Data Intake Report

# Name: G2M insight for Cab Investment firm

Report date: 08.05.2023

Internship Batch: LISUM20

Version:<1.0

Data intake by: Berna Gündoğdu Özbaşaran

Data intake reviewer: -

Data storage location: https://github.com/gunberna/Final\_project/upload/main

**Tabular data details:**

**Cab\_Data.csv**

|  |  |
| --- | --- |
| **Total number of observations** | 1218 rows |
| **Total number of files** | <Number of files received> |
| **Total number of features** | 12 columns |
| **Base format of the file** | .csv |
| **Size of the data** | 52 kb |

**Description of Problem:**

As retailers and stores need customers to survive, the amount of clients are very crucial for their strategy. For instance, if sales are low one month, the retailer must decide if this is due to some external forces like the status of economy or whether something is wrong with the product.

**Timeline:**

Hence, as the fluctuations throughout a timetable are crucial, my aim will be to find some patterns and trends that could provide some insight to the retailer.

For this purpose I will firstly cleanse the data where I will figure out the missing values, outliers and what to do with them. This could be done with the pandas and matplotlib library where I will use both. I will then illustrate the descriptive statistics to see the feature of the data in hand. Also I will decide what part of the data might be useful for my analysis. Then I will establish a make three machine learning models. This will give me the insight to make recommendations. The final step is to portrait my suggestions and analysis on the presentation I will make.