

# Data Intake Report

Name: Week 2 “G2M Marketing Insights for the Pink and Yellow Cab Companies”

Report date: 12/13/2022

Internship Batch: LISUM16

Version: 1.0

Data intake by: Olivia Foster

Data intake reviewer: Data Glacier

Data storage location: <location URL eg: github, cloud>

## Tabular data details: Cab\_Data.csv

<b>Total number of observations</b>	359392
<b>Total number of files</b>	1
<b>Total number of features</b>	7
<b>Base format of the file</b>	csv
<b>Size of the data</b>	19.2 MB

## Tabular data details: City.csv

<b>Total number of observations</b>	20
<b>Total number of files</b>	1
<b>Total number of features</b>	3
<b>Base format of the file</b>	csv
<b>Size of the data</b>	608 bytes

## Tabular data details: Customer\_Data.csv

<b>Total number of observations</b>	49171
<b>Total number of files</b>	1
<b>Total number of features</b>	4
<b>Base format of the file</b>	csv
<b>Size of the data</b>	1.5 MB

## Tabular data details: Transaction\_ID.csv

<b>Total number of observations</b>	440098
<b>Total number of files</b>	1
<b>Total number of features</b>	3
<b>Base format of the file</b>	csv
<b>Size of the data</b>	10.1 MB

**Note:** Replicate same table with file name if you have more than one file.

**Proposed Approach:**

- There are no missing values in any of the datasets.
- We assume that that all costs are accounted for in the costs column of the Cab\_Data dataframe.

**Note: Convert this doc in pdf and provide the link of pdf file in your dashboard.  
Please do not forget to remove this section while converting the file into pdf.**