Data Intake Report

Name: EDA Cab Data and US holiday dataset
Report date: 8 March 2021
Internship Batch:LISP01
Version:<1.0>
Data intake by: Afshan hashmi
Data intake reviewer:<intern who reviewed the report>
Data storage location: https://github.com/Inquisitive4Data/EDA_CAB-data

Tabular data details:
Cab_data.csv

Total number of observations	359392
Total number of files	4
Total number of features	11
Base format of the file	.csv
Size of the data	20.1MB

Transaction_ID.CSV

Total number of observations	44098
Total number of files	4
Total number of features	3
Base format of the file	.csv
Size of the data	8.58 MB

Customer_ID.csv

Total number of observations	49171
Total number of files	4
Total number of features	4
Base format of the file	.csv
Size of the data	1MB

City.csv

Total number of observations	19
Total number of files	4
Total number of features	2
Base format of the file	.csv
Size of the data	4 KB

US holidaydates(2004-21)

Total number of observations	342
Total number of files	1
Total number of features	6
Base format of the file	.csv
Size of the data	16 KB

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

Proposed Approach:
- Mention approach of dedup validation (identification)
- Mention your assumptions (if you assume any other thing for data quality analysis)