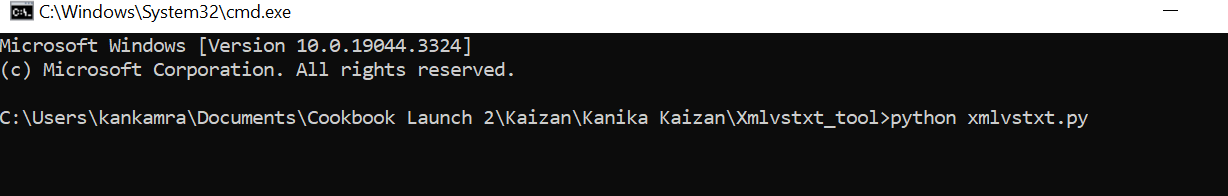
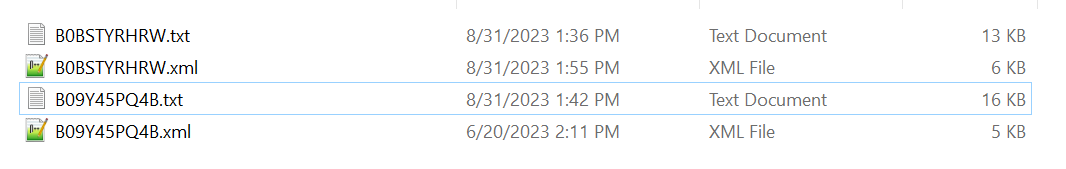
**XML vs TEXT COMPARISON Tool**

**Steps to Run the tool:**

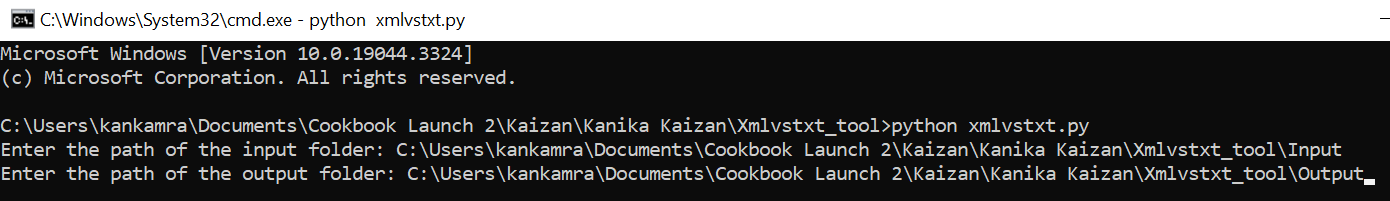
1. Download the XML and Text File Comparison Tool.
2. Extract the contents and ensure you have the necessary libraries installed: 'bs4', 'difflib', 'pandas', 'lxml'(Use pip install bs4, pip install difflib and pip install pandas, pip install lxml respectively).
3. Run the script (Open Command Prompt> type python xmlvstxt.py and Press Enter). OR If you are running .exe file, Double tap on the .exe file.



1. Then provide the input folder path containing the .txt and .xml files you want to compare as shown below and Press Enter. Note: You can multiple xml & txt files at once.



1. After that, provide an output folder path where the results will be stored and Press Enter.



**Important Points to Remember:**

1. Ensure there are no spaces in the folder names and file names within the input folder.
2. If any missing headers are detected in the text files, they will be logged with their locations in an Excel file for further analysis.
3. If there are any duplicate headers being detected in the text file, no update will be made, but the instances will be logged with the locations in the excel file.

**Special Character Replacement Tool**

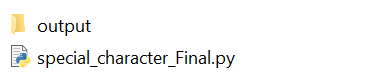
**Steps to Run the tool:**

* Create a folder and insert files(xml, txt, final.xml)  where special characters are available.
* Make sure you have the necessary libraries installed: 'bs4', 'difflib', 'pandas', 'lxml'(Use pip install bs4, pip install difflib and pip install pandas, pip install lxml respectively).
* Open cmd prompt and enter the command:

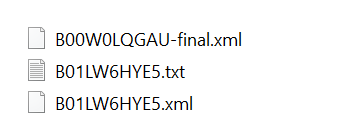
python special\_character\_Final.py -e C:\Users\guntupak\Desktop\test\3



* Once you run the tool, automatically output folder will be generated in the path where you pasted .py file(special\_character\_Final.py)



* If you open output folder it will have txt, xml and final.xml file with recipe headers where special characters are available.



* To check the files where headers got replaced go to the path where you have given in cmd prompt. Once open them in notepad++ if you have closed them, or else if files are already open just refresh it, so all the files get updated with html codes.





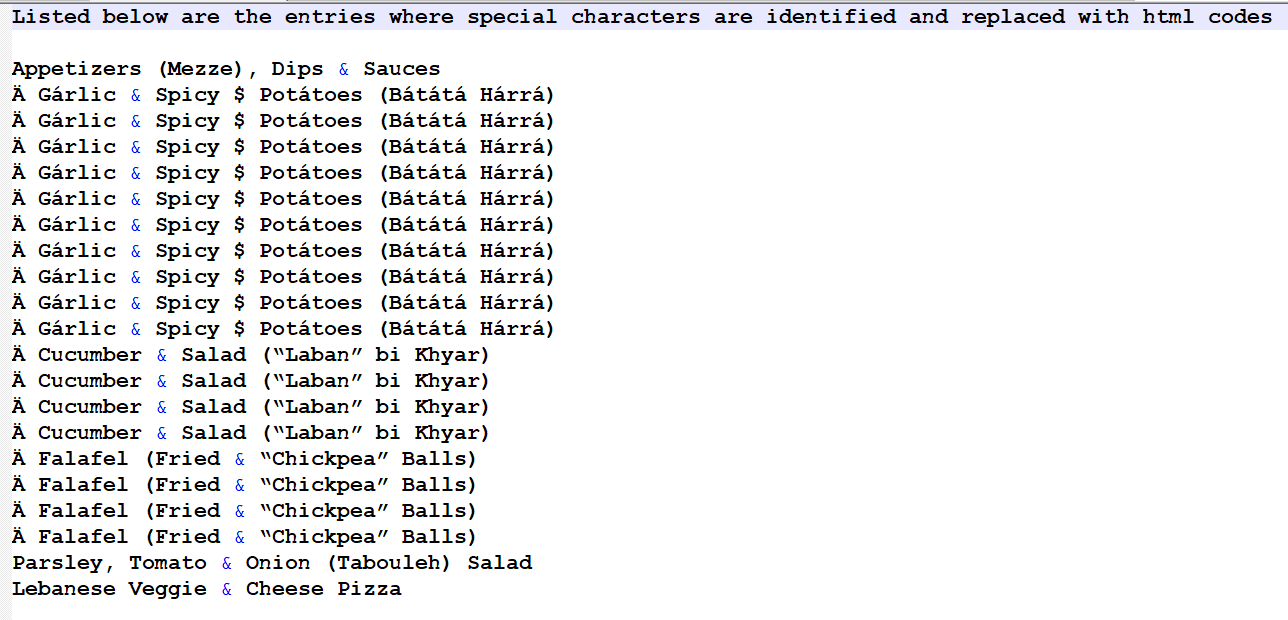








* Output folder shows the headers for special characters. Single recipe header is showing multiple times because it counts no. of special characters available in single header. It notifies us in the first line saying that below entries are special characters in all the files.



* If you run same Asin again, new instances will be available in the same file after the previous headers. If you run a new Asin, files will be available in output folder by differentiating with Asin name.