Exploratory Data Analysis

G2M Insight for Cab Investment Firm

21-06-2022

Agenda

- Executive Summary
- Approach
- EDA
- EDA Summary
- Recommendations

Executive Summary

- XYZ is a private equity firm in US. Due to remarkable growth in the Cab Industry in last few years and multiple key players in the market, XYZ is planning for an investment in Cab industry.
- XYZ has requested an investment recommendation, based on the data of two cab companies.
- Comparison Exploratory Data Analysis of all data related to the companies, cities, customers and transactions.
- Considering the company's profit, customer retention, city dominance, potential growth and total customer userbase, I recommend investing in the Yellow Cab Company.

Approach

- Basic Data Exploration
- Data Cleaning and Preprocessing
- Exploratory Data Analysis
- Hypothesis Testing
- Recommendation

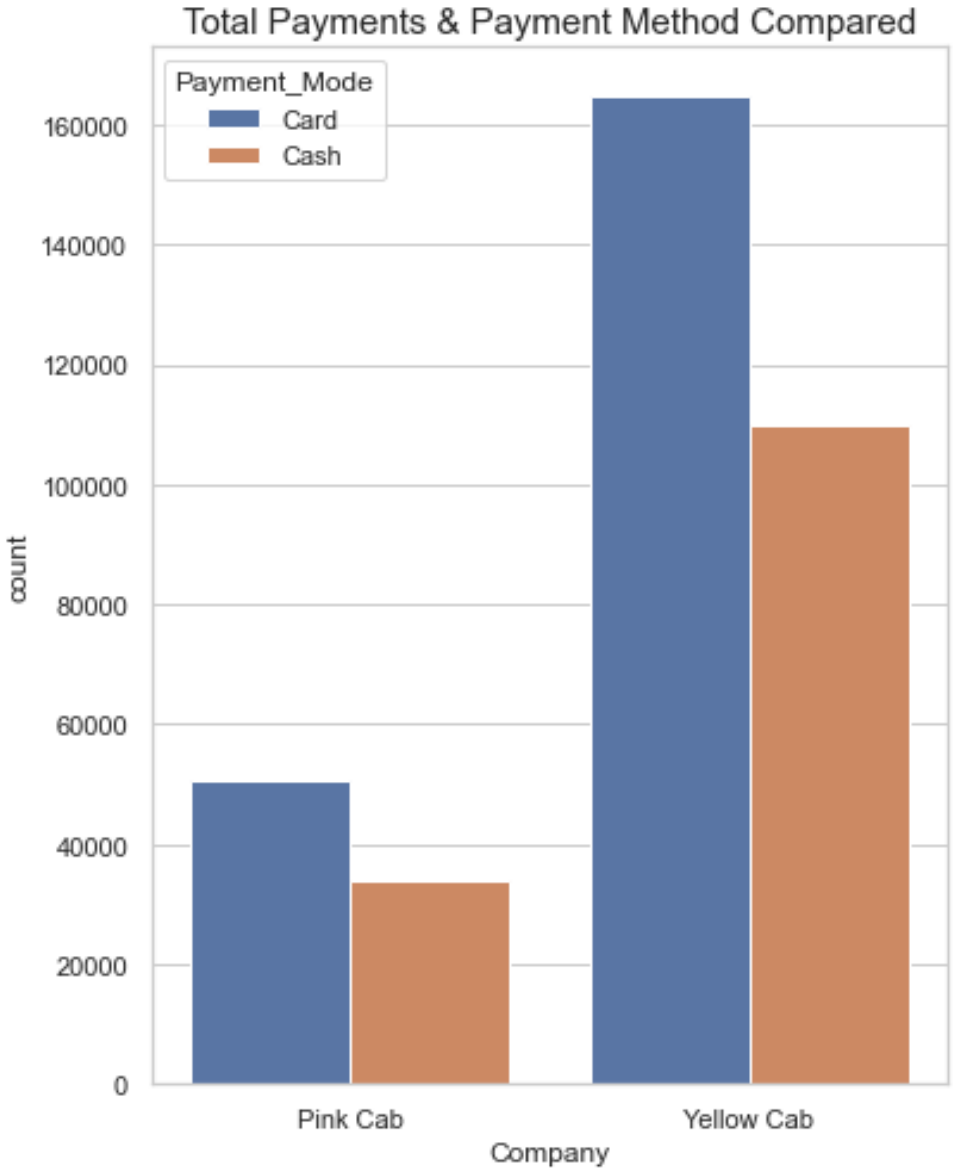
Exploratory Data Analysis

- Aim of the EDA is to extract useful insights that will ultimately help make an investment recommendation.

