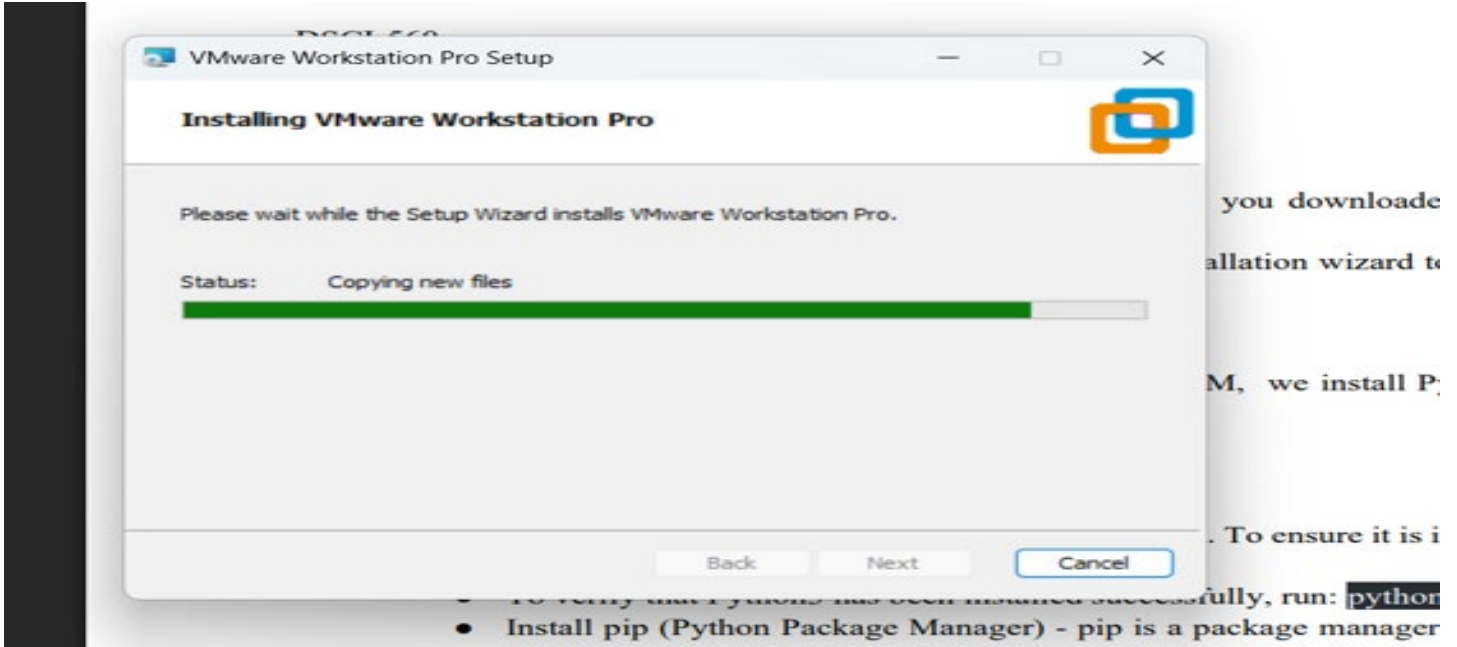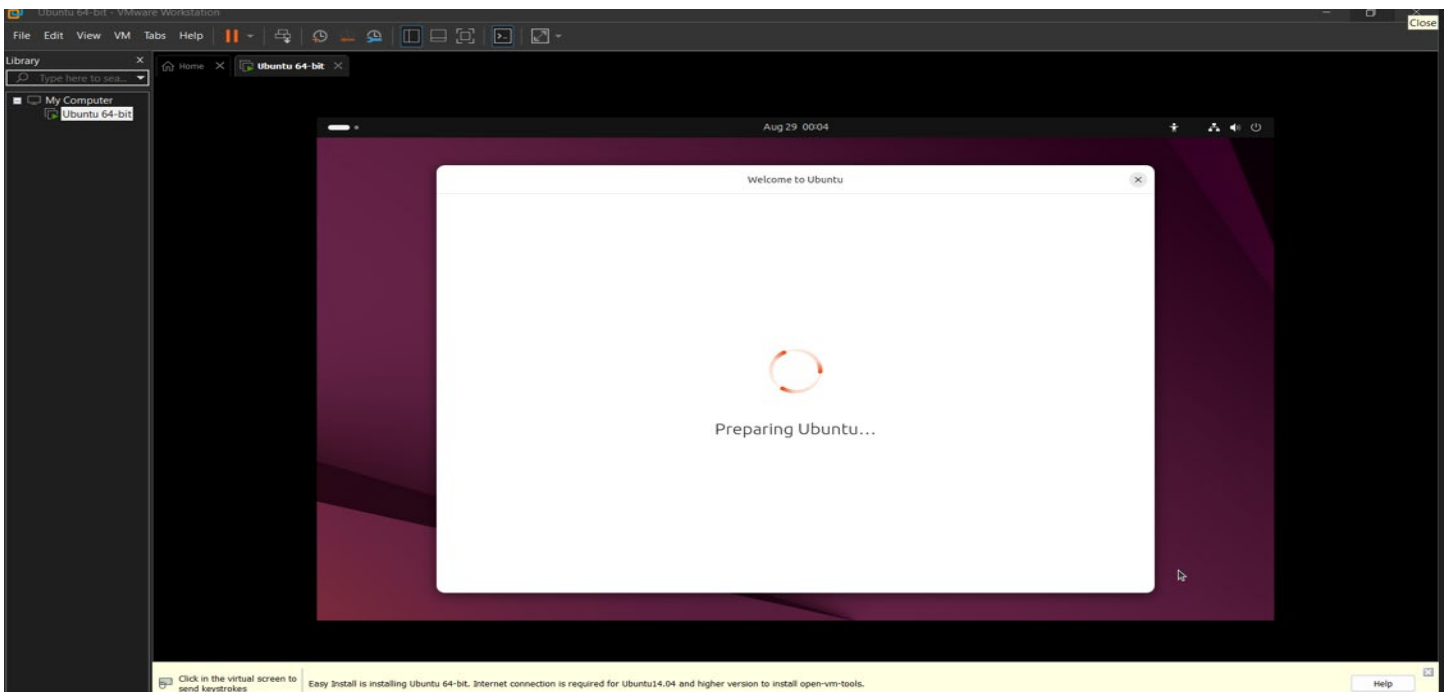## 1. Installation and Setup

