

Data Intake Report

Name: Go-to-Market (G2M) insight for Cab Investment firm

Report date: 7/21/2022

Internship Batch: LISUM11

Version: 1.0

Data intake by: Akshar Chaklashiya

Data intake reviewer: Data Glacier Team

Data storage location: [GIT](#)

Tabular data details:

Table Name	Cab_Data
Total number of observations	359392
Total number of files	1
Total number of features	14
Base format of the file	.csv
Size of the data	20.1 MB

Table Name	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

Table Name	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

Table Name	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

Proposed Approach:

- No duplicate value present but few records from transaction table where associated user trips were not found, we drop those records.