Data Intake Report

Name: G2M Case Study
Report date: 15-08-2025
Internship Batch: LISUM48
Version: 2.0
Data intake by: Peter Abban
Data intake reviewer:<intern who reviewed the report>
Data storage location: https://github.com/DataGlacier/DataSets

Tabular data details:
File name: Cab_Data

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

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)

Tabular data details:
File name: Customer_ID

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

Tabular data details:
File name: Transaction_ID

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

Tabular data details:
File name: City

Total number of observations	20
Total number of files	1file
Total number of features	3
Base format of the file	csv
Size of the data	size in B