Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 14- February -2023

Internship Batch: LISUM08

Version:

Data intake by: Chooladeva Piyasiri

Data intake reviewer: Data Glacier

Data storage location: *https://github.com/Chooladeva/Data-Science-Internship-at-Data-Glacier/tree/main/Week%202*

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	20.1 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	8.58 MB

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 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	4 KB

Proposed Approach:

- There are no missing values in any of the four datasets mentioned above.

- Merged all 4 datasets