Test case solution explanation:

STEP 1: Reading the excel file and performing Web Scraping

**Read the Excel File:**

* The input Excel file contains two columns: URL\_ID and URL.

**Fetch Webpage Content:**

* Use Python’s requests and BeautifulSoup libraries to send HTTP requests and extract article text.

**Save Extracted Text:**

* Store the scraped text in individual .txt files named using their corresponding URL\_ID

STEP 2: Cleaning the extracted text

Includes converting the text to lowercase, removing punctuations and its tokenization.

Removing the given stopwords.

Saving cleaned text in a directory “cleaned\_articles”

STEP 3: Performing sentiment analysis

From the given positive and negative words, calculating positive and negative scores. Further using them to calculate Polarity score and subjectivity score.

STEP 4: Calculating all the metrics

Calculating the values for using their given formulae

1. AVG SENTENCE LENGTH
2. PERCENTAGE OF COMPLEX WORDS
3. FOG INDEX
4. AVG NUMBER OF WORDS PER SENTENCE
5. COMPLEX WORD COUNT
6. WORD COUNT
7. SYLLABLE PER WORD
8. PERSONAL PRONOUNS
9. AVG WORD LENGTH

STEP 5 : Exporting results to output.xlsx file

Write Data to Excel:

* Use pandas to store and save the results

**Technologies & Libraries Used**

* **Web Scraping:** requests, BeautifulSoup
* **Text Processing:** nltk, re, string
* **Data Handling:** pandas