

# Data Intake Report

Name: G2M Insight for Cab Investment Company

Report date: 01/11/2022

Internship Batch: LISUM17

Version:1.0

Data intake by: Shruthi Madgi

Data intake reviewer: Data Glacier

Data storage location:

<https://github.com/shru0405/DataGlacierInternship/tree/main/Week%202/Data%20For%20EDA>

## Tabular data details:

### Data Taken: Cab

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	21.8 MB

### Data Taken: Customer

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

### Data Taken: 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	4 KB

### Data Taken: Transactions

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

**Data Taken: US Holidays**

<b>Total number of observations</b>	3288
<b>Total number of files</b>	1
<b>Total number of features</b>	10
<b>Base format of the file</b>	csv
<b>Size of the data</b>	2.43 KB

**Proposed Approach:**

- No missing values in any of the data.
- Merged all the data into a master dataset using Data frame.