Sentiment Analysis and Readability Analysis Instructions

1. Folder Structure:

- Create a folder named "scraped\_data" and place your text files containing scraped content in this folder.

- Create a folder named "analyzed\_data" where the analyzed results will be stored.

2. Data Files:

- Ensure you have the following data files in the same directory as your script:

- positive-words.txt: A list of positive words.

- negative-words.txt: A list of negative words.

3. Running the Analysis:

- Open the Python script provided in your preferred code editor or IDE.

4. Modify File Paths (Optional):

- If your scraped data files or data files are in different directories, update the 'input\_directory' and 'output\_directory' variables with the correct paths.

5. Running the Script:

- Run the script to perform sentiment analysis and readability analysis on the scraped text files.

6. Analysis Results:

- The script will process each text file in the "scraped\_data" folder.

- For each file, it will perform sentiment analysis and readability analysis.

- The results will be printed to the console. You can modify the script to save the results to a file or database if needed.

7. Interpreting Results:

- Sentiment Analysis:

- Positive Score: Number of positive words found in the text.

- Negative Score: Number of negative words found in the text.

- Polarity Score: A score between -1 and 1 indicating the polarity of the text (positive or negative).

- Subjectivity Score: A score between 0 and 1 indicating the subjectivity of the text (objective or subjective).

- Readability Analysis:

- Average Sentence Length: Average number of words per sentence.

- Percentage of Complex Words: Percentage of words that are considered complex.

- Fog Index: A readability index calculated using sentence length and complex word percentage.

- Average Words Per Sentence: Average number of words per sentence.

8. Further Customization (Optional):

- You can customize the analysis by modifying the stopwords, adding more dictionaries, or adjusting the analysis formulas as needed.