

EDA and FE will be done on Exported Data

Handle missing values, remove duplicate records,and data formatting like string to numeric data or date time data,

After reading data we need to validate whether all the things present as we discussed earlier like No of colums and data types and is there missing values and if yes is it in permissible limit,

Data reading from shared address or server where data kept by client

1. Validating batch file which received from client
2. Validating file name as per format
3. If validation unsuccess will move file to bad data folder if success move to good data folder