Data Intake Report

Name: G2M Case Study
Report date: 14th July 2022
Internship Batch:LISUM11
Version:<1.0>
Data intake by: sAMIAT bOLA-mATANMI
Data intake reviewer:<intern who reviewed the report>
Data storage location: <location URL eg: github, cloud>

Tabular data details:

Total number of observations	359392
Total number of files	4
Total number of features	7
Base format of the file	.csv
Size of the data	23MB
File Name	Cab_data

Total number of observations	20
Total number of files	4
Total number of features	20
Base format of the file	.csv
Size of the data	759 Bytes
File Name	City_data

Total number of observations	49171
Total number of files	4
Total number of features	4
Base format of the file	.csv
Size of the data	1.1 MB
File Name	Customer ID

Total number of observations	440098
Total number of files	4
Total number of features	3
Base format of the file	.csv
Size of the data	9 MB
File Name	Transaction ID

Proposed Approach:

Approach using in this data analysis was understanding of data first, data cleaning and then data analysis. The data was first read using pandas and cleaning was done using pandas built in libraries and analysis was also done using pandas.