The first step involves setting up a development environment. VMware Workstation was installed to host a virtualized operating system. Within VMware, an Ubuntu Linux distribution has been deployed.
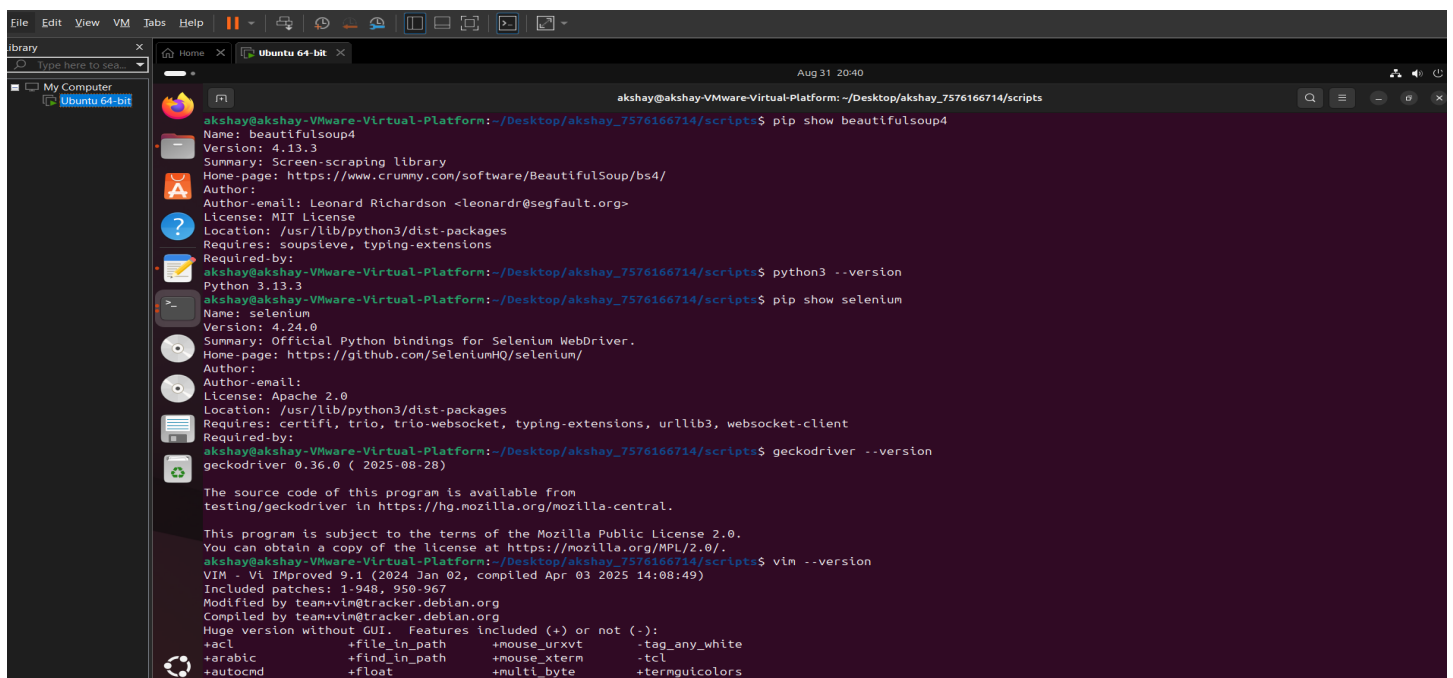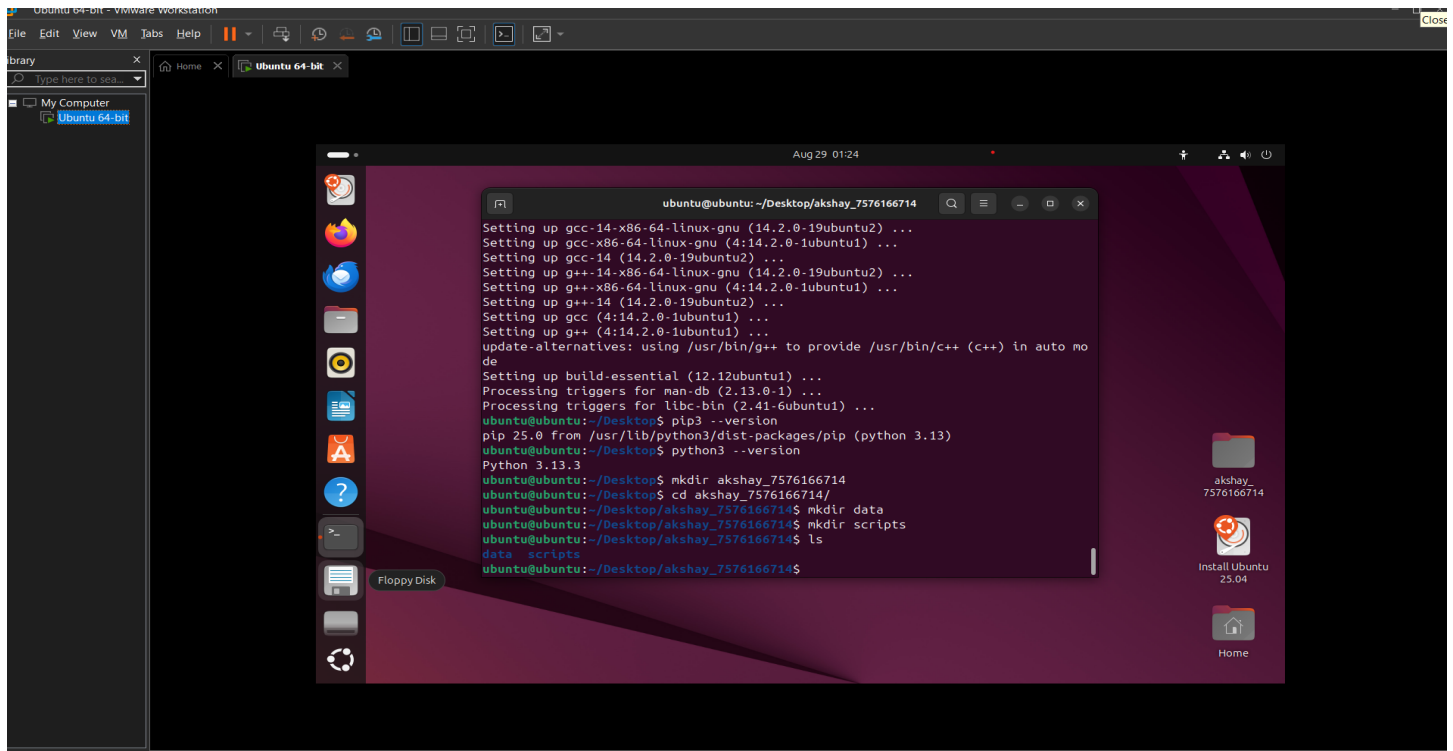

