

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 2024-02-14

Internship Batch: LISUM30

Version: 1.0

Data intake by: Olena Panchenko

Data intake reviewer:

Data storage location:

https://github.com/helenapanchenko/DS_Internship_Glacier/tree/master/Week2

Tabular data details: Cab_Data

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	20.1 MB

Tabular data details: 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	1 KB

Tabular data details: Customer_ID

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

Tabular data details: Transaction_ID

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

Tabular data details: US_Holidays

Total number of observations	55
Total number of files	1
Total number of features	6
Base format of the file	.csv
Size of the data	3 KB

Assumptions:

- **There are no missing data in all files listed above.**
- **We assume the data is full or incomplete with preserved data proportion and distribution.**