**Instructions:**

The deliverables should contain

1. This instructions file
2. Input excel file
3. Output data structure excel file
4. Output\_Populator python file
5. Text\_File\_Generator python file
6. Trie\_Base python file
7. Utils python file
8. Text\_Files Folder
9. logs Folder

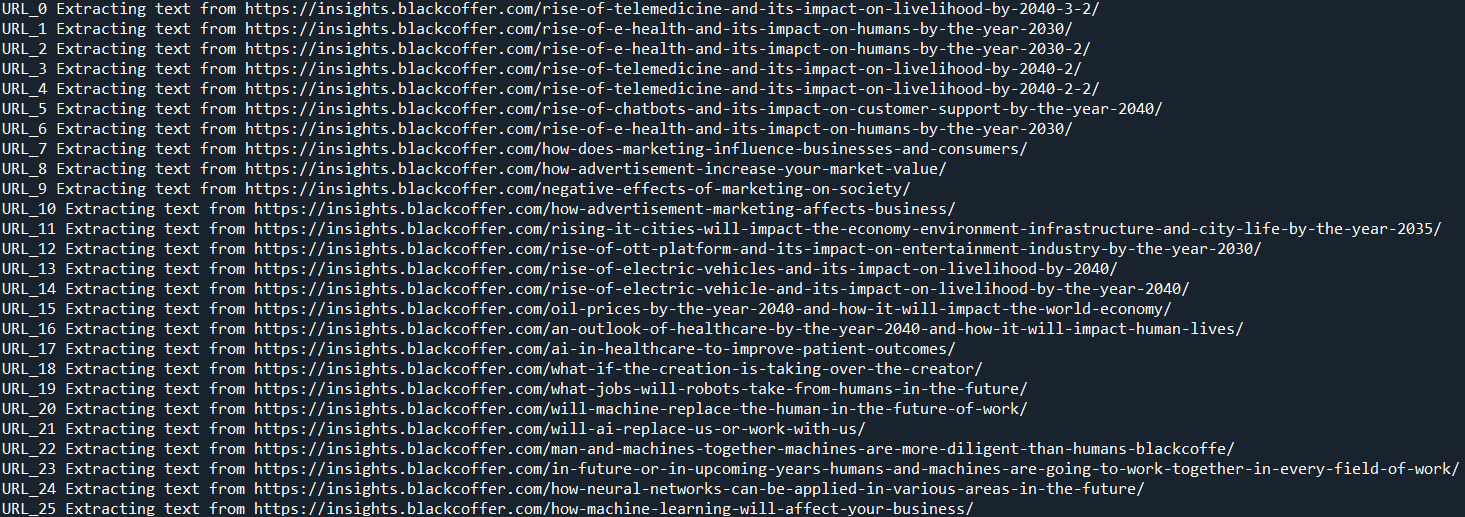
Go inside the deliverables folder and just execute the **Output\_Populator.py** file.

This automatically, generates,

1. the text-files which have the text generated from the URLs present in the input file into the **Text\_Files** Folder.
2. and later performs the analysis
3. and populates the output file too.

Note: You just need to execute **Output\_Populator.py** file. No other file needs to be executed. Every other file will be called from inside of this file.

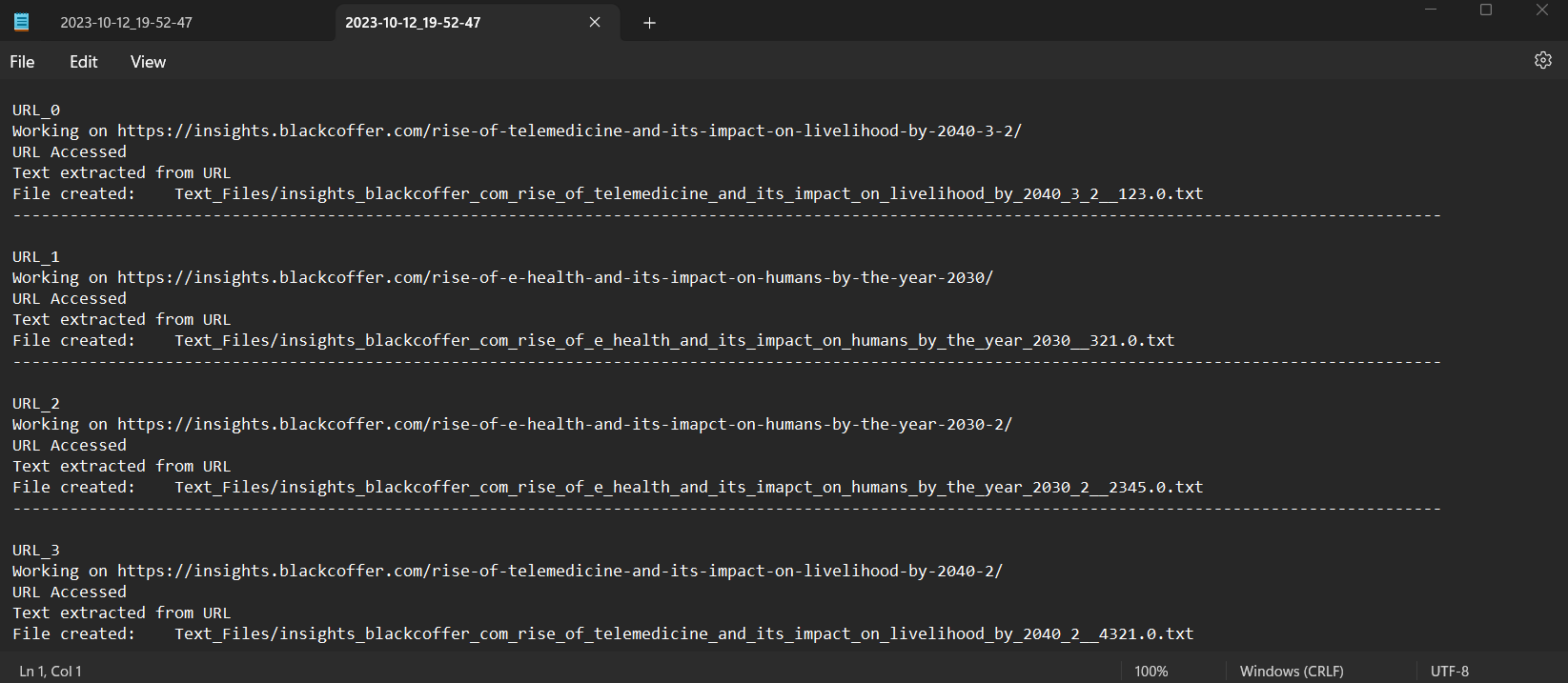
The console will look like below when the code is executing:



In **Trie\_Base python file**, I have created 3 Tries one each for Stop\_Words, Positive\_Words, Negative\_Words. I have chosen tries because using the **trie makes the search operations faster** when we are dealing with huge amounts of text

In Utils, there are modules of general utilities like making logs and Trie operations and cleaning words using stop word module of nltk, counting syllables etc.

You can check what happened with each URL by going inside the logs folder and checking the new log formed. For example the sample entry in log file will look like,



You can easily search what happened for a particular URL by using its index

Check the Output Data Structure for knowing the final analysis and the output variables which we got for each of the URL