

Data Intake Report

Name: G2M Cab Investment Firm

Report date: June 23rd 2022

Internship Batch: LISUM10

Version:<1.0>

Data intake by: Mohammed Maqsood

Data intake reviewer:

Data storage location: https://github.com/Maqsood8/CabInvestment_EDA_Project.git

Tabular data details:

Total number of observations	359393
Total number of files	1
Total number of features	7
Base format of the file	.csv
Size of the data	20.6MB

Tabular data details: Cab_data

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	9MB

Tabular data details: Transaction_id

Total number of observations	49171
Total number of files	1
Total number of features	3
Base format of the file	.csv
Size of the data	1MB

Tabular data details: Customer_ID

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	1KB

Tabular data details: City

Proposed Approach:

- The datasets were all combined to create one master dataset.
- 5 features were derived from the datasets available:
- Month and Year: these were derived from the month_of_travel feature.
- Profit: this is the difference between price charged and cost_of_trip features
- Age_range: the ages of the customers were allocated to different bins.
- Percentage_users: this is a ratio in percentile of the users in each city to the population of that city.
- Exploratory Data Analysis approach utilized to draw insights from the data.
- The analysis has been divided into 5 parts: ➤ Data Exploration. ➤ EDA ➤ finding the most profitable Cab Company. ➤ Hypothesis Testing. ➤ Recommendations for investment.

Assumption:

- As we see yellow cab is much popular company in larger population, which may be inferred as a potential indicator of future market growth mixing with total user increase in each year. So it can be said that yellow cab has a bigger potential market comparing to pink cab.
- Yellow cab's average profit per KM is almost three times the average profit per KM of the Pink cab.
- Yellow cab is mostly liked company in larger population of male and female, Profit is also on increment each year in yellow cab comparing pink.
- In general, the Payment_Mode is card.
- 3 years data shows that the most profitable company was the Yellow Cab compared to the Pink Cab.
- The average profit contributed by people with different incomes taking cabs is nearly the same, But the average profit of Yellow Cab is much larger than that of Pink Cab.
- According to the detailed analysis, XYZ firm should invest in Yellow Cab. Given the losses, XYZ firm should invest more to Yellow Cab in the New York.
- Pink cab is still in transition of joining the market, as we can see smaller market share in comparison to Yellow Cab even the regular increase in total users each year. This idea can also be supported by pink cab's lower profit per KM, to serve as an incentive for future market disruption.