

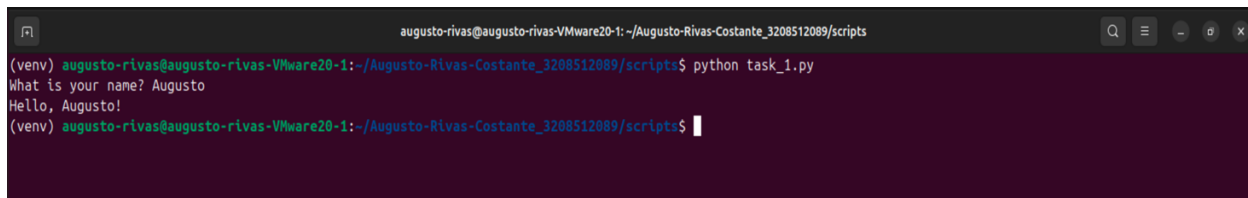
Laboratory Assignment 1

This report showcases the results of the work accomplished as part Lab 1. It includes a screenshot for all outputs generated from each Python script. The work was done using Ubuntu inside a Linux virtual machine in VMWare Fusion.

All the work related to this assignment has been included in this GitHub repository:
<https://github.com/arivas191/dsci-560>.

Basic Python Script

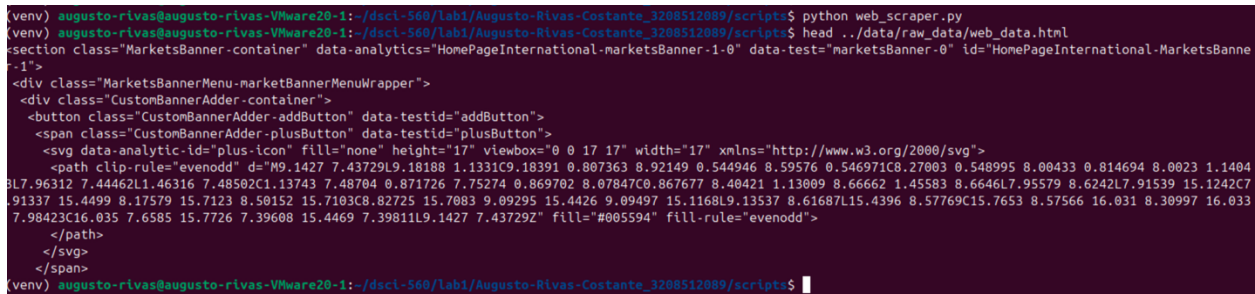
The first task was to create a simple Python script to prompt the user for their name and return a greeting message accordingly:

A terminal window with a dark background. The prompt is 'augusto-rivas@augusto-rivas-VMware20-1: ~/Augusto-Rivas-Costante_3208512089/scripts'. The user runs 'python task_1.py'. The script prompts 'What is your name?' and the user enters 'Augusto'. The script then outputs 'Hello, Augusto!'.

```
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/Augusto-Rivas-Costante_3208512089/scripts$ python task_1.py
What is your name? Augusto
Hello, Augusto!
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/Augusto-Rivas-Costante_3208512089/scripts$
```

Python Web-scraping Task

The second task involved scraping specific data from the provided news website of CNBC. Specifically, the data for the markets banner at the top of the website and the data for the latest news was scraped and collected into an HTML file inside the “raw_data” directory. The screenshot below shows the first 10 lines of such HTML file:

A terminal window with a dark background. The prompt is 'augusto-rivas@augusto-rivas-VMware20-1: ~/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/scripts'. The user runs 'python web_scraper.py'. The script outputs the first 10 lines of an HTML file. The HTML content includes a 'MarketsBanner-container' section, a 'MarketsBannerMenu-marketBannerMenuWrapper' div, and a 'CustomBannerAdder-container' div with a button and a span containing an SVG clip rule.

```
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/scripts$ python web_scraper.py
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/scripts$ head ../data/raw_data/web_data.html
<section class="MarketsBanner-container" data-analytics="HomePageInternational-marketsBanner-1-0" data-test="marketsBanner-0" id="HomePageInternational-MarketsBanner-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 17" width="17" xmlns="http://www.w3.org/2000/svg">
            <path clip-rule="evenodd" d="M9.1427 7.43729L9.18188 1.1331C9.18391 0.807363 8.92149 0.544946 8.59576 0.546971C8.27083 0.548995 8.00433 0.814694 8.0023 1.1404
            8.1337 15.4499 8.17579 15.7123 8.50152 15.7103C8.82725 15.7083 9.09295 15.4426 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 7.43729Z" fill="#005594" fill-rule="evenodd">
          </path>
        </svg>
      </span>
    </div>
  </div>
</section>
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/scripts$
```

Data Filtering Task

The final task involved writing a filtering script for collecting the relevant data within the markets and latest news sections, and finally outputting this data into corresponding CSV files. The screenshot below shows information logs that were printed during the filtering process:

```
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/scripts$ python data_filter.py
Reading web_data.html...
Filtering fields...
Storing market data...
CSV created: /home/augusto-rivas/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/data/processed_data/market_data.csv
Storing news data...
CSV created: /home/augusto-rivas/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/data/processed_data/news_data.csv
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/scripts$
```

Lastly, the screenshot below shows the outputs that were consolidated into the two CSV files:

```
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/data/processed_data$ cat market_data.csv
Symbol,StockPosition,ChangePct
ASX 200*,8,810.20",-1.02%
NIKKEI*,42,203.17",-0.25%
NIFTY 50*,24,579.60",UNCH
HSI*,25,545.79",+0.19%
SHANGHAI*,3,840.793",-0.45%
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/data/processed_data$ cat news_data.csv
Timestamp,Headline,Link
44 Min Ago,Australia's GDP expands more than expected; fastest pace since September 2023,https://www.cnn.com/2025/09/03/australias-gdp-expands-more-than-expected-f
astest-pace-since-september-2023.html
45 Min Ago,China's Xi says the world faces 'peace or war' as Trump claims Beijing conspiring against U.S.,https://www.cnn.com/2025/09/03/china-military-day-parade-
xi-trump-beijing-us.html
1 Hour Ago,In photos: China's military might on display with Putin and Kim Jong Un in attendance,https://www.cnn.com/2025/09/03/chinas-military-might-on-display-wi
th-putin-and-kim-jong-un-in-attendance.html
1 Hour Ago,Trump says stocks "want the tariffs" - but it might be Treasurys instead,https://www.cnn.com/2025/09/03/cnn-daily-open-trump-says-stocks-want-tariffs-b
ut-it-might-be-treasurys.html
3 Hours Ago,Shut windows and empty offices: China's military parade is closed to the public,https://www.cnn.com/2025/09/02/china-military-parade-victory-day.html
3 Hours Ago,Asia markets trade mixed as tariff concerns weigh,https://www.cnn.com/2025/09/03/asia-stock-markets-today-nikkei-225-asx-200-kospi-hang-seng-csi-300-se
nsex-nifty-50.html
(venv) augusto-rivas@augusto-rivas-VMware20-1:~/dsci-560/lab1/Augusto-Rivas-Costante_3208512089/data/processed_data$
```