

# Data Intake Report

Name: Cab Investment  
Report date: 13/11/2022  
Internship Batch:LISUM15  
Version:<1.0>  
Data intake by: Roxana Virlan  
Data intake reviewer:<intern who reviewed the report>  
Data storage location: github

## Tabular data details:

<b>Total number of observations</b>	359392
<b>Total number of files</b>	1
<b>Total number of features</b>	12
<b>Base format of the file</b>	csv
<b>Size of the data</b>	35.6MB

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

## Proposed Approach:

- Data validation has been achieved by checking that the data type of variables is assigned correct. Data that was not useful and not compatible with the other data has been removed- City.csv dataset due to not representing the correct population of the cities for all three years.
- Data quality has been achieved by removing any duplicates, NA and outliers.