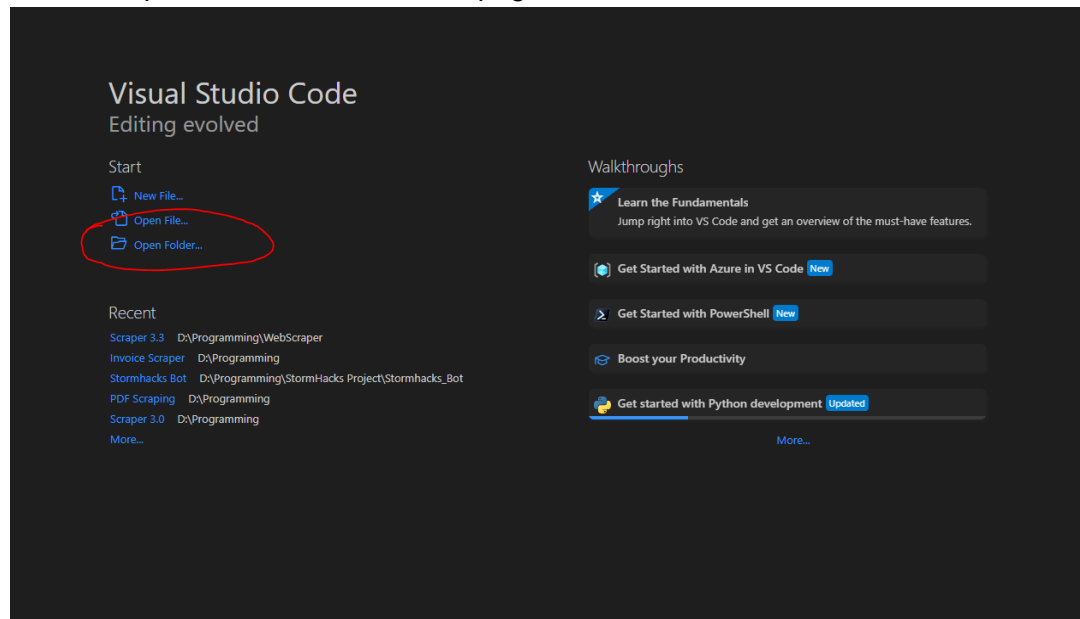# Instructions for the installations of the Jondo Web Scraper (PC VERSION)

1. Download and put the web scraper file inside a designated folder in your local device to allocate and organize for future created files.

2. Install Visual Studio Code from Google and open the scraper file with Visual Studio Code (VS Code)

3. Open the folder in VS Code

    a. Open VS Code
    b. Click the open folder if the welcome page is shown

    c. Otherwise, Click on File on the top left corner
    d. Click on Open Folder
    e. Open the folder in which the web scraper was put into

4. Download the newest version of Python from the microsoft store.

    a. Press Windows key
    b. Search for Microsoft Store and open it
    c. Type Python in the search bar
    d. Download the newest version

5. Press the Windows key and type PowerShell and run Windows PowerShell as

administrator.

6. Type these following 3 codes into the powershell to install the necessary packages for the program to run.

      a. pip install requests

      b. pip install beautifulsoup4

      c. pip install selenium

7. Go back to VS Code, and type ctrl + shift + p

8. Type in "Python Select Interpreter"

9. Select the Python with the file path of "\AppData\Local\Microsoft\WindowsApps\pythonm3.9.exe", or the one with the identification of (windows store) beside its name.

10. Download Geckodriver from this URL, select the version that matches your local device OS

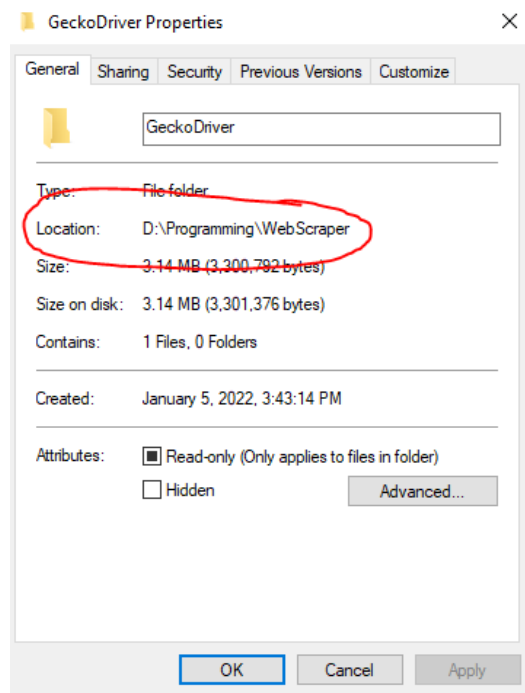   https://github.com/mozilla/geckodriver/releases

