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:**

|  |  |
| --- | --- |
| **File name** | Cab\_Data |
| **Total number of observations** | 359392 |
| **Total number of files** | 1 |
| **Total number of features** | 7 |
| **Base format of the file** | .csv |
| **Size of the data** | 20.7 MB |

|  |  |
| --- | --- |
| **File name** | City.csv |
| **Total number of observations** | 20 |
| **Total number of files** | 1 |
| **Total number of features** | 3 |
| **Base format of the file** | .csv |
| **Size of the data** | 1 KB |

|  |  |
| --- | --- |
| **File name** | cities\_centroids |
| **Total number of observations** | 20 |
| **Total number of files** | 1 |
| **Total number of features** | 4 |
| **Base format of the file** | .geojson |
| **Size of the data** | 5 KB |

|  |  |
| --- | --- |
| **File name** | Customer\_ID |
| **Total number of observations** | 49171 |
| **Total number of files** | 1 |
| **Total number of features** | 4 |
| **Base format of the file** | .csv |
| **Size of the data** | 1 MB |

|  |  |
| --- | --- |
| **File name** | Transaction\_ID |
| **Total number of observations** | 440098 |
| **Total number of files** | 1 |
| **Total number of features** | 3 |
| **Base format of the file** | .csv |
| **Size of the data** | 8.8 MB |

|  |  |
| --- | --- |
| **File name** | US Holiday Dates (2004-2021) |
| **Total number of observations** | 57 |
| **Total number of files** | 1 |
| **Total number of features** | 6 |
| **Base format of the file** | .csv |
| **Size of the data** | 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.