

# Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 14/02/2024

Internship Batch: LISUM30

Version: 1.0

Data intake by: Konstantinos Soufleros

Data intake reviewer: Data Glacier

Data storage location: [https://github.com/kostas696/DG\\_Intern](https://github.com/kostas696/DG_Intern)

## 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>	20.1 MB

## Tabular data details: 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>	759 bytes

## Tabular data details: 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>	1.00 MB

## Tabular data details: 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>	8.58 MB

**Tabular data details: us-holiday-dates-2004-2021**

<b>Total number of observations</b>	342
<b>Total number of files</b>	1
<b>Total number of features</b>	6
<b>Base format of the file</b>	.csv
<b>Size of the data</b>	15.7 kB

**Tabular data details: us-weather-events**

<b>Total number of observations</b>	8627181
<b>Total number of files</b>	1
<b>Total number of features</b>	14
<b>Base format of the file</b>	.csv
<b>Size of the data</b>	1.09 GB

**Proposed Approach:**

- There are no missing values in the data sets