**INSTRUCTION**

**1. My approach towards the solution**

* I approached the problem step by step, focusing on manageable tasks.
* Initially, I recognized the need to extract information from web pages and analyze the text content.
* To handle this, I divided the task into smaller chunks such as retrieving data, processing text, and creating a structured output.
* I relied on popular tools like requests for fetching web page data, BeautifulSoup for understanding the HTML structure of web pages, and pandas for organizing and managing data.

**2. How to run the .py file to generate Output**

* Open your command prompt or terminal on your computer.
* Navigate to the directory where the Python script is located using the 'cd' command.
* Once you're in the correct directory, type python filename.py and hit Enter. Replace filename.py with the name of the Python script you want to run.
* The script will start executing, and upon completion, you'll find the output file in the same directory.

**3. Required Dependencies**

* The script relies on specific tools called dependencies to function correctly:
  + requests: This library helps in making HTTP requests to fetch web page content.
  + BeautifulSoup: It's used for parsing HTML and extracting data from web pages.
  + pandas: This library is used for data manipulation and analysis, especially for handling structured data like Excel files.
  + nltk: This library aids in various natural language processing tasks, such as tokenization and sentiment analysis.
* You can easily install these dependencies using pip, the Python package installer, by running pip install requests beautifulsoup4 pandas nltk in your command prompt or terminal. Make sure to install these before running the script to ensure it works smoothly.