Data Intake Report

Name: G2M insight for Cab Investment firm
Report date: 9/12/24
Internship Batch: LISUM37
Version:1.0
Data intake by:Furkan Ay
Data intake reviewer:Data Glacier

Tabular data details: Transaction_ID

Total number of observations	440099
Total number of files	1
Total number of features	3
Base format of the file	csv
Size of the data	8.58 MB

Tabular data details: Customer_ID

Total number of observations	49172
Total number of files	1
Total number of features	4
Base format of the file	csv
Size of the data	1 MB

Tabular data details: Cab_Data

Total number of observations	35393
Total number of files	1
Total number of features	7
Base format of the file	csv
Size of the data	20.1 MB

Tabular data details: City

Total number of observations	21
Total number of files	1
Total number of features	3
Base format of the file	csv
Size of the data	4 KB

Proposed Approach:

       My approach going into this data was to analyze the several given data. I approached this data in this way by first combining the data sets by using pandas.merge function. After combining these datasets I plan to analyze the data and gain insights. I will look into analyzing the profits of the companies and hypothesis testing to confirm my ideas and hypothesis. I will also use pandas to make reports making visualizations of the data further allowing me to make insights about the data. I plan to hypothesize on the age compared to profits, the two companies profits and compare gender profits. These ideas and insights will later be put on a presentation to assert these ideas to the VenturePath Investments.