Get Familiar with Linux and Python

Course Assignment Report

Name: Sen Pang
USC ID: 8598139533
Operating System: Ubuntu 24.04.3 LTS (ARM64)
Environment: VMware Fusion on macOS

1. Introduction

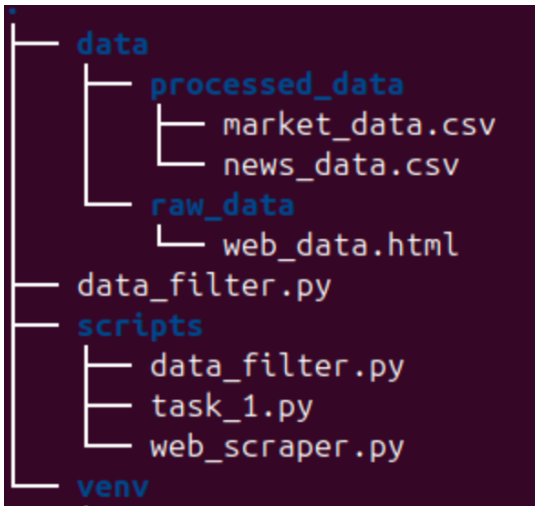
This assignment focuses on gaining hands-on experience with Linux terminal commands and basic Python scripting. The tasks include directory management, script creation, user input handling, and simple web data processing.

2. Playing Around with Linux Terminal

The Linux terminal was opened to perform basic file system operations.

A project directory SenPang_8598139533 was created on the Desktop, with data and scripts as subdirectories. An empty Python file task_1.py was added to the scripts folder. Directory structure was verified using commands such as ls and tree.

Screenshot 1: Directory Structure Creation

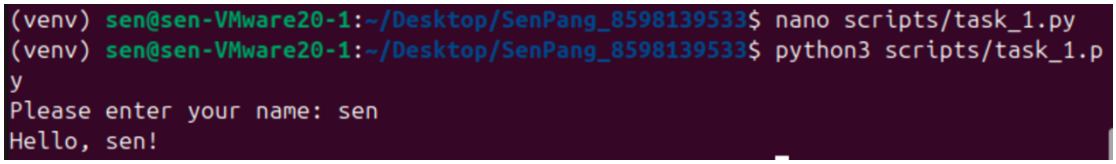


3. A Basic Python Script

The task_1.py file was opened in a terminal editor (nano). A short Python script was added to prompt the user for their name and print a greeting message..

The script was executed successfully from the terminal and displayed the correct output.

Screenshot 2: Python Script Execution

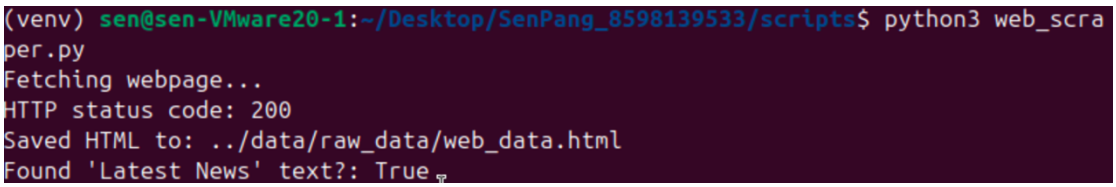


4. Web Scraping and Data Processing

A Python script fetched a webpage and saved the HTML content locally, returning an HTTP 200 status code. The raw HTML file was stored in data/raw_data.

A second script parsed the HTML and extracted the required fields into CSV files saved in data/processed_data.

Screenshot 3: Web Scraper Running Successfully



Screenshot 4: Printing first 10 lines of HTML file

Screenshot 5: Data Filtering and CSV Generation

```
(venv) sen@sen-VMware20-1:~/Desktop/SenPang_8598139533/scripts$ python3 data_f
ter.py
Reading HTML: ../data/raw_data/web_data.html
Filtering Market banner fields...
Market rows found: 0
Market CSV created: ../data/processed_data/market_data.csv
Filtering Latest News fields...
News rows found: 30
News CSV created: ../data/processed_data/news_data.csv
Done.
```

The market banner returned zero rows due to dynamic content loading on the CNBC webpage at runtime.

Screenshot 6: Generated Files Verification

```
(venv) sen@sen-VMware20-1:~/Desktop/SenPang_8598139533/scripts$ ls ../data/proce
ssed_data
market_data.csv  news_data.csv
```

5. Conclusion

This assignment offered practical exposure to Linux terminal commands, Python scripting, and basic data processing workflows. Completing the tasks improved my familiarity with the Linux environment and strengthened my ability to write Python programs for automation and data handling.