

Data Intake Report

Name: G2M Insights for Cab Investment

Report date: July 26, 2024

Internship Batch: LISUM35

Version: 1.0

Data intake by: Sheldon Gordon

Data intake reviewer: N/A

Data storage location: N/A

Tabular data details:

Name: cab_data

Total number of observations	359392
Total number of files	
Total number of features	7
Base format of the file	.csv
Size of the data	19.2MB

Tabular data details:

Name: city

Total number of observations	20
Total number of files	
Total number of features	3
Base format of the file	.csv
Size of the data	612.0 B

Tabular data details:

Name: customer_ID

Total number of observations	49171
Total number of files	
Total number of features	4
Base format of the file	.csv
Size of the data	1.5 MB

Tabular data details:

Name: customer_ID

Total number of observations	49171
Total number of files	
Total number of features	4
Base format of the file	.csv
Size of the data	1.5 MB

Tabular data details:**Name: transaction_ID**

Total number of observations	440098
Total number of files	
Total number of features	3
Base format of the file	.csv
Size of the data	10.1 MB

Approach:

Check for duplicate rows and fields.

Perform inner merge to merge subset or common data.