SS1:The screenshot above shows VMware installation

SS2: The screenshot above shows Ubuntu being installed on the virtual machine.

2. Get familiar with Linux

To become comfortable with the Linux environment, basic commands were executed through the terminal. This included navigating the filesystem using commands such as ls, pwd, and cd, as well as creating new directories with mkdir. I also used simple package management commands , including updating repositories (sudo apt update) and installing essential utilities (e.g., vim, python3-pip).

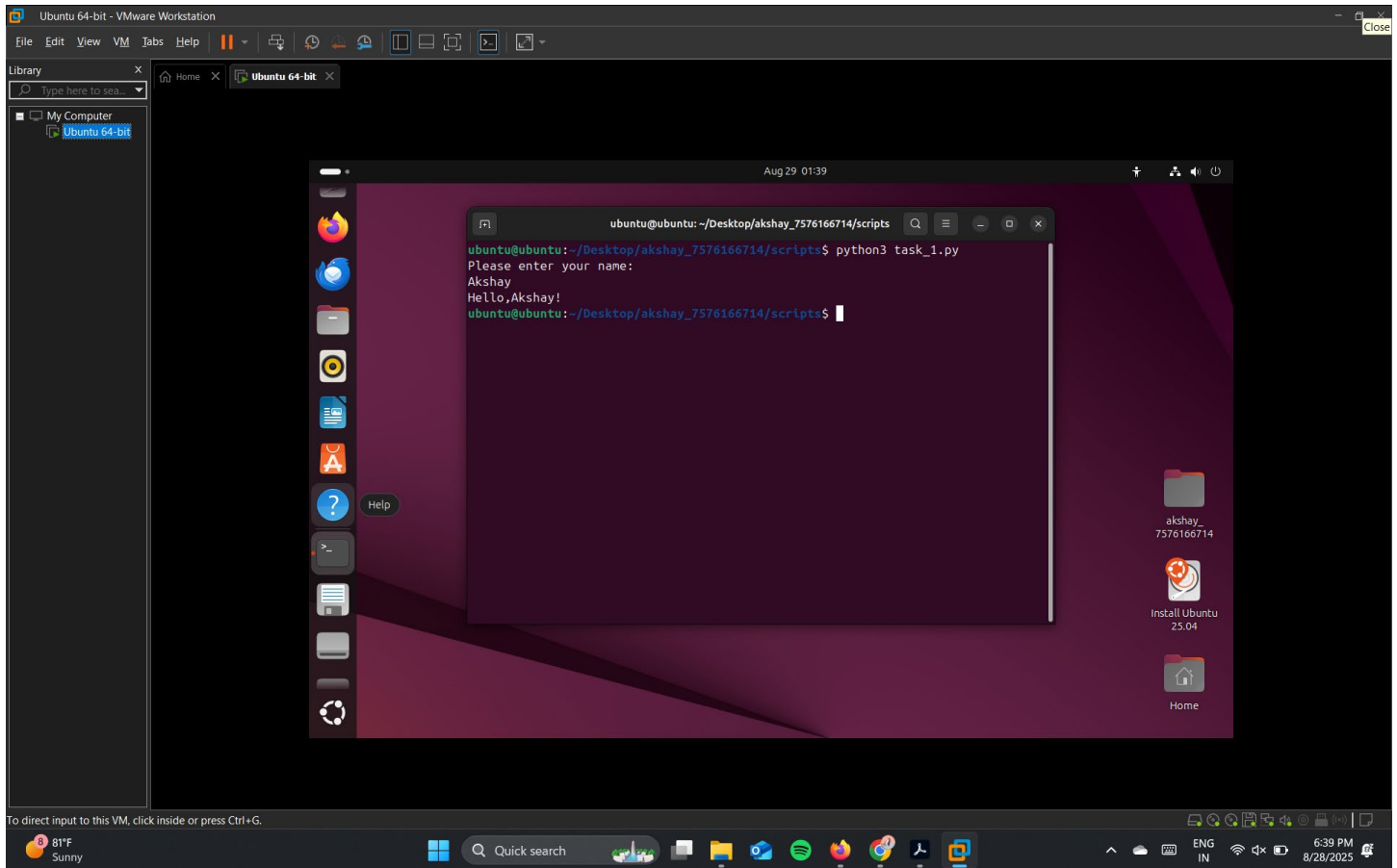The screenshot below shows getting familiar with Linux,creating the required directories, and installing packages