Payment Analysis

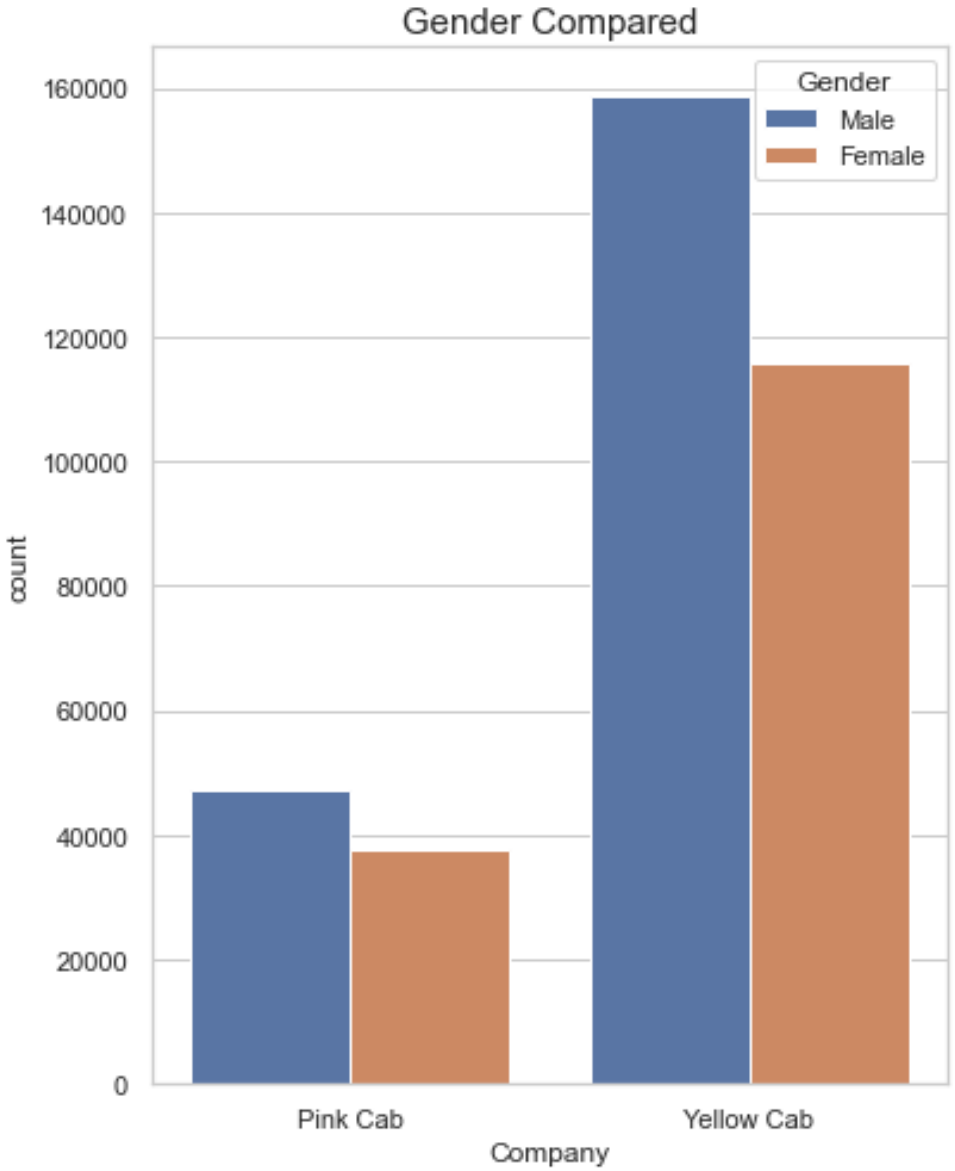
- **Yellow Cab**: Cards are most used as payment method by a large margin, also clearly has made more transactions in the same time period compared to Pink Cab.

- **Pink Cab:** Cards are most used as payment method. Although by a smaller margin compared to Yellow Cab.

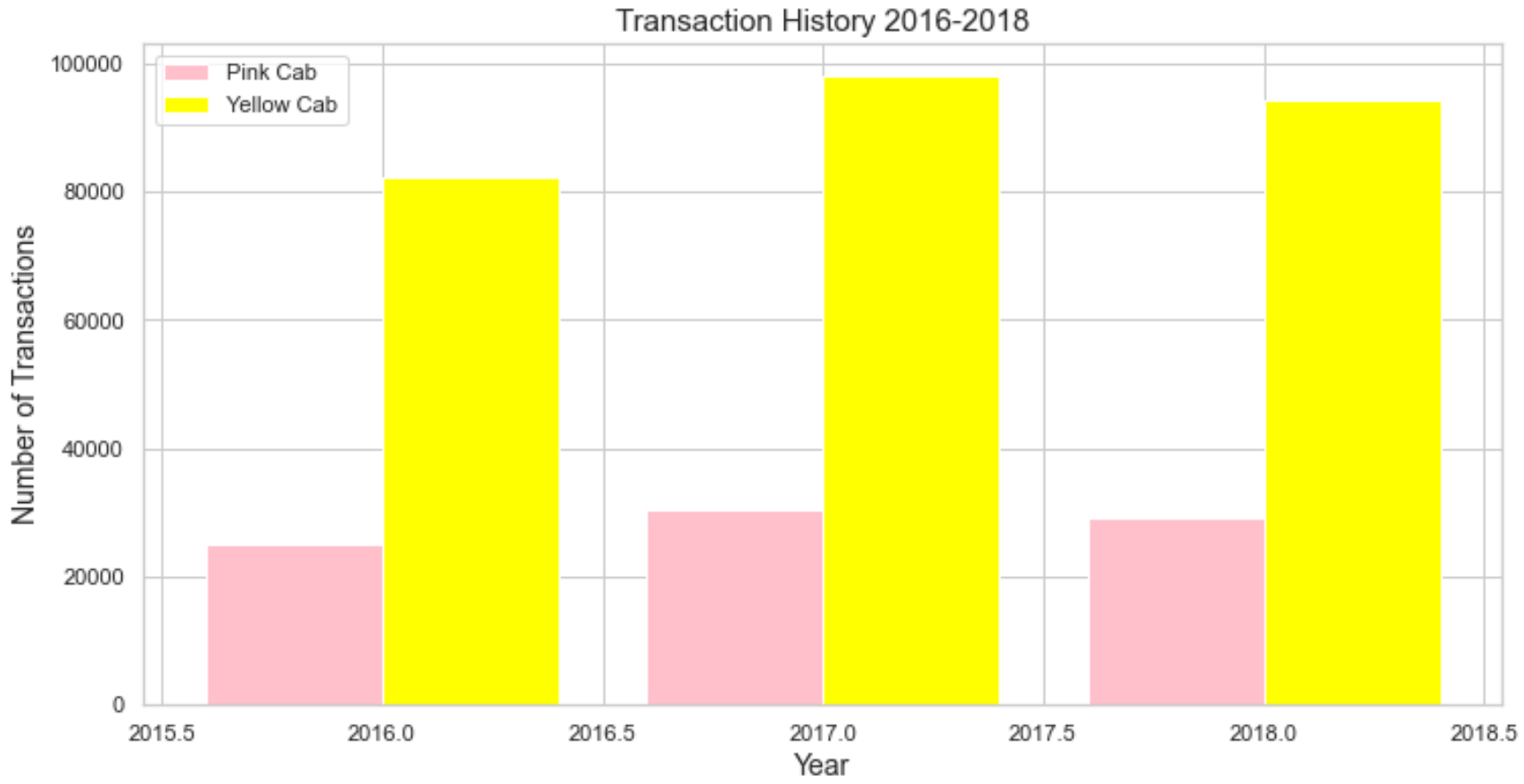


Gender Analysis

- **Yellow Cab**: Males are the most frequent gender using the yellow cab service.
- **Pink Cab:** Males are also the most frequent gender using the pink cab service, by a smaller margin. *(More balanced than Yellow Cab)*

