

# Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 07/21/2023

Internship Batch: LISUM23

Version:1.0

Data intake by: Madoka Fujii

Data intake reviewer:

Data storage location:

## Tabular data details:Cab\_Data.csv

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	21MB

## Tabular data details:City.csv

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	1KB

## Tabular data details:Customer\_ID.csv

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	1MB

## Tabular data details:Transaction\_ID.csv

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	8.9MB

## Tabular data details:US\_Holiday\_Dates.csv

Total number of observations	342
Total number of files	1
Total number of features	6
Base format of the file	.CSV
Size of the data	16KB

**Proposed Approach:**

- Checked if there is any missing data in the all of the files.
- Cleaned the data with correct data types such as date and object.
- Eliminated comma and space.
- Merged the all data.