Data Intake Report

Name: G2M insight for Cab Investment firm
Report date: 12-June-2022
Internship Batch: LISUM10
Version: 1.0
Data intake by: Efe KARASIL
Data intake reviewer:
Data storage location: https://github.com/ekarasil/DataGlacier-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: 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: 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: 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

Tabular data details: US Holiday Dates (2004-2021)

Total number of observations	342
Total number of files	1
Total number of features	6
Base format of the file	.csv
Size of the data	16 KB

Proposed Approach:
- There was no duplicate of data in any of the datasets
- There was no missing values in all five datasets that mentioned above