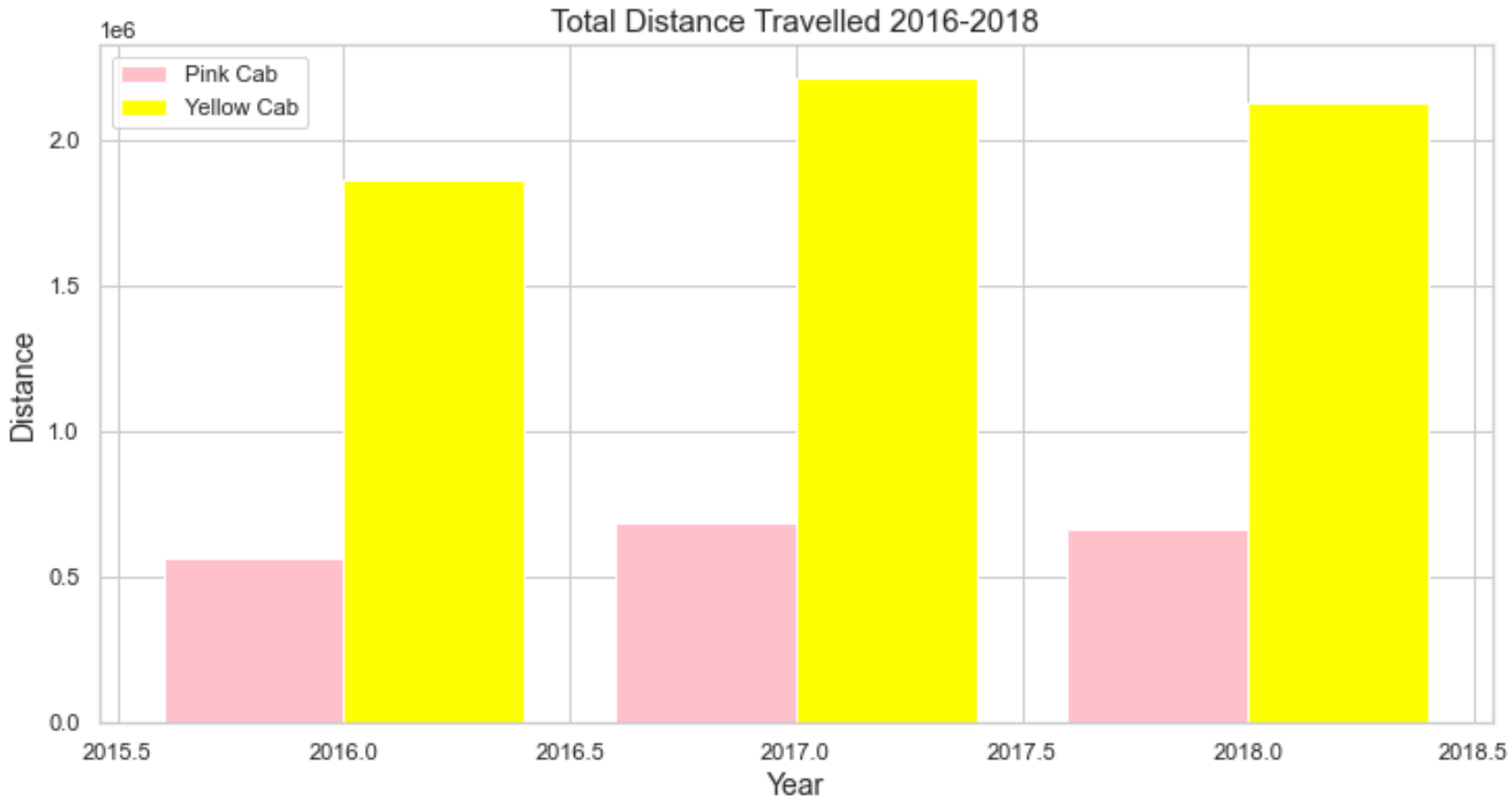


Transaction Analysis



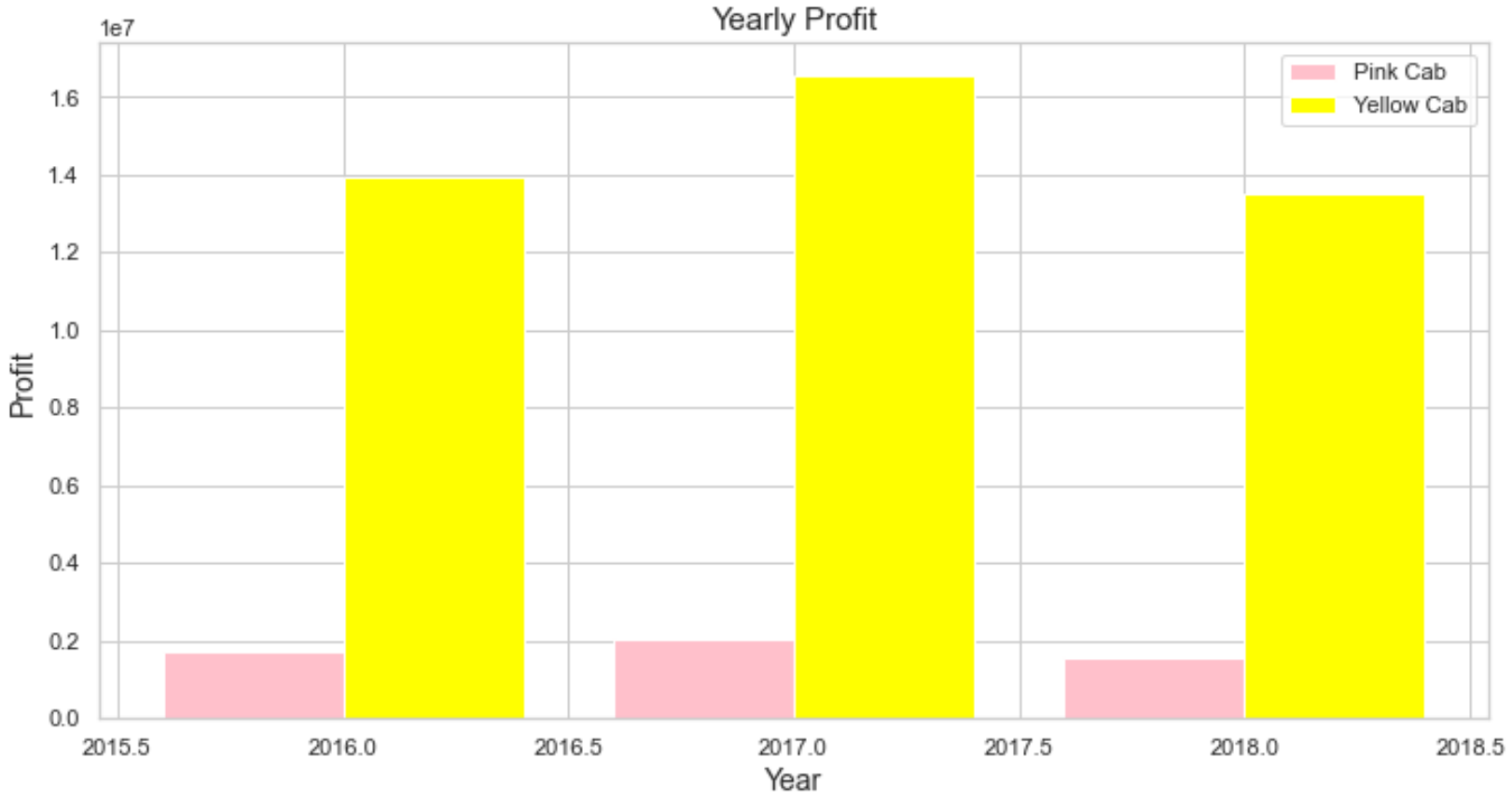
- From 2016-2018 Yellow Cab made considerably more transactions compared to Pink Cab.

