INSTRUCTIONS FOR CODE

By Sohan Mishra

Ph: 8939142489

email:sohanm226@gmail.com

Instructions:

1. The submission contains a jupyter source file that contains the source code (.ipynb) extension and a .py file

2. To check the submission either run each of the code snippets on machine or check the output that has been obtained by me.

3. Read the comments and the markdowns to get better understanding of each function.

4. Run the Taskfile.py file to validate the output.

5. Check out the output data structure_1.xlsx file that contains the final output excel file, containing the requested output variables as given in instructions.

6. Please note that there were some rows that have "This para is empty as URL showed error" ...this was due to the reasons that during the execution of web crawler, the URLs produced error 404.

Please Note that the .ipynb file has been attached for convenient code verification