Data Intake Report

Name: G2M insight for Cab Investment firm
Report date: 04/09/2021
Internship Batch: LISUM03
Version: 1.0
Data intake by: Motamen MohammedAhmed
Data intake reviewer:<intern who reviewed the report>
Data storage location: GitHub: https://github.com/DataGlacier/DataSets

Tabular data details:

File name	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	21.5 MB

File name	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	759 B

File name	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	1.00 MB

File name	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

Assumptions:
- We assume that the users feature in City.csv file consists of Yellow Cab, Pink Cab users, and other companies. This is proven by finding the number of male and female customers in each city and comparing these numbers with the users. Sum of males and females in each city is less than users.
- Customer ID and Transaction ID attributes are anonymous; they don't present any useful information.