

# Data Intake Report

**Name:** G2M insight for Cab Investment firm

**Report date:** 14-April-2023

**Internship Batch:** LISUM20

**Version:**1.0

**Data intake by:** Salih Eren Yüzbaşıoğlu

**Data intake reviewer:** Data Glacier

**Data storage location:**

<https://github.com/SalihErenYzb/DataGlacierInternship/tree/newBranch/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>	20.1 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:** 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 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>	4 KB

**Tabular data details: Population**

<b>Total number of observations</b>	3235
<b>Total number of files</b>	1
<b>Total number of features</b>	20
<b>Base format of the file</b>	csv
<b>Size of the data</b>	185 KB

**Proposed Approach:**

- There is no missing values in all four datasets that mentioned above.