

Data Intake Report

Name: G2M Case Study

Report date: 11/17/2023

Internship Batch: LISUM27

Version: 1.0

Data intake by: Doruk Dincer

Data intake reviewer:

Data storage location:

https://github.com/dorukdincer/DG_Virtual_Internship/tree/main/Week%202

Tabular data details: Cab_Data.csv

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	19.2 MB

Tabular data details: City.csv

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	612.0 Bytes

Tabular data details: Customer_ID.csv

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	1.5 MB

Tabular data details: Transaction_ID.csv

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	10.1 MB

Proposed Approach:

- Merged all the related data into one master data
- Altered some of the types of how data stored to a more easier to analyze version