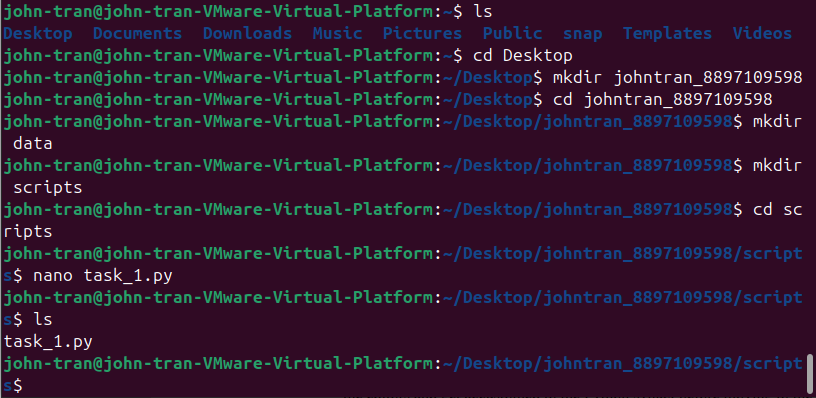
**John Tran**

**8897109598**

**2.1)**

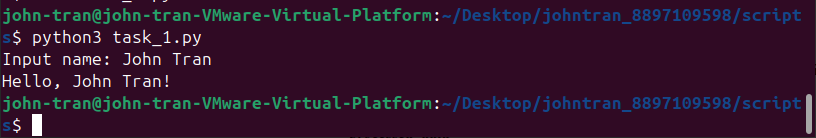


I created my class folder, the folder **data** and **scripts**, and the **task\_1.py** file inside of **scripts**.

**2.2)**

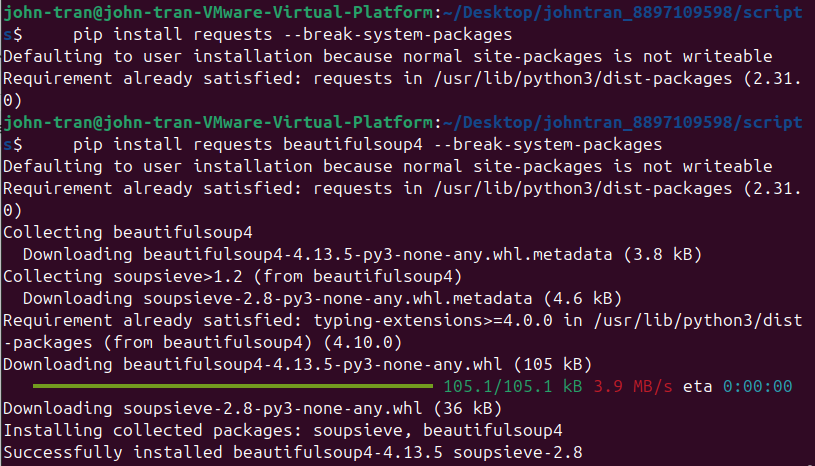


I wrote the python script that reads a user's name as input and greets the user with "Hello, [name]!".



The script prints out ‘Hello, John Tran!’ when I entered my name ‘John Tran’.

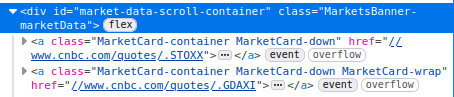
**2.3)**



I installed **requests** and **beautifulsoup4**.

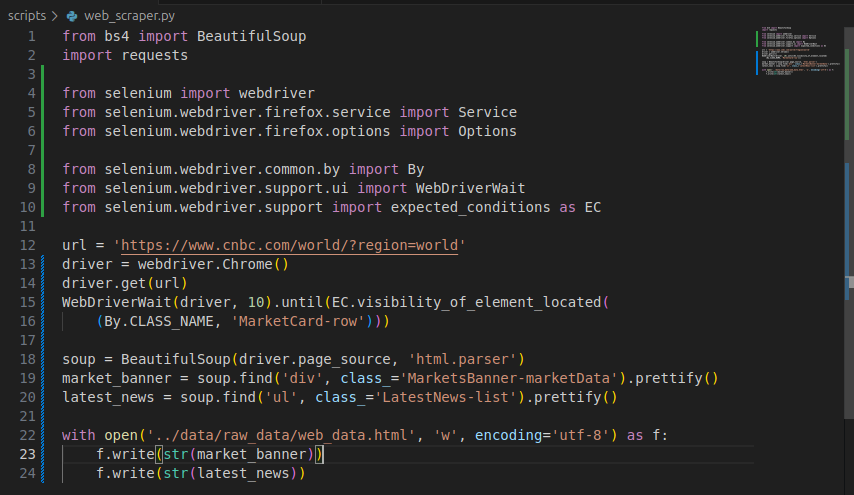


I created the folders **processed\_data** and **raw\_data**.

