

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 13th August 2024

Internship Batch: LISUM36

Version:1.0

Data intake by: Nonye Nweke

Data intake reviewer:<intern who reviewed the report>

Data storage location: [NonnyCodes/G2M-insights: G2M insight for cab investment firm \(github.com\)](#)

Tabular data details: Cab_data

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

Tabular data details: Customer_ID

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

Tabular data details: Transaction_ID

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

Tabular data details: City

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

Proposed Approach:

- There were no duplicates in the dataset.
- It is Assumed that for the age group young ranges from 18-29, middle age ranges from 30-44 and elderly ranges from 45-65
- It is assumed that for income group, low income ranges from 2000-9999, medium ranges from 10000 – 19999 and high 20000- 35000
- Mention your assumptions (if you assume any other thing for data quality analysis)