

Data Intake Report

Name: G2M Insight for Cab Investment Firm

Report date: 14/04/2023

Internship Batch: LISUM20

Version:<1.0>

Data intake by: Yashwanth Bharadwaj V

Data intake reviewer:<intern who reviewed the report>

Data storage location: <https://github.com/Yashu2699/DG-Week2>

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	19.2 MB

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	698 bytes

Tabular data details: Customer_Data

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	1.5 MB

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	10.1 MB

PROPOSED APPROACH:

- There were a few modifications made to the datasets, like changing the data type of users and population from object to int and changing the data type of date. There were no missing values, and the source seems to be authentic