Data Intake Report

Name: Cab Data Analysis
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: < github

Tabular data details:

Total number of observations	359392
Total number of files	4
Total number of features	18
Base format of the file	.csv
Size of the data	1.98 MB

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

Proposed Approach:
- Joined all the 4 files together
- Drop the rows with null values
- Detect any outliers Present
- Splitted the date into day , month and year
- Normalized some of the Numeric column to get proper result
- Rename some of the columns as required
- Changed the Data Type as needed

3. EDA recommendation and hypothesis results

- Most of the travelling is done in the month of December around Christmas so Festive Offer could be given by the company to increase the no. of customer
- Yellow cab has much more customer share and reliability so it's better to invest in that
- Pink Cab has increased its female customer gradually from 2016 to 2018
- yellow cab can Provide some additional discount or facilities to attract more female customer
- New York followed by Silicon valley has the highest no. of cab users and Profit for both the company
- There are equal proportion of customer using cash and card in both the company