Exploratory Data Analysis

G2M Insight for Cab Investment Firm

Represented by: Abdukhakimov Asatilla

2025.02.19

Agenda

PROBLEM STATEMENT

APPROACH

DATA EXPLORATION

EDA

EDA SUMMARY

RECOMMENDATIONS

Data Glacier

Your Deep Learning Partner

Problem Statement

- XYZ, a private firm in the US, is evaluating investment opportunities in the rapidly growing cab industry.

- To make an informed decision, XYZ needs a comprehensive market analysis of two key competitors.

- The challenge is to identify which company— Yellow Cab or Pink Cab — offers the best growth potential and profitability for investment.

- Key factors: revenue, profitability, customer demographics, market expansion potential, and financial projections.

- XYZ seeks data-driven insights to refine its Go-to-Market (G2M) strategy and maximize returns.
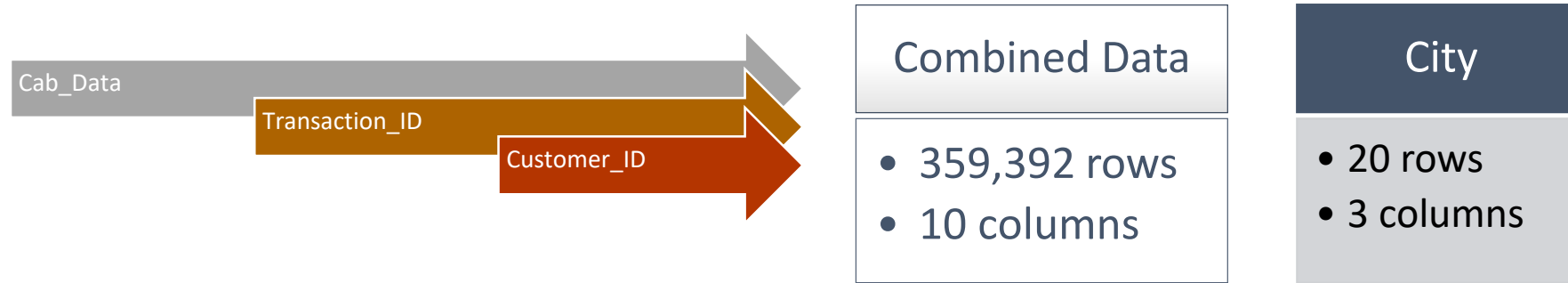
Approach

Data Exploration: Preparing, understanding and formatting data.

Exploratory Data Analysis (EDA): Revenue, trip distance, customer demographics, payment trends, seasonality, and geographic performance.

Comparative Analysis: Evaluate financial performance and user behavior differences between the two companies.

Predictive Forecasting: 2019 profit projections for both companies.

Data Exploration

Cab_Data		
	Transaction_ID	
		Customer_ID

Combined Data
- 359,392 rows
- 10 columns

City
- 20 rows
- 3 columns

- Merged Key Tables – Cab, Customer, and Transaction data for efficient analysis.

- Data Cleaning – Removed irrelevant columns (Transaction ID & Customer ID).

- Checked Missing Values & Outliers

- Date Formatting – Converted "Date of Travel" and extracted the year, month, and weekday columns.

Compare Revenue by Company
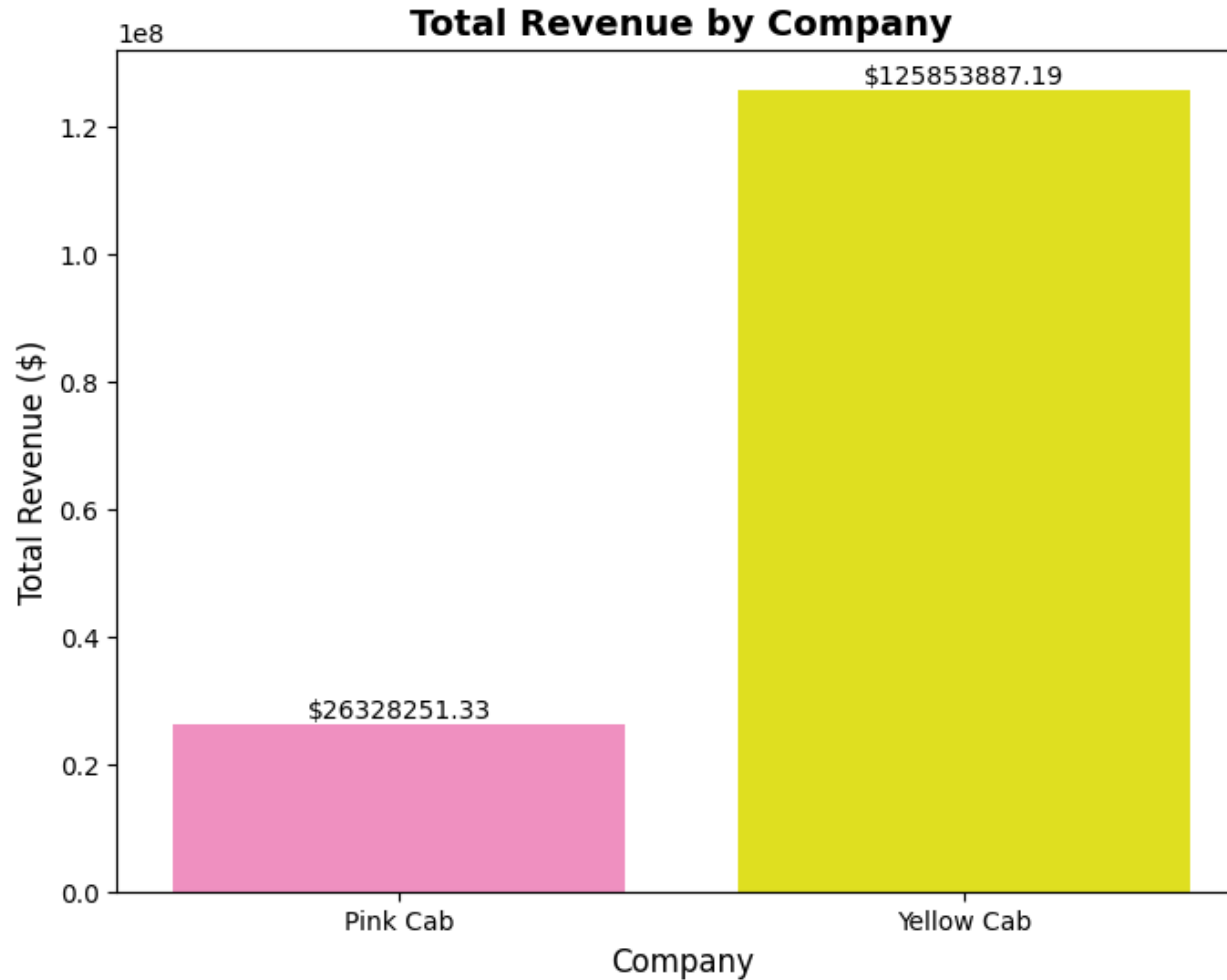
Yellow Cab's revenue is approximately five times greater than Pink Cab's

Compare Average Trip Distance

Average Trip Distance by Company

22.56 km   22.57 km

Average Trip Distance (km)

20

15

10

5

0

Pink Cab   Yellow Cab
