

Data intake report

Name: Insight for Cab Investment firm

Date: 3/3/2021

Report by: Nada Alzahrani

Problem:

XYZ is a private firm in US. Due to remarkable growth in the Cab Industry in last few years and multiple key players in the market, it is planning for an investment in Cab industry and as per their Go-to-Market(G2M) strategy they want to understand the market before taking final decision.

Dataset:

- Cab_Data
- Customer_ID
- Transaction_ID
- City

Data details:

- 17 Features
- Timeframe of the data : 31/01/2016 to 31/12/2018
- Total data points : 359392
- Format : csv

Steps:

- Merged Cab and City using "City" column
- Merged Transaction and Customer using "Customer ID" column
- Then merged 2 dataset using "Transaction ID" column
- Convert 'Date of Travel' using to_datetime
- Separated date to 3 columns (day, month and year)
- Create new column which is profit

EDA:

- Show statistical of each column
- Find null and duplicates values
- Graphs
-

Hypothesis:

- KM Travelled
- Price Charged
- Age
- Income (USD/Month)
- Profit