

Data Intake Report

Name: G2M insight for Cab Investment Firm

Report date: 04-May-2023

Internship Batch: LISUM21

Version:1.0

Data intake by: Ayushi Malaviya

Data intake reviewer:

Data storage location:

<https://github.com/Ayushimalaviya/VC/tree/Project/scripts/week2>

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	21.9 MB

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	4KB

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	1.1MB

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	9MB

Tabular data details: us_holidays

Total number of observations	3288
Total number of files	1
Total number of features	10
Base format of the file	.CSV
Size of the data	191KB

Proposed Approach:

Exploratory Data Analysis [derived features, scaled data, statistical analysis]

Assumptions

1. Average Profit by Company Based on Weekdays.
2. Average Profit by Company Based on month.
3. Average Profit by Company Based on year.
4. Users on holidays travelling cab
5. Average profit by Company Based on City Area
6. Profit per city per KM by each company
7. Total Rides by user per year for each company
8. Customer retention analysis for each company