Distance Analysis



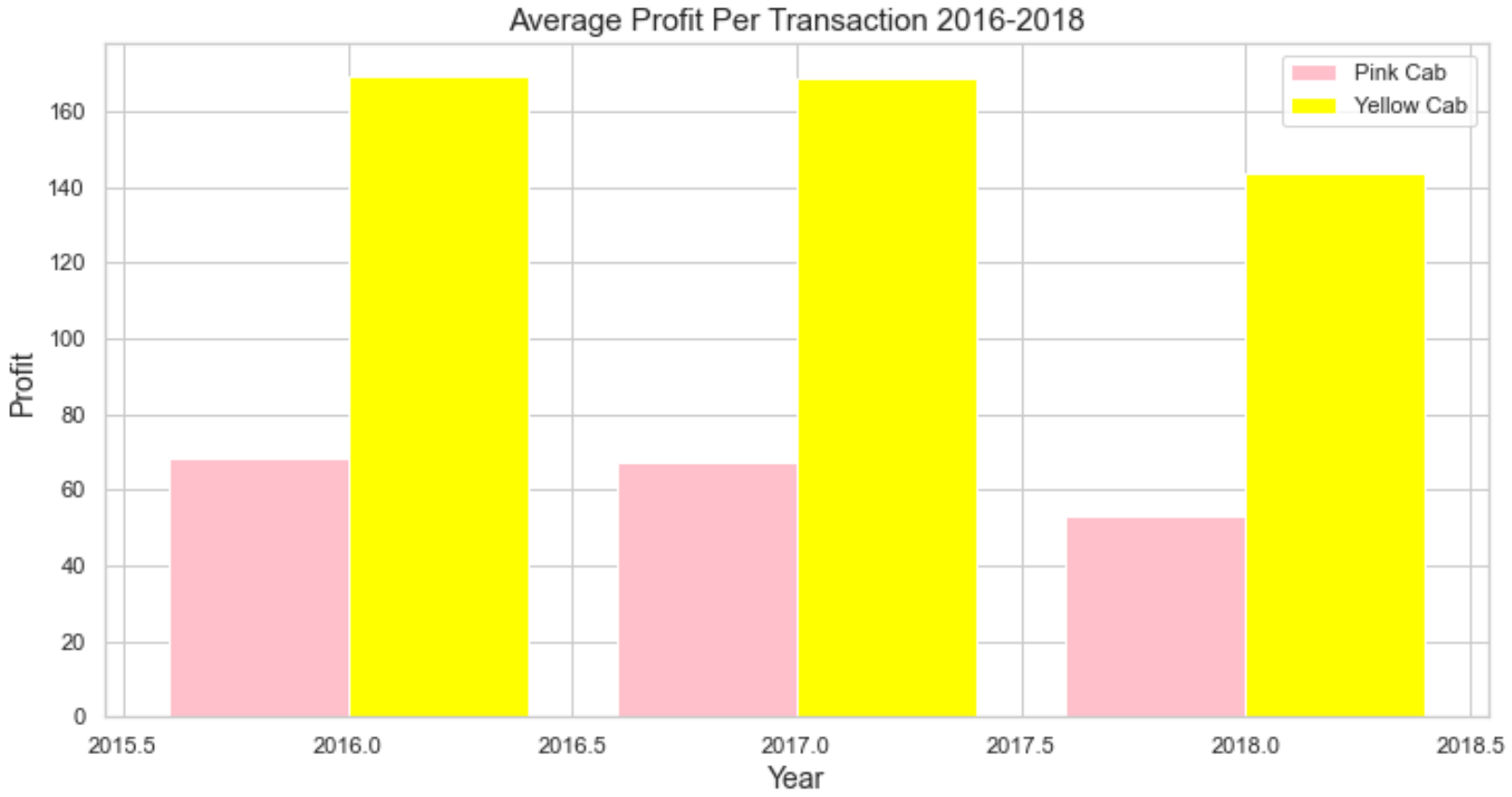
- From 2016-2018 Yellow Cab travelled considerably more compared to Pink Cab.

