READ ME

INDEX FOR THE FILES ATTACHED:

Please refer to the below list for information about all the attached files:

|  |  |  |
| --- | --- | --- |
| FILE NAME | DOC TYPE | DESCRIPTION |
| DATA228\_Group4\_CLV\_Project\_Report | pdf | Project documentation |
| DATA228\_Group4\_CLV\_Project\_Workbook | ipynb | Project Jupyter Notebook |
| DATA228\_Group4\_Project\_Presentation | ppt | Project Presentation |
| online\_retail\_file\_1, online\_retail\_file\_2 | csv | Present in google drive, link given below |

SOURCE FILE:

The dataset used (“online\_retail\_file\_1”, ”online\_retail\_file\_2”) is uploaded into google drive. Can be accessed using the link below:

<https://drive.google.com/drive/folders/1MKiexcKNo88I1r0_Pk3ddWsrsjY-Ifpy?usp=sharing>

INSTRUCTIONS FOR RUNNING THE JUPYTER NOTEBOOK:

Please read the following before proceeding with the execution.

Installations of below packages are necessary

pip install xgboost

pip install lifetimes

pip install plotly

pip install pyspark==3.2.0