

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 14 September, 2023

Internship Batch: LISUM25

Version:1.0

Data intake by: Sai Supriya Varugunda

Data intake reviewer: NA

Data storage location: [SupriyaVaru/Cab-Datasets: Cab-Datasets \(github.com\)](#)

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	20663 KB

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

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	1027 KB

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	8788 KB

Assumptions:

Date of Travel column in Cab_Data.csv represents ordinal values of days. So, we need to convert it into usable date format by using "DateOffset" function. Time period of data is from 31/01/2016 to 31/12/2018(1066 days) but we have 1095 unique values in Date of Travel column. So, we need to set base date as '1899-12-29' and time period of data would be 01/01/2016 to 30/12/2018.