



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

G2M Case Study

21.08.2023

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Executive Summary

The Client

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.

Problem Statement

- Review the Source Documentation
- Understand the field names and data types
- Identify relationships across the files
- Field/feature transformations
- Determine which files should be joined versus which ones should be appended
- Create master data and explain the relationship
- Identify and remove duplicates
- Perform other analysis like NA value and outlier detection

EDA

EDA

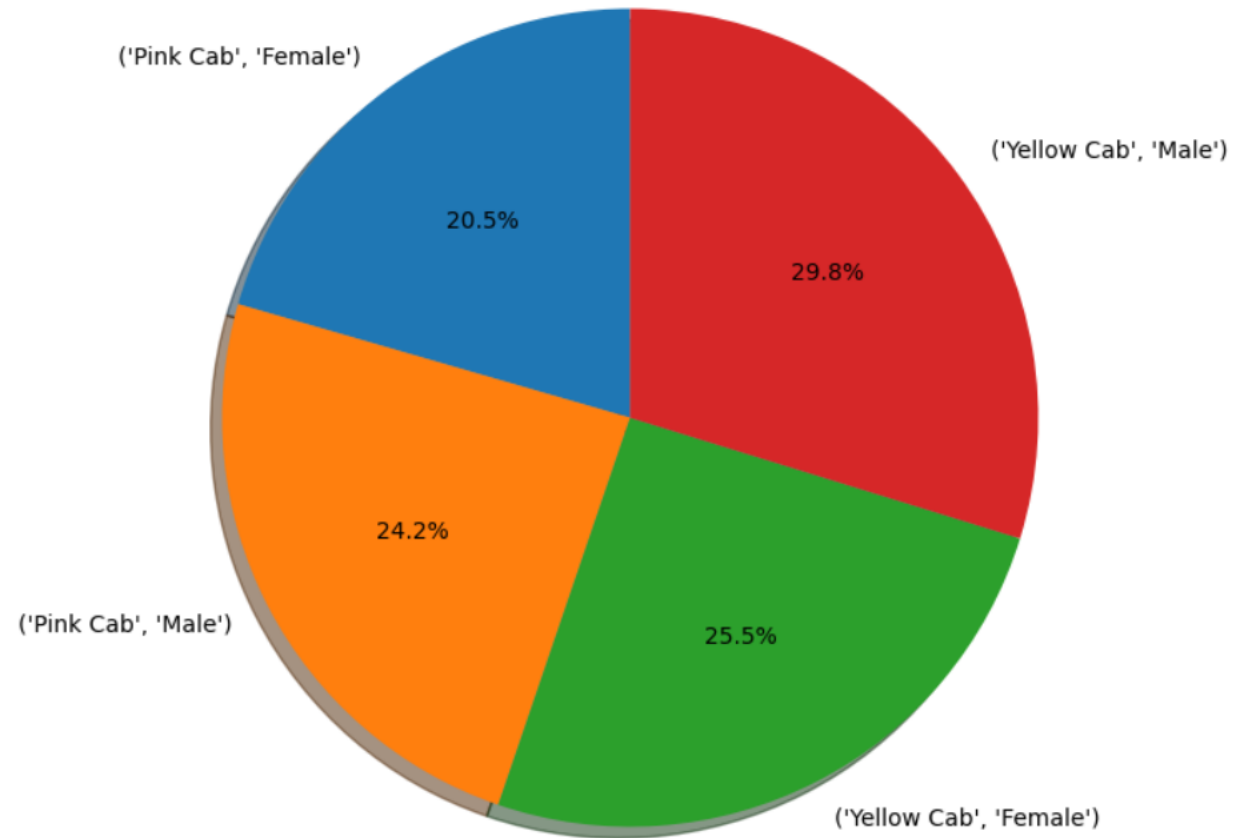
```
[88]: df= cab_df.merge(tran_df, on= 'Transaction ID').merge(cust_df, on ='Customer ID').merge(city_df, on = 'City')
df.head()
```

```
[88]:
```

	Transaction ID	Date of Travel	Company	City	KM Travelled	Price Charged	Cost of Trip	Customer ID	Payment_Mode	Gender	Age	Income (USD/Month)	Po
0	10000011	42377	Pink Cab	ATLANTA GA	30.45	370.95	313.6350	29290	Card	Male	28	10813	
1	10351127	43302	Yellow Cab	ATLANTA GA	26.19	598.70	317.4228	29290	Cash	Male	28	10813	
2	10412921	43427	Yellow Cab	ATLANTA GA	42.55	792.05	597.4020	29290	Card	Male	28	10813	
3	10000012	42375	Pink Cab	ATLANTA GA	28.62	358.52	334.8540	27703	Card	Male	27	9237	
4	10320494	43211	Yellow Cab	ATLANTA GA	36.38	721.10	467.1192	27703	Card	Male	27	9237	

EDA Summary

Customer share per gender per cab



Recommendations

According to the analysis in Section 3, we recommend XYZ company to invest in Yellow Cab.

Thank You