Name: Rongzi Xie USC ID:9100096985

Github link and push records:

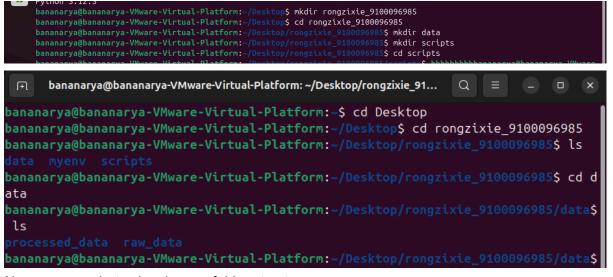
https://github.com/Bananarya/DSCI560 lab1

1.3 Install Python on Linux

```
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop$ sudo apt install python3
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
python3 is already the newest version (3.12.3-0ubuntu2).
python3 set to manually installed.
O upgraded, O newly installed, O to remove and 21 not upgraded.
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop$ python3--version
python3--version: command not found
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop$ python3 --version
Python 3.12.3
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop$ mkdir rongzixie 9100096985
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop$ cd rongzixie_9100096985
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop/rongzixie_9100096985$ mkdir data
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop/rongzixie_9100096985$ mkdir scripts
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop/rongzixie_9100096985$ cd scripts
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop/rongzixie_9100096985/scripts$
```

Above I installed python3 in ubuntu

2.2 Playing around with Linux Terminal



Above screenshots showing my folder structure.

```
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop/rongzixie_9100096985/scrip
ts$ ls
data_filter.py task_1.py web_scraper.py
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop/rongzixie_9100096985/scrip
ts$
```

Above showing the Is command and its results.

2.2 A basic Python Script

```
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop/rongzixie_9100096985/scrip
ts$ python3 task_1.py
please provide you name
Arya
Hello, Arya!
```

Above showing the input and result of my python script

2.3 Python Web-scraping Task

```
bananarya@bananarya-VMware-Virtual-Platform: ~/Desktop/rongzixie_91...
bananarya@bananarya-VMware-Virtual-Platform:~/Desktop/rongzixie_9100096985/data/
raw_data$ head -n 10 web data.html
<section class="MarketsBanner-container" data-analytics="HomePageInternational-m</pre>
arketsBanner-1-0" data-test="marketsBanner-0" id="HomePageInternational-MarketsB
anner-1">
<div class="MarketsBannerMenu-marketBannerMenuWrapper">
  <div class="CustomBannerAdder-container">
   <button class="CustomBannerAdder-addButton" data-testid="addButton">
    <span class="CustomBannerAdder-plusButton" data-testid="plusButton">
     <svg data-analytic-id="plus-icon" fill="none" height="17" viewbox="0 0 17 1
7" width="17" xmlns="http://www.w3.org/2000/svg">
      <path clip-rule="evenodd" d="M9.1427 7.43729L9.18188 1.1331C9.18391 0.8073</pre>
63 8.92149 0.544946 8.59576 0.546971C8.27003 0.548995 8.00433 0.814694 8.0023 1.
14043L7.96312 7.44462L1.46316 7.48502C1.13743 7.48704 0.871726 7.75274 0.869702
8.07847C0.867677 8.40421 1.13009 8.66662 1.45583 8.6646L7.95579 8.6242L7.91539 1
5.1242C7.91337 15.4499 8.17579 15.7123 8.50152 15.7103C8.82725 15.7083 9.09295 1
5.4426 9.09497 15.1168L9.13537 8.61687L15.4396 8.57769C15.7653 8.57566 16.031 8.
30997 16.033 7.98423C16.035 7.6585 15.7726 7.39608 15.4469 7.39811L9.1427 7.4372
9Z" fill="#005594" fill-rule="evenodd">
      </path>
     </svq>
    </span>
```

Above is the first ten line of the scraped data

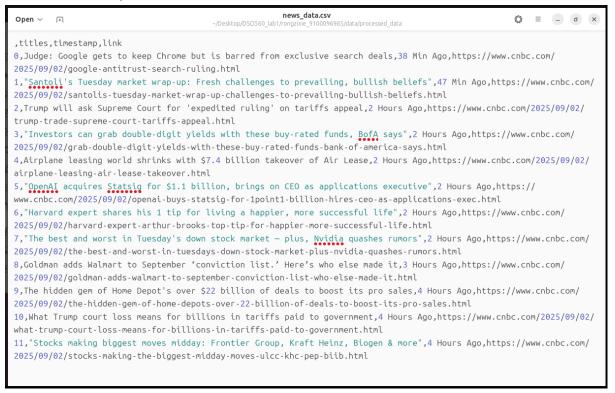
2.4 Data Filtering Task

```
(myenv) bananarya@bananarya-VMware-Virtual-Platform:~/Desktop/rongzixie_91000969
85/scripts$ python3 data_filter.py
reading data from html
filtering market data
storing market data
filtering news data
storing news_data
storing news_data
```

Terminal output lines



merkets data scraped and filtered.



newsdata scraped and filtered