Profit Analysis



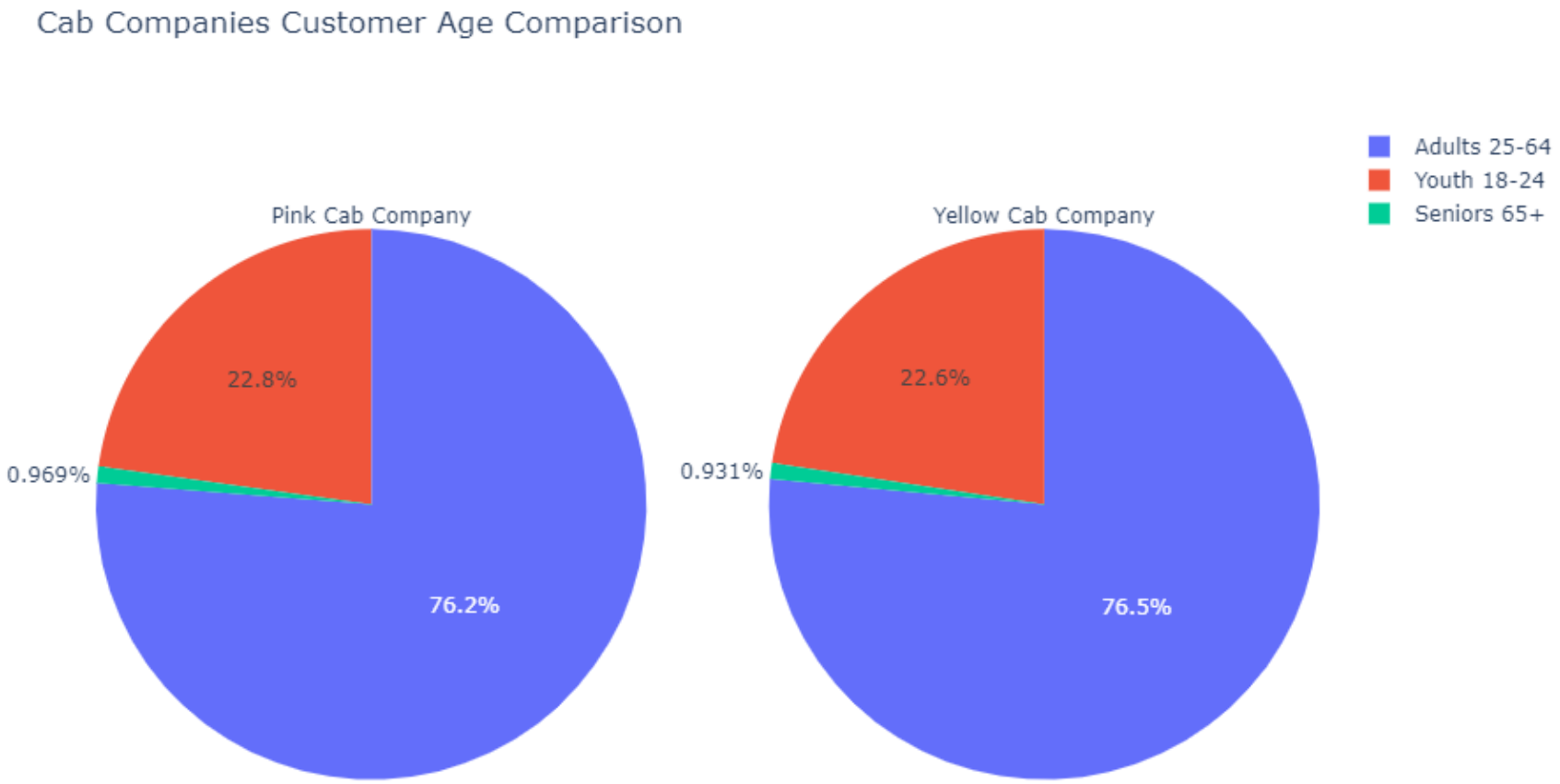
- Total Yearly Profit of Yellow Cab is numerous times more than Pink Cab.

Profit Analysis cont.



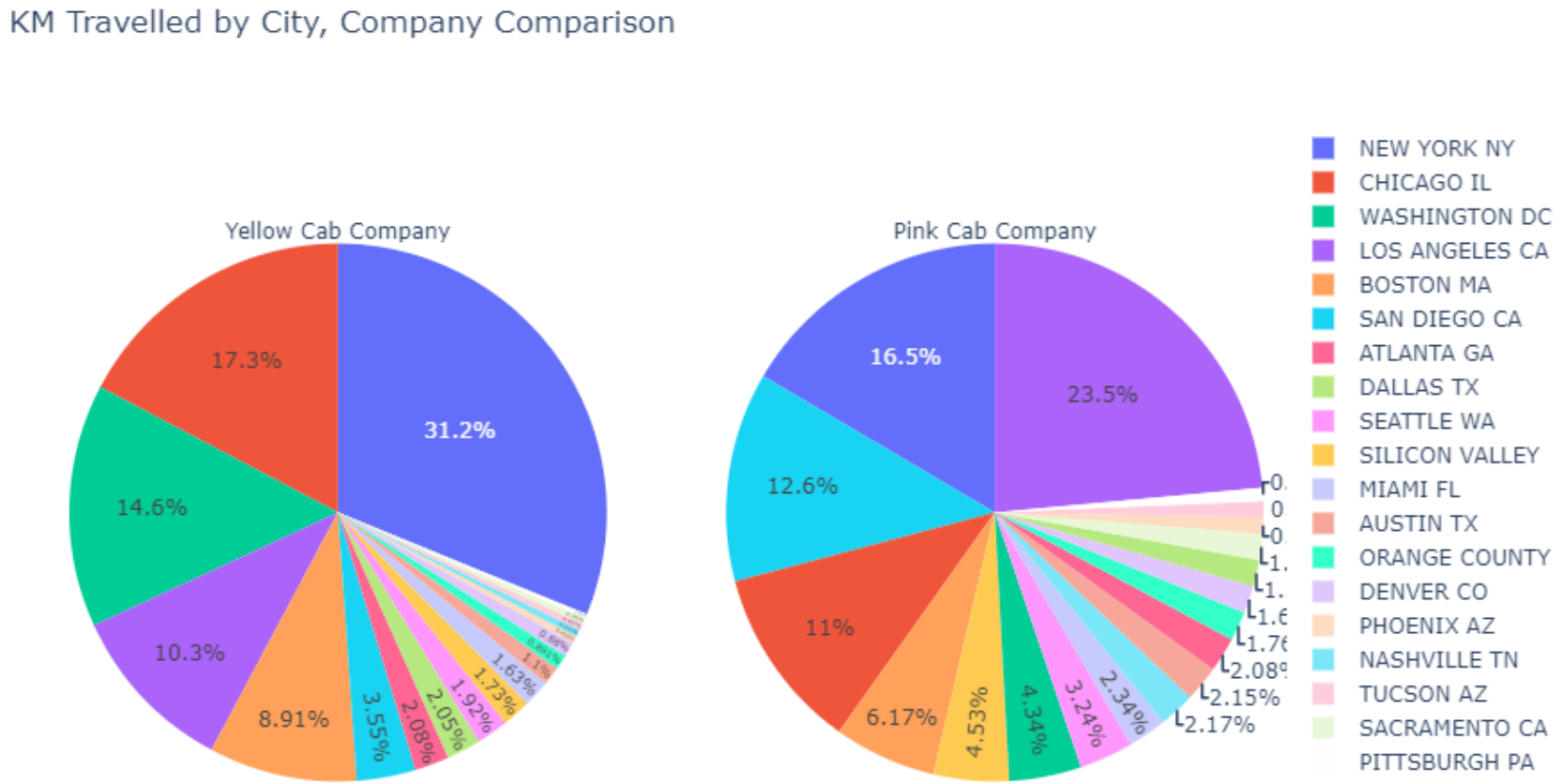
- Yellow Cab made more profit on average per transaction from 2016-2018, compared to Pink Cab.

