Data Intake Report

Name: G2M insight for Cab Investment firm
Report date: 13th September, 2022
Internship Batch: LISUM13:30
Version:
Data intake by: Jay Panara
Data intake reviewer:
Data storage location: https://github.com/jaypanara/Data_Glacier_Internsip.git

Tabular data details:

Cab_Data.csv

Total number of observations	359392
Total number of files	1
Total number of features	7
Base format of the file	.csv
Size of the data	20.2MB

City.csv

Total number of observations	20
Total number of files	1
Total number of features	3
Base format of the file	.csv
Size of the data	759 Bytes

Customer_ID.csv

Total number of observations	49171
Total number of files	1
Total number of features	4
Base format of the file	.csv
Size of the data	1MB

Transaction_ID.csv

Total number of observations	440098
Total number of files	1
Total number of features	3
Base format of the file	.csv
Size of the data	8.58MB

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)

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.