Data Intake Report

Name: Go-to-Market Case Study
Report date: 14/05/2023
Internship Batch: LISUM21
Version: 1.0
Data intake by: Bogdan-Remus Pintilie
Data intake reviewer:
Data storage location:

Tabular data details:

Total number of observations	359392
Total number of files	1 (Cab_Data)
Total number of features	7
Base format of the file	.csv
Size of the data	20663KB

Total number of observations	49171
Total number of files	1 (Customer_ID)
Total number of features	4
Base format of the file	.csv
Size of the data	1027KB

Total number of observations	440098
Total number of files	1 (Transaction_ID)
Total number of features	3
Base format of the file	.csv
Size of the data	8788KB

Proposed Approach:
- I used a data tool in Excel called Remove Duplicates and selected the join key range of values. The returned result was that there are no duplicates. Since I created a master data ensuring the join key has no duplicates is sufficient to use the data for analysis.
- It is assumed that the data has been correctly entered, recorded, and processed.
- It is assumed there are no variations or inconsistencies in the way data is recorded.
- I assume the data is free from bias.