Data Intake Report

Name: G2M insight for Cab Investment firm
Report date: 11/11/2023
Internship Batch: LISUM25
Version:1.0
Data intake by: Farnaz Amiri

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	20.663 MB

Proposed Approach:
- No Duplicated row, No missing values
- Feature "Date of Travel" type is not consistent with expectations, so we need to convert it to meaningful date objects.

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	1.027 MB

Proposed Approach:
- No Duplicated row, No missing values
- Data adheres to the expected format and structure.

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	8.788 MB

Proposed Approach:

- No Duplicated row, No missing values
- Data adheres to the expected format and structure.

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

Proposed Approach:
- No Duplicated row, No missing values
- Features "Population" and "Users" data types are not consistent with expectations, so we need to convert them to integer.