Data Intake Report

Name: G2M Insight for Cab Investment Firm
Report date: 2/8/2022
Internship Batch: LISUM06
Version:
Data intake by: Catherine G Sanda
Data intake reviewer:
Data storage location: https://github.com/cate6495/cab

Tabular data details: cabdata

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	19.2 MB

Tabular data details: citydata

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	608.0 bytes

Tabular data details: customerdata

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	1.5MB

Tabular data details: transactiondata

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	10.1MB

Proposed Approach:
- There are no null values in all the 4 datasets
- Outliers are ignored