

Data Intake Report

Name: G2M insight for Cab Investment Firm

Report date: 10/14/2023

Internship Batch: LISUM26

Version:<1.0>

Data intake by: Joseph Pang

Data intake reviewer: N/A

Data storage location: <https://github.com/MangoJoey/G2M-Insight-for-Cab-Investment-Firm.git>

Tabular data details: Cab_Data

Total number of observations	358879
Total number of files	1
Total number of features	7
Base format of the file	.csv
Size of the data	21.2 MB

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	1.1 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	9 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	759 bytes

Proposed Approach:

- Assumed that the outliers for price charge data were necessary and should not be removed because higher income customers may have purchased better vehicle/services
- No null values in any of the data