3. Basic Python Script

A Python script that asks for a username and then says hello was executed. Below is the execution output.

4. Web scraping task

For this task, two scripts were created: one to get the raw data and another to filter the data. Raw data is extracted from the CNBC page. To handle webpages that load content dynamically, Selenium was used in combination with the Firefox browser and GeckoDriver.

The script for launched a headless Firefox browser(headless runs in background), navigating to the target URL, waiting until market and news sections were fully loaded(wait is very important I faced issues in earlier versions), and then capturing the complete HTML file.
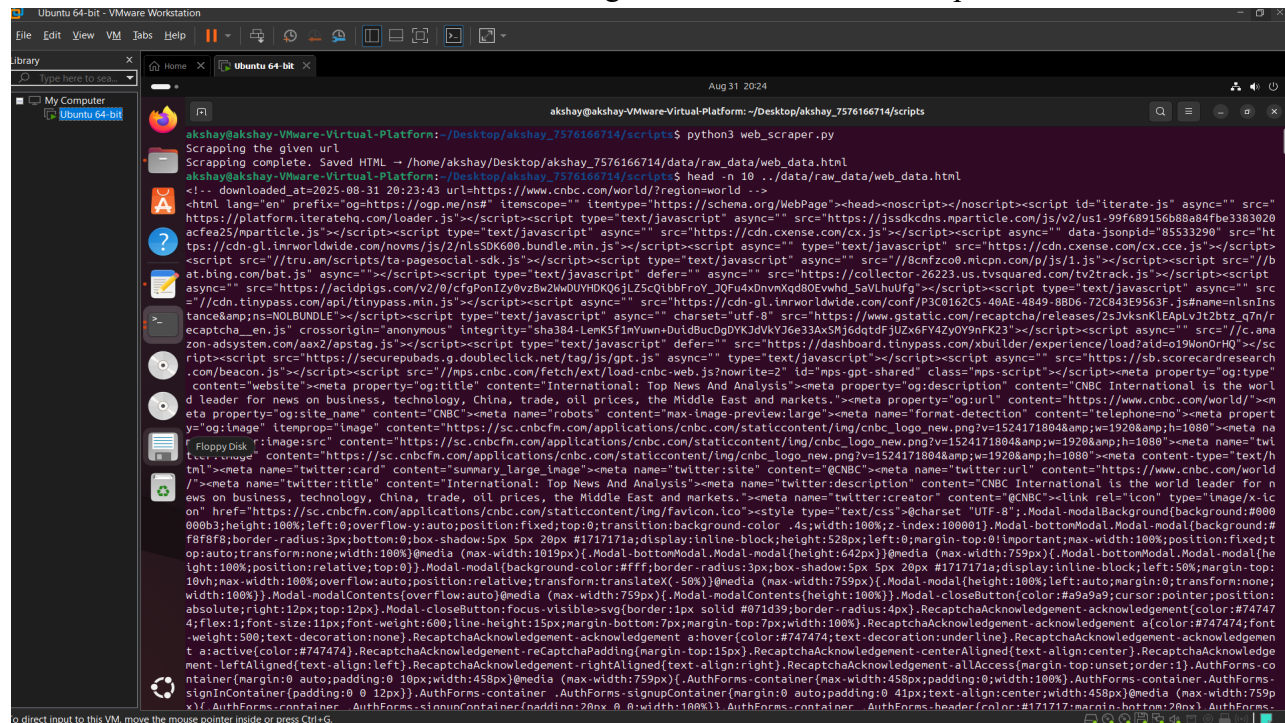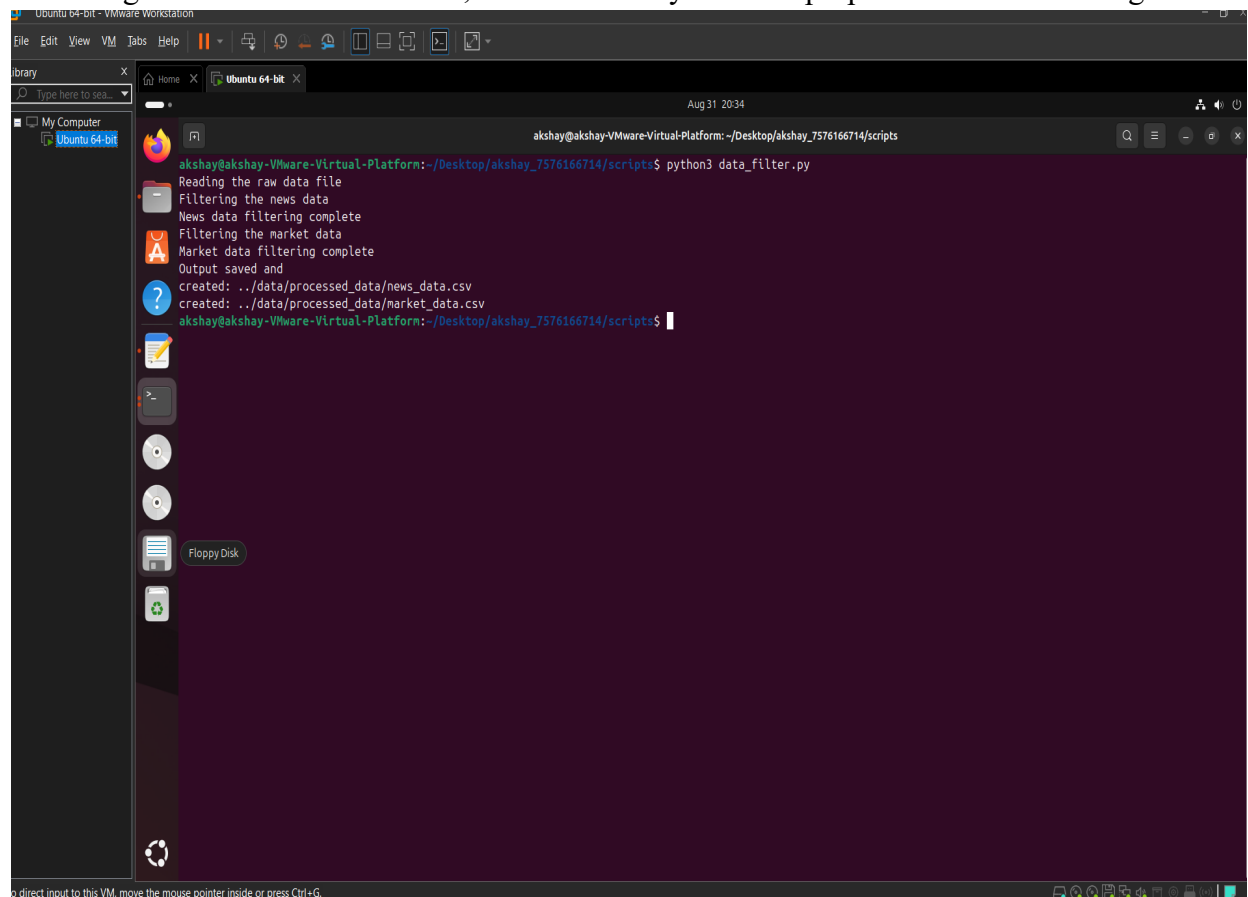

