

# Data Intake Report

Name: G2M insight for Cab Investment firm

Internship Batch: LISUM13: 30

Version:1.0

Data intake by: Rayan Yassminh

Data intake reviewer:-

Data storage location: <https://github.com/yassmin1/Data-Science-Intern-at-Data-Glacier/Week2>

## Tabular data details:

### Cab\_Data

<b>Total number of observations</b>	359392
<b>Total number of files</b>	1
<b>Total number of features</b>	7
<b>Base format of the file</b>	csv
<b>Size of the data</b>	19.2+MB

### City

<b>Total number of observations</b>	20
<b>Total number of files</b>	1
<b>Total number of features</b>	3
<b>Base format of the file</b>	csv
<b>Size of the data</b>	608.0+ bytes

### Customer ID

<b>Total number of observations</b>	49171
<b>Total number of files</b>	1
<b>Total number of features</b>	4
<b>Base format of the file</b>	csv
<b>Size of the data</b>	1.5+ MB

### Transaction\_ID

<b>Total number of observations</b>	440098
<b>Total number of files</b>	1
<b>Total number of features</b>	3
<b>Base format of the file</b>	csv
<b>Size of the data</b>	10 MB

## Proposed Approach:

- The csv file was joined into one csv file based on the Primary ID. No duplicated records.
- Time of Travel was converted to datetime format.