Instructions to use the file –

1. Requirements: here is the list of dependencies you need to install in your laptop environment so that the code runs smoothly without generating errors.
   1. Pandas
   2. Beautiful Soup
   3. Requests
   4. Html5lib
   5. nltk
   6. Textstat
   7. Textblob

Install the required dependencies using your command line prompt.

1. **DATA EXTRACTION:** run the data\_extraction.py file in any desired python-supported IDE to perform the data extraction on the input file.
   1. Import the excel file ‘Input.xlxs’ in your IDE and run the code to achieve the desired output.
   2. Save the output in your device.
2. **DATA ANALYSIS:** run the text\_analysis.py file for performing text analysis on the article texts. After the above steps are completed, kindly fill the required metrics in the ‘output data structure' file.
3. **Tools Used-** Anaconda Navigator and Jupyter Notebook, MS Excel, MS Word.