Data Intake Report

Name: XYZ Cab Investment
Report date: 12th Jan 2025
Internship Batch: LISUM41
Version: 1.0
Data intake by: Dinesh Saini
Data intake reviewer:
Data storage location: <location URL eg: github, cloud>

Tabular data details:

Total number of observations	359,392
Total number of files	1
Total number of features	7
Base format of the file	Cab_data.csv
Size of the data	21.8 MB

Total number of observations	20
Total number of files	1
Total number of features	3
Base format of the file	City.csv
Size of the data	759 bytes

Total number of observations	491,71
Total number of files	1
Total number of features	4
Base format of the file	Customer_ID.csv
Size of the data	1 MB

Total number of observations	440,098
Total number of files	1
Total number of features	3
Base format of the file	Transaction_ID.csv
Size of the data	8.58 MB

A private firm XYZ is planning for investment in the cab industry and as per their Go-to-market strategy they want to understand the market before making decisions.

Data Transformation:
- Created the profit column from price charged and cost of trip.
- Covert the age and distance column into bins.
- Convert the integer dates into corrected formatted dates.

Hypothesis:

- What is the market share for each Cab investment and what is their profit over time.
- Growth rate of each Cab Company for the given time.
- Which is the demand of each cab company change over time.
- Repeat customer based on their ride experience
- Distance-based analysis of the profit
- City based presence of the cab and their profit share
- Price charged vs cost of trip analysis based on the city to improve the profit share.