**Instructions for Running the Text Analysis Script**

**Approach :**

The provided Python script performs text analysis on a set of URLs, extracting relevant information and computing various variables as specified in the "Text Analysis.docx" file. Here's a Key steps of approach :

**1.Data Extraction:**

* Using requests library to extracting html content from each url.
* After extracting beautifulsoup library is used for parsing the html content and extracting the article from html content.

**2. Text Analysis :**

* Developed functions for sentiment analysis, sentence and word count, syllable count, complex word detection, and other relevant metrics.
* Used nltk for tokenization as specified in ‘text analysis.docx’ file.
* Calculated the values for variables with the help of formula and definition as specified in "Text Analysis.docx" file.

**3. Output :**

* Organized the results into a pandas DataFrame for easy manipulation.
* Using pandas Stored the results in excel file

**How to Run the .py File:**

**1. Dependencies:**

* Ensure you have Python installed (version 3.x recommended).
* Install the required Python libraries by executing the following commands in your terminal or command prompt:
* Pip install requests .
* Pip install beautifulsoup4
* Pip install pandas
* Pip install nltk

**2. NLTK Resources:**

* Before running the script, download the necessary NLTK resources by executing the following commands
* Import nltk
* nltk.download(‘punkt’)
* nltk.download(‘stopwords’)

**3. Script Execution:**

* Open a terminal or command prompt.
* Navigate to the directory containing your Python script.
* Execute the script by running the command:
* python script\_name.py

**4. Output :**

* Upon successful execution, the script will generate an Excel file named 'Output.xlsx' containing the computed variables for each URL.