

# Data Intake Report

Name: Cab Companies EDA

Report date: 14/10/2022

Internship Batch: LISUM14

Version: 1.0

Data intake by: Sebastian Bucheli

Data intake reviewer:

Data storage location:

## Tabular data details:

<b>File name</b>	Cab Data
<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>	20.7 MB

<b>File name</b>	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>	1 KB

<b>File name</b>	cities centroids
<b>Total number of observations</b>	20
<b>Total number of files</b>	1
<b>Total number of features</b>	4
<b>Base format of the file</b>	.geojson
<b>Size of the data</b>	5 KB

<b>File name</b>	Customer ID
<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 MB

<b>File name</b>	Transaction ID
<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>	8.8 MB

<b>File name</b>	US Holiday Dates (2004-2021)
<b>Total number of observations</b>	57
<b>Total number of files</b>	1
<b>Total number of features</b>	6
<b>Base format of the file</b>	.csv
<b>Size of the data</b>	16 KB

**Proposed Approach:**

- This data is referred to two companies of Cabs and their behaviors in 20 cities of USA.
- I assume that the geographical data is very important for the explanation of this data.