Data Intake Report

Name: G2M insight for cab investment
Report date: 12/06/2022
Internship Batch: *LISUM10:30*
Version:1.0
Data intake by: Wangu Ndungu
Data intake reviewer: Wangu Ndungu
Data storage location: https://github.com/DataGlacier/DataSets

Tabular data details:

Cab_Data

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	20663 KB

City

Total number of observations	20
Total number of files	4
Total number of features	3
Base format of the file	.csv
Size of the data	1 KB

Customer_ID

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	1027 KB

Transaction_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	8788 KB

Proposed Approach:

- Use of Exploratory Data Analysis preceded by thorough data wrangling to draw appropriate conclusions from the data
- Assumptions: Profit is the equivalent of the difference between 'Price Charged' and 'Cost of Trip'