11. Unzip the downloaded zip folder to a location on your computer.

12. Copy the address of that geckodriver file and change the specified location in the webscraper.py file

      a. Go to the geckodriver file location in your local device
      b. Right click on geckodriver.exe

  c.   Click on Properties
  d.   Copy the location information as shown below

  e.   Paste inside the web scraper file at the specified location
  f.   Add "\geckodriver.exe" after it
    i.   In the example case, the full location would be
      "D:\Programming\WebScraper\GeckoDriver\geckodriver.exe"

13. Download Firefox browser
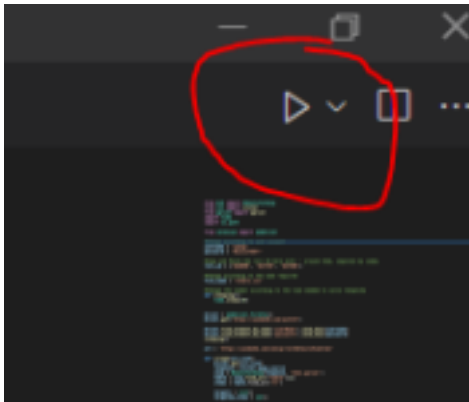
14. Add the roll numbers to the list roll_id

  a.   For each roll, it should be represented by a 6 digit number
  b.   Type them into the roll_id with single quotation marks around each number
  c.   Separate the numbers by comma if there are multiple roll numbers
  d.   Afterwards, it should look like the image below

```
 5     import os.path
 6     from selenium.webdriver.common.by import By
 7     import unicodedata
 8
 9     from selenium import webdriver
10
11     #Change according to user account
12     username = 'VincentK'
13     password = 'Nwinc17667!'
14
15     #Copy and Paste the roll ID here with '' around them, separate by comma.
16     roll_id = ['887955','887954','887964']
17     #Change according to the name required
18     file_name = 'orders.csv'
19
20     #URL of jondo website that is being accessed
21     url = 'https://jondohd.com/zencp/rollDetailsView?id='
22
23     #Change the number according to the time needed to solve reCaptcha
24  ∨ def sleeping():
25         time.sleep(30)
26
```

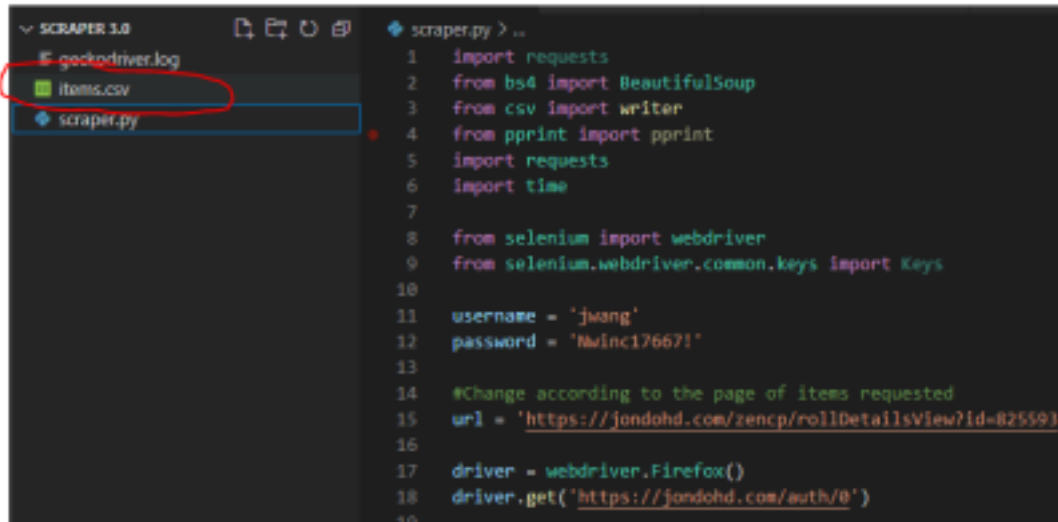15. Run the program by clicking the button on the top right

16. The file with all the information about the orders should be called "orders.csv", can it can be found within the same folder as the web scraper.

Instructions for the usage of the web scraper

1. Run the program, and it should open up a FireFox tab, where the login should already have been entered.
2. Click on the reCaptcha and try to solve it within the time span provided by the line of code time.sleep(# of seconds)
3. Click the login button and log into the main page
4. If not logged into the main page on time, restart the program by clicking run again to retry.
5. After a while, the program should automatically scrape through the list of items from the

given list of rolls without any need of operation

6. When the program is finished, the pages should stop refreshing, and there should be a csv file within the folder section on VS Code that looks like the one below

7. Right click the csv file and choose reveal in file explorer to get access to the file through windows file explorer to be copied for other usage.