

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 11/14/2024

Internship Batch: LISUM39

Version: 1.0

Data intake by: Madoka Fujii

Data intake reviewer: Data Glacier

Data storage location:

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	20.7MB

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	1 KB

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	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	8.78MB

Proposed Approach:

- Cleaning up and Joining datasets with MySQL
- EDA with Jupyter Notebook