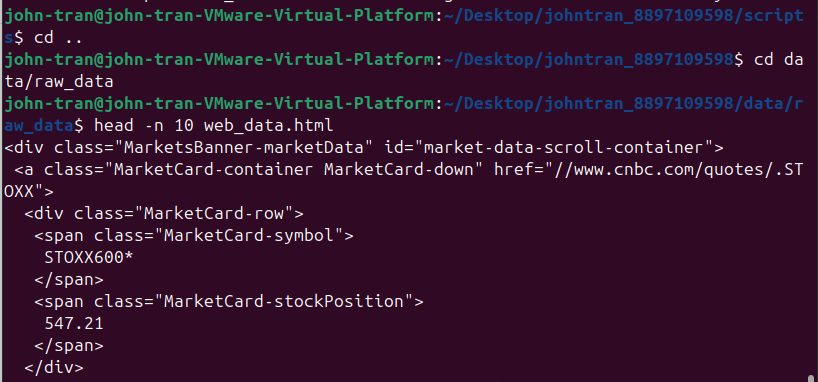




I found the tags for both the Market Banner and the Latest News section.

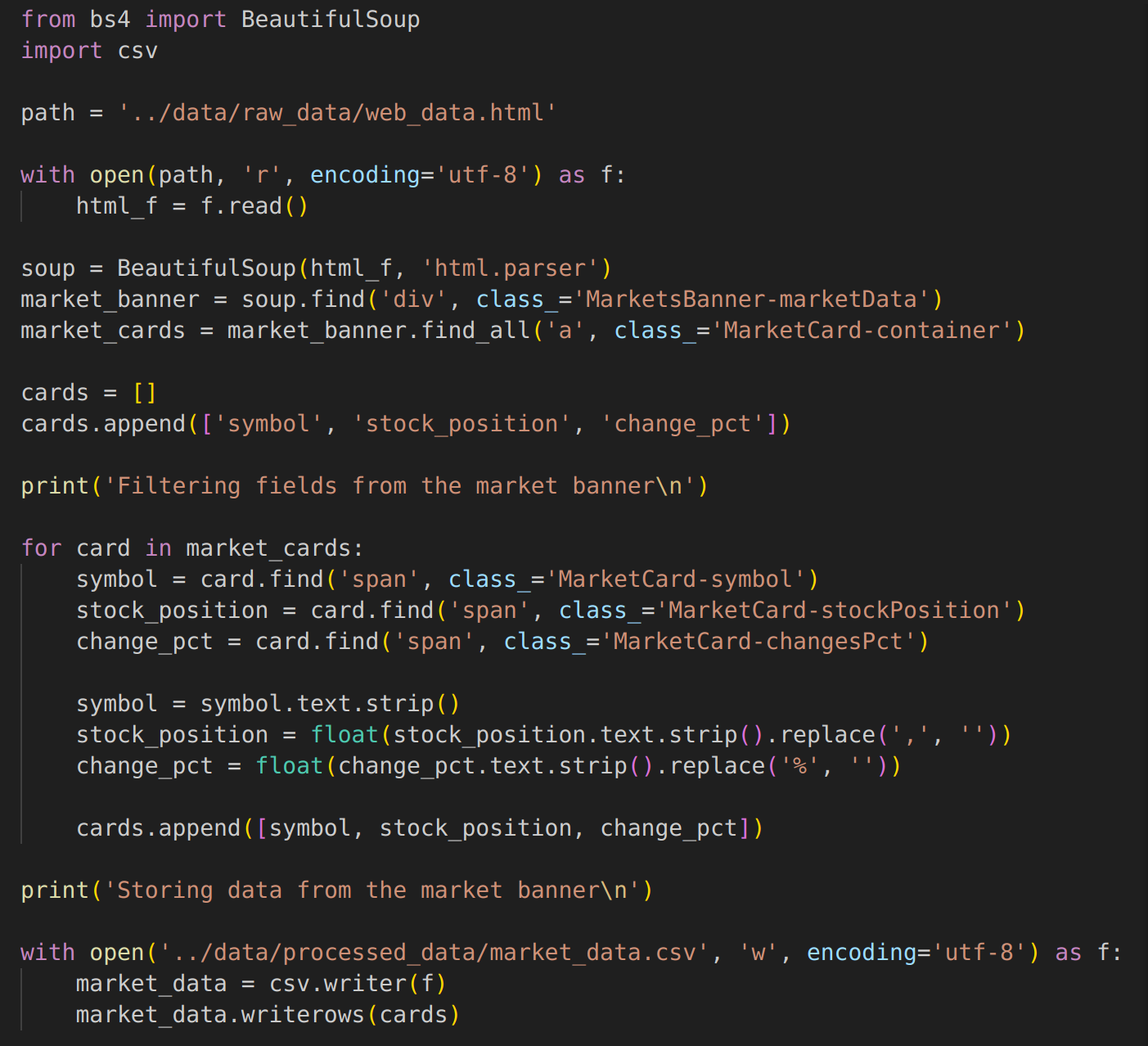


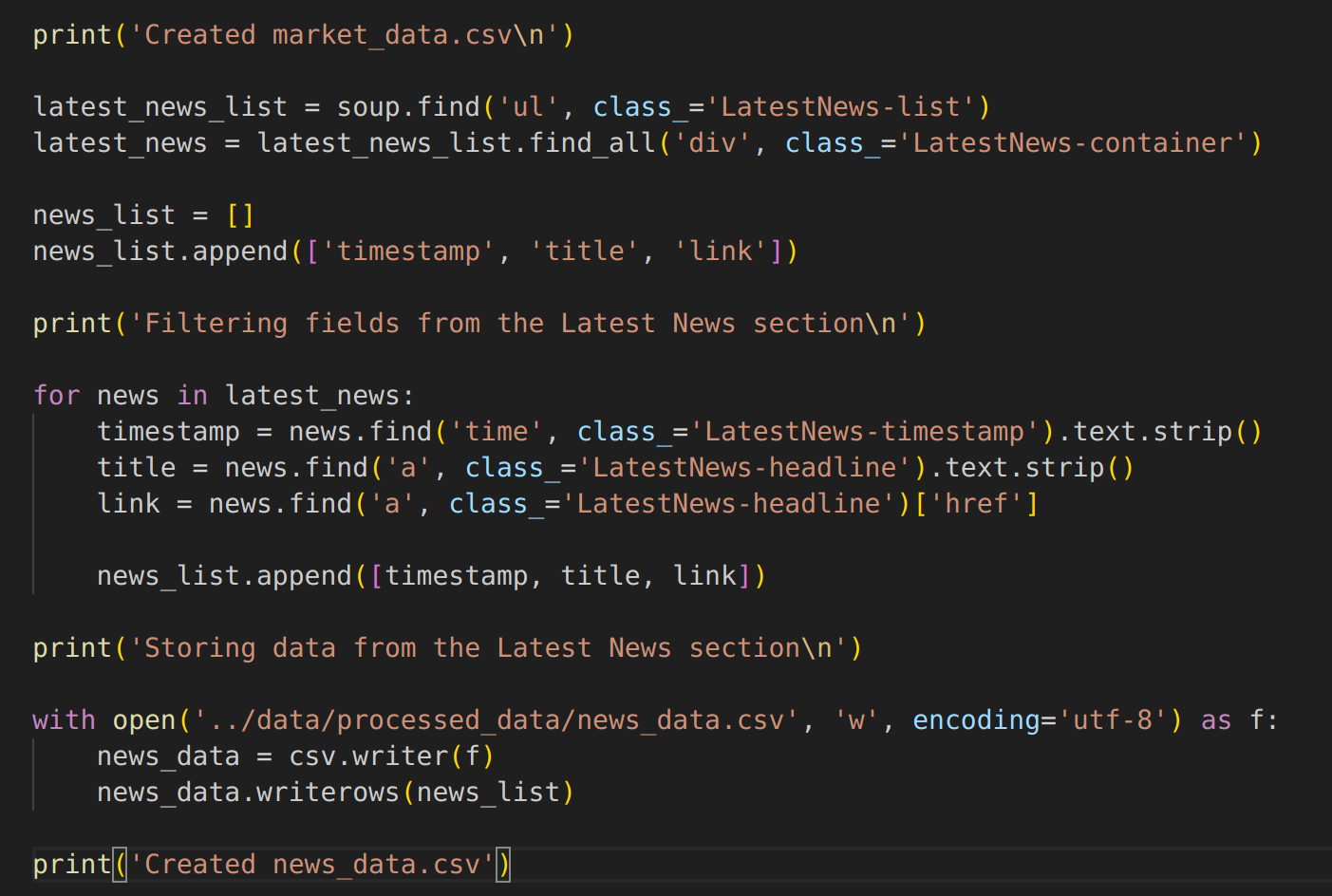
Since the Market Banner includes fields that are dynamically loaded, after some research, I used the **selenium** library in order to scrape data. After scraping all the tags from the Market Banner and Latest News section, I wrote them to **web\_data.html** that is located in the **raw\_data** within **data**.



I printed the first 10 lines of **web\_data.html**.

**2.4)**





I read **web\_data.html** and used **BeautifulSoup** to find the relevant information for both the Market Banner and the Latest News section. After cleaning the information retrieved, I add them to lists and store them under 2 csv files: **market\_data.csv** and **news\_data.csv**. I also added messages to be printed in the console.

**GitHub link**

<https://github.com/johntusc/dsci-560>

**GitHub commit screenshots**

