Data Intake Report

Name: Data Collection Pipeline
Report date: 31/01/2023
Internship Batch: LISUM15
Version:2.0
Data intake by abdulrahmanafifi33@gmail.com
Data intake reviewer:<intern who reviewed the report>
Data storage location: [AfifiGhost2000/DataCollectionPipelineProject (github.com)](AfifiGhost2000/DataCollectionPipelineProject)

Tabular data details:

Total number of observations	3
Total number of files	1
Total number of features	9
Base format of the file	xls
Size of the data	6 KB

Note: Replicate same table with file name if you have more than one file.

Proposed Approach:
- Mention approach of dedup validation (identification)
- Mention your assumptions (if you assume any other thing for data quality analysis)

  My approach for dedup check is I called duplicated() function from pandas library on the pandas dataframe created from n google forms. And then called sum() to find total number of duplicates in our datframe. Fortunately, there was no duplicated items.

Note: Convert this doc in pdf and provide the link of pdf file in your dashboard.
     Please do not forget to remove this section while converting the file into pdf.