The screenshot above shows the url to be scraped

a. Getting raw data

The screenshot below shows the raw data being fetched and first 10 lines printed

b. Filtering the raw data

After I got the raw HTML data, the second Python script parsed the data using BeautifulSoup.

c. Output after filtering

The data filter script targeted two types of information:

Market banner data such as ticker symbols, price positions, and percentage changes.

Latest news headlines, including titles, timestamps, and associated links.

The parsed data was structured into Python objects and then saved in CSV format .


```
marketCard_symbol,marketCard_stockPosition,marketCard_changePct
ASX 200*,8918.90,-54.20
NIKKEI*,41849.82,-868.65
NIFTY 50*,24426.85,
HSI*,25586.19,+508.57
SHANGHAI*,3874.593,+16.666
```



```
LatestNews_timestamp,title,link
11 Min Ago,"China's Xi urges AI cooperation, rejects 'Cold War mentality' at SCO summit",https://www.cnbc.com/2025/09/01/sco-summit-xi-ai-cooperation-reject-cold-war-mentality-putin-modi-china-india-russia.html
1 Hour Ago,BYD's Hong Kong shares fall nearly 8% after quarterly profit drop,https://www.cnbc.com/2025/09/01/byd-hong-kong-shares-fall-nearly-8percent-after-quarterly-profit-drop.html
1 Hour Ago,"China's August factory activity beats estimates, expanding at fastest pace in five months: private survey ",https://www.cnbc.com/2025/09/01/chinas-august-factory-activity-beats-estimates-expanding-at-fastest-pace-in-5-months.html
2 Hours Ago, Tariff uncertainty intensifies with the U.S. court's ruling,https://www.cnbc.com/2025/09/01/cnbc-daily-open-tariff-uncertainty-intensifies-with-the-us-courts-ruling.html
3 Hours Ago,Asia markets trade mixed as investors assess China's RatingDog manufacturing data,https://www.cnbc.com/2025/09/01/2025/09/01/asia-stock-markets-today-live-updates-nikkei-225-asx-200-kospi-hang-seng-csi-300-sensex-nifty-50.html
4 Hours Ago,Spanish stocks have surged this year — and there could be more upside ahead,https://www.cnbc.com/2025/09/01/why-spains-stock-market-is-booming-and-what-might-come-next.html
7 Hours Ago,Rudy Giuliani hospitalized with serious injuries after car accident,https://www.cnbc.com/2025/08/31/rudy-giuliani-hospitalized-after-car-accident-spokesman.html
11 Hours Ago,Trump trade adviser blasts tariff ruling: It's 'weaponized partisan injustice',https://www.cnbc.com/2025/08/31/trump-trade-adviser-tariff-ruling-peter-navarro-.html
13 Hours Ago,"The top 10 jobs offering remote, part-time work in 2025, says new report",https://www.cnbc.com/2025/08/31/the-top-job-titles-for-remote-part-time-work-according-to-flexjobs-report.html
14 Hours Ago,I've studied over 200 kids—6 'magic phrases' that make children listen to parents,https://www.cnbc.com/2025/08/31/child-psychologist-parents-who-raise-kids-who-listen-use-6-magic-phrases.html
14 Hours Ago,"After losing her husband, a therapist's second marriage opened 'new possibilities'",https://www.cnbc.com/2025/08/31/after-losing-her-husband-therapists-
```