Age Analysis



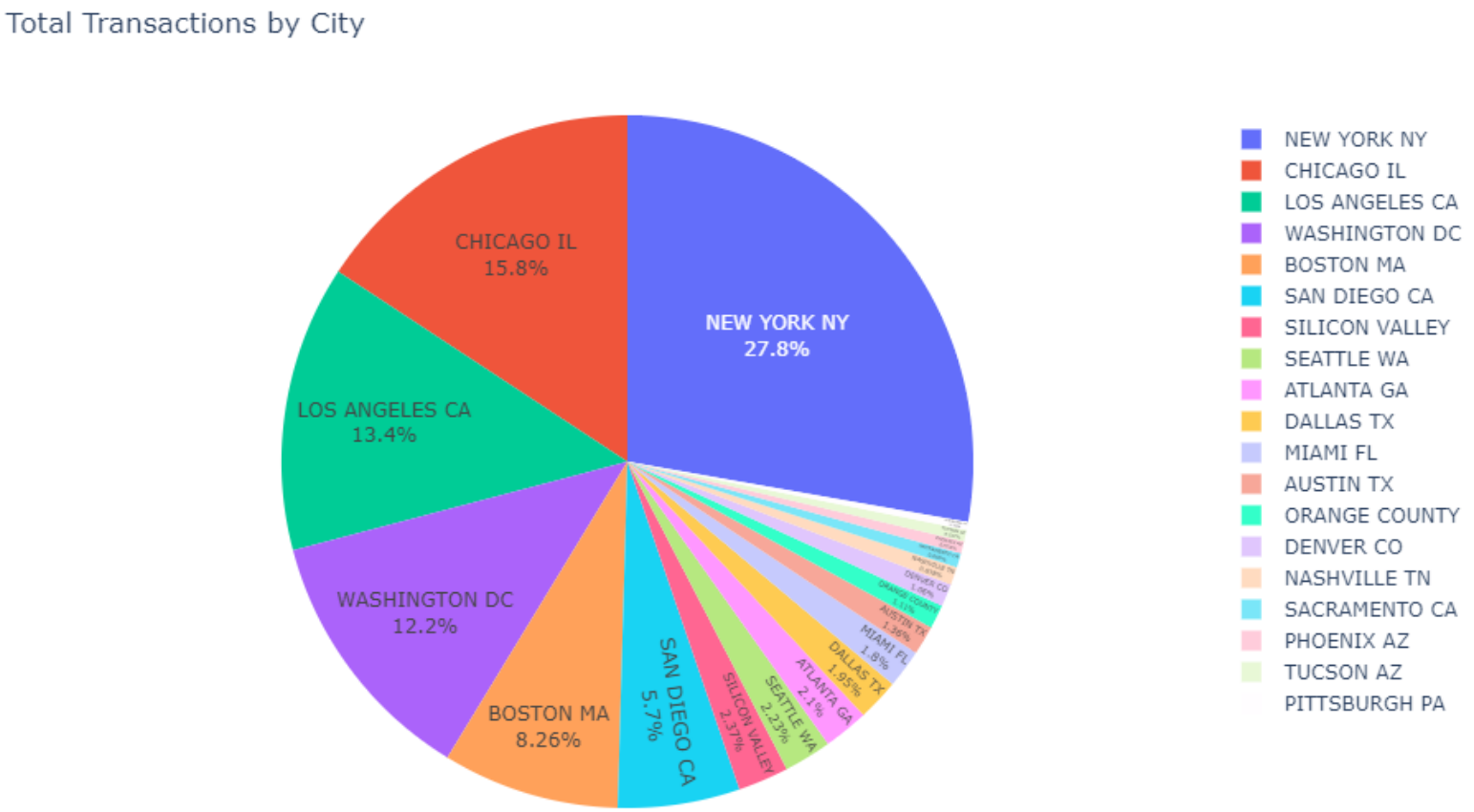
- The biggest market for both companies appears to be Adults, followed by Youth and a small portion of Senior customers.

City Analysis



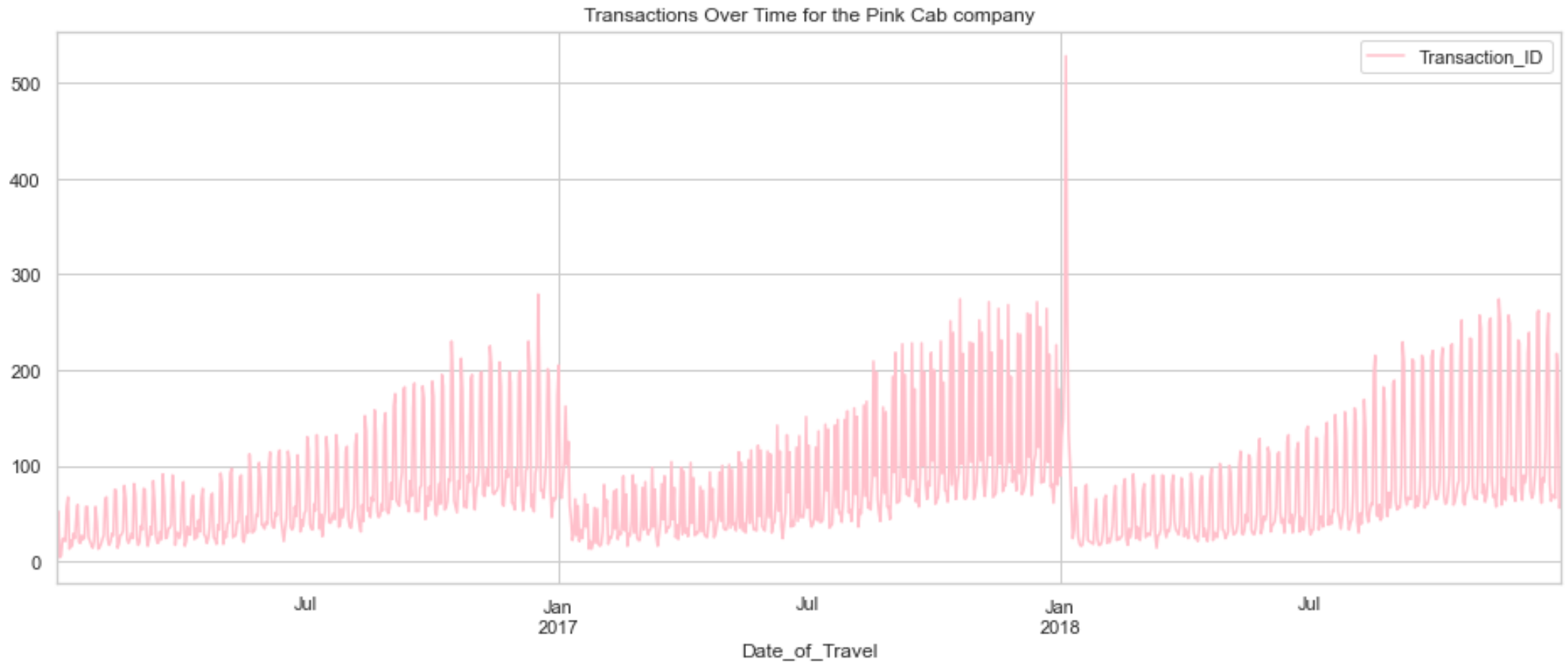
- **Pink Cab Company:** Company is mostly active in Los Angeles, New York, San Diego, Chicago and Boston.
- **Yellow Cab Company:** Company is mostly active in New York, Chicago, Washington, Los Angeles and Boston.

City Analysis cont.



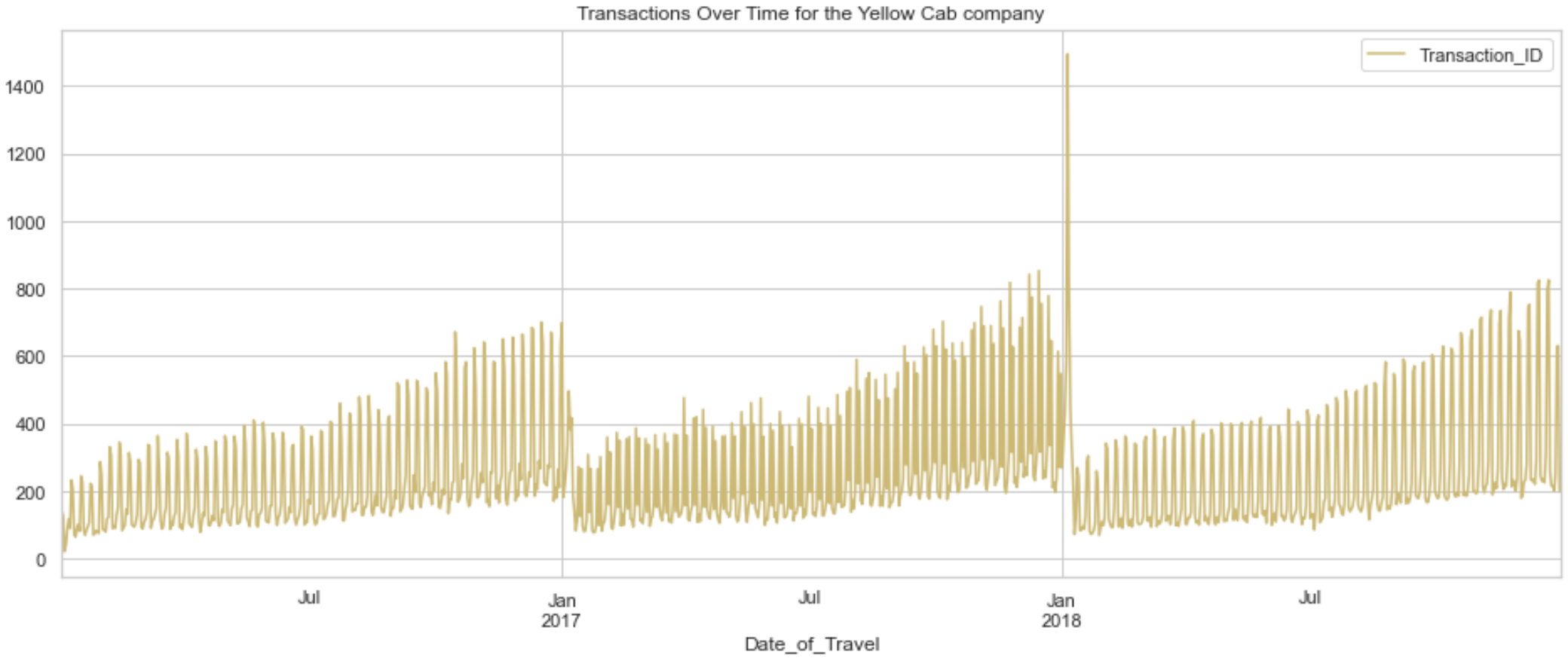
- New York has the most transactions, followed by Chicago, Los Angeles, Washington and Boston etc.

Time Series Analysis : Pink Cab – Transactions over time



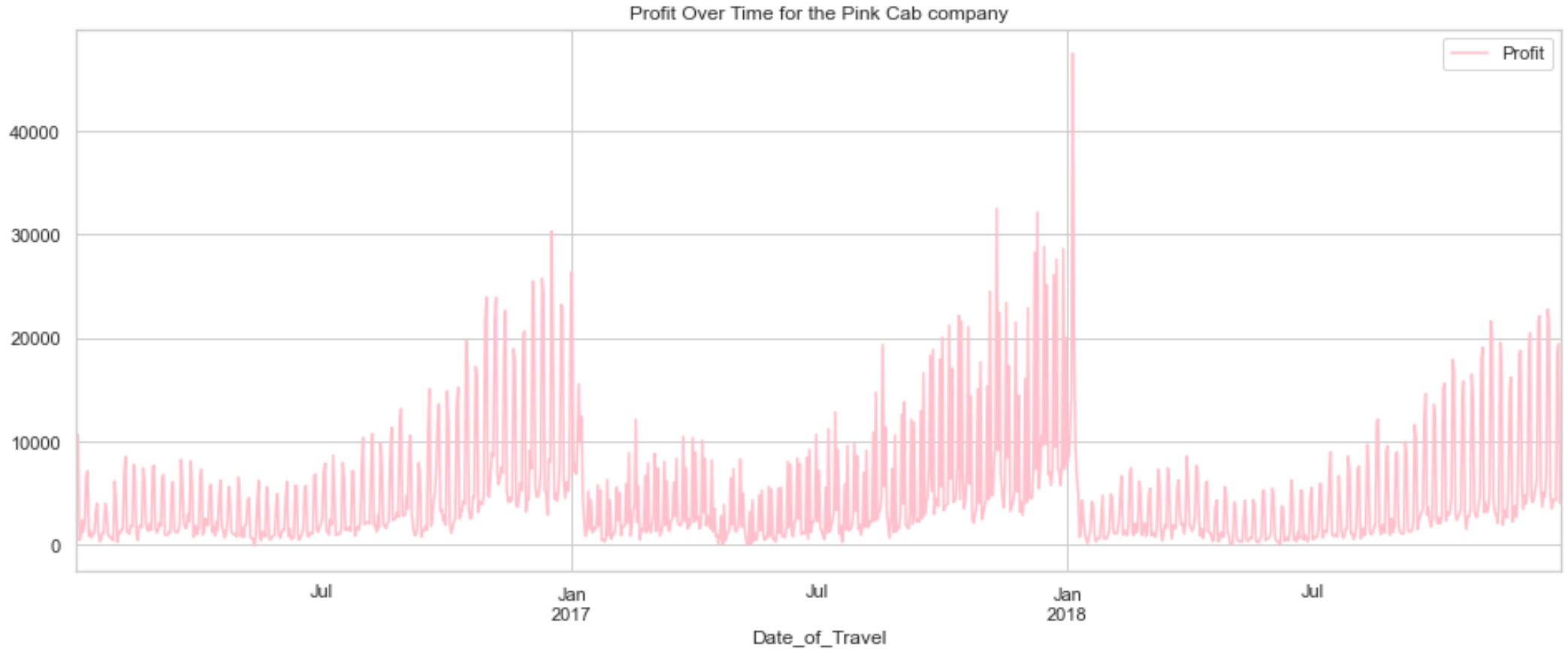
- Clear upward trend in transactions throughout each year.
- October - December appears to be the most busy time

Time Series Analysis : Yellow Cab – Transactions over time



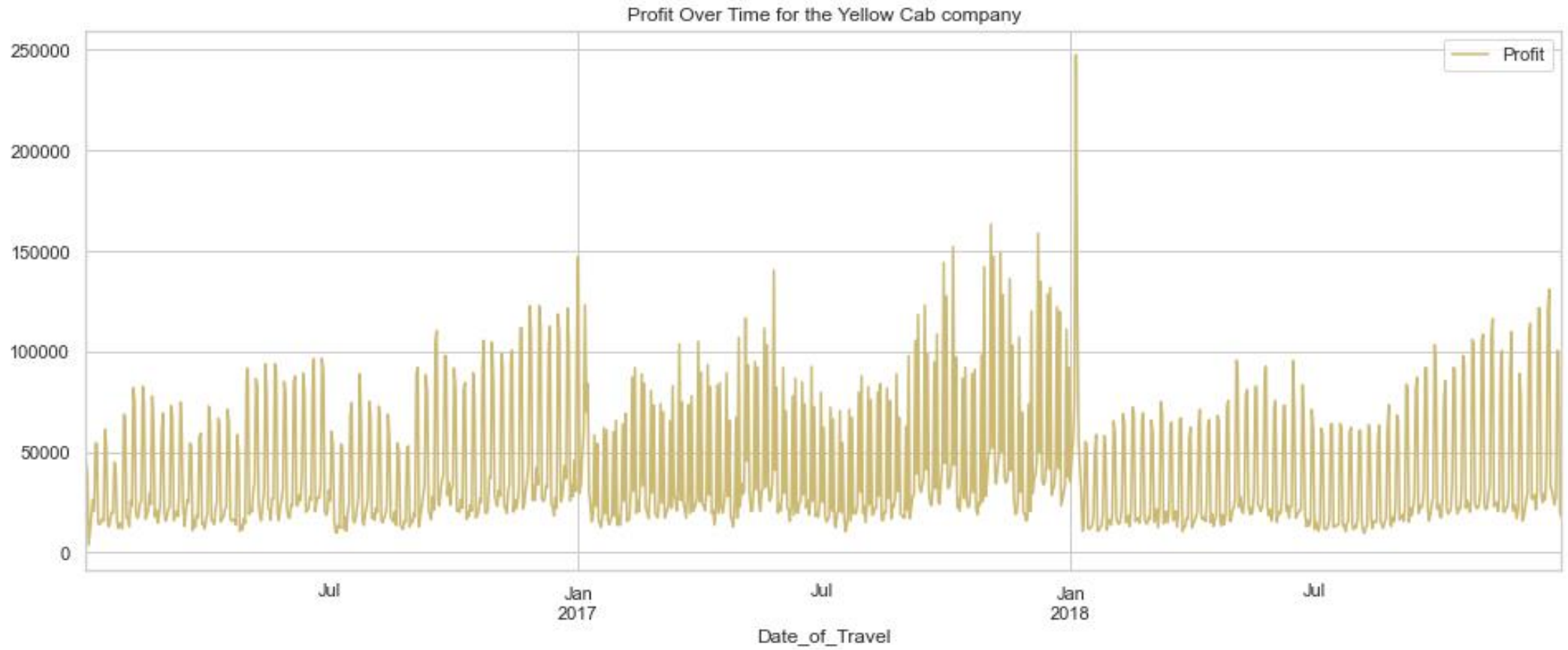
- Upward trend in transactions throughout each year.

Time Series Analysis : Pink Cab – Profit over time



- Upward trend in profit throughout each year.
- October - December appears to be the most profitable time.

Time Series Analysis : Yellow Cab – Profit over time



- Some stability and an increase in October – December.

EDA Summary

- **Transaction History:** Yellow Cab has more than **double the transactions every year** from 2016 – 2018 compared to Pink Cab.
- **Profitability:** Yellow Cab made **more than double the profit per transaction every year** from 2016-2018 compared to Pink Cab.
- **Age Reach**: Yellow Cab is popular with Young Adults (*Age < 25)*, although its **biggest market lies with Adults *(Age 25-64)*** and it has some reach with the Seniors (Age > 65).
- **Stability & Demand**: Yellow Cab's income is much more stable and **in demand throughout the year**, compared to Pink Cab, which is heavily spiked at the end of the year.
- **City Dominance:** Yellow Cab is **dominating the market** in some of the **biggest cities**, including New York, Chicago and Washington DC.

Recommendation

Considering both company's transaction history, profitability, age reach, city dominance, stability and demand, I highly recommend **investing in the Yellow Cab Company.**

Your Deep Learning Partner

Thank You