First, I converted data from input.xlsx into dataframe so i can access the data in a better way..

For this solution,in the function create\_files(), I used BeautifulSoup library to extract the data using web scraping and then find under which tag the article text is present and then convert the text into normal string.

Then i wrote the data in a new file under the Articles folder using file.write and saved it for each

Then i created a function called stopwords\_finder() which reads data from all the files inside the stopwords folder provided by blackcoffer and then stored all the contents inside stpwords variable.

Then i created a function called pos() which reads data from the positive words file provided by blackcoffer and then stored all the contents inside positive variable.

Then i created a function called neg() which reads data from the negative words file provided by blackcoffer and then stored all the contents inside negative variable.

I also created a function called count\_syllables which finds the number of syllables in a word passed to it.

Then coming to the main part of the program that is the function sentimental\_analysis(), i first opened each and every file inside the Articles/ directory to perform sentimental analysis on it.

After converting all the text in lowercase and removing all the punctuations, i stored it inside the text variable then i used nltk.word\_tokenize() and stored the tokenized words in tokenized\_words variable.

After removing all the stopwords provided by blackcoffer, i stored the remaining text in words variable.

We then find the number of positive and negative words in the article with reference to master dictionary provided by blackcoffer to find the positive score and negative score of the text. Then we can find polarity and subjectivity easily by following the formula.

To find the number of complex words we check i f the word contains more than two syllables using the count\_syllables function created by me. Then we find the fog\_index and related terms easily. Then we find the number of pronouns using regex and stored it inside matches. Then we found all the other requirement using simple formulas.

I stored all this data inside list and at the end i converted it into data frame and then into excel file and exported it !!

**Dependencies required -** pandas,requests, nltk,bs4,BeautifulSoup,os,string,re,nltk.download('punkt'),

nltk.download('stopwords'), nltk.corpus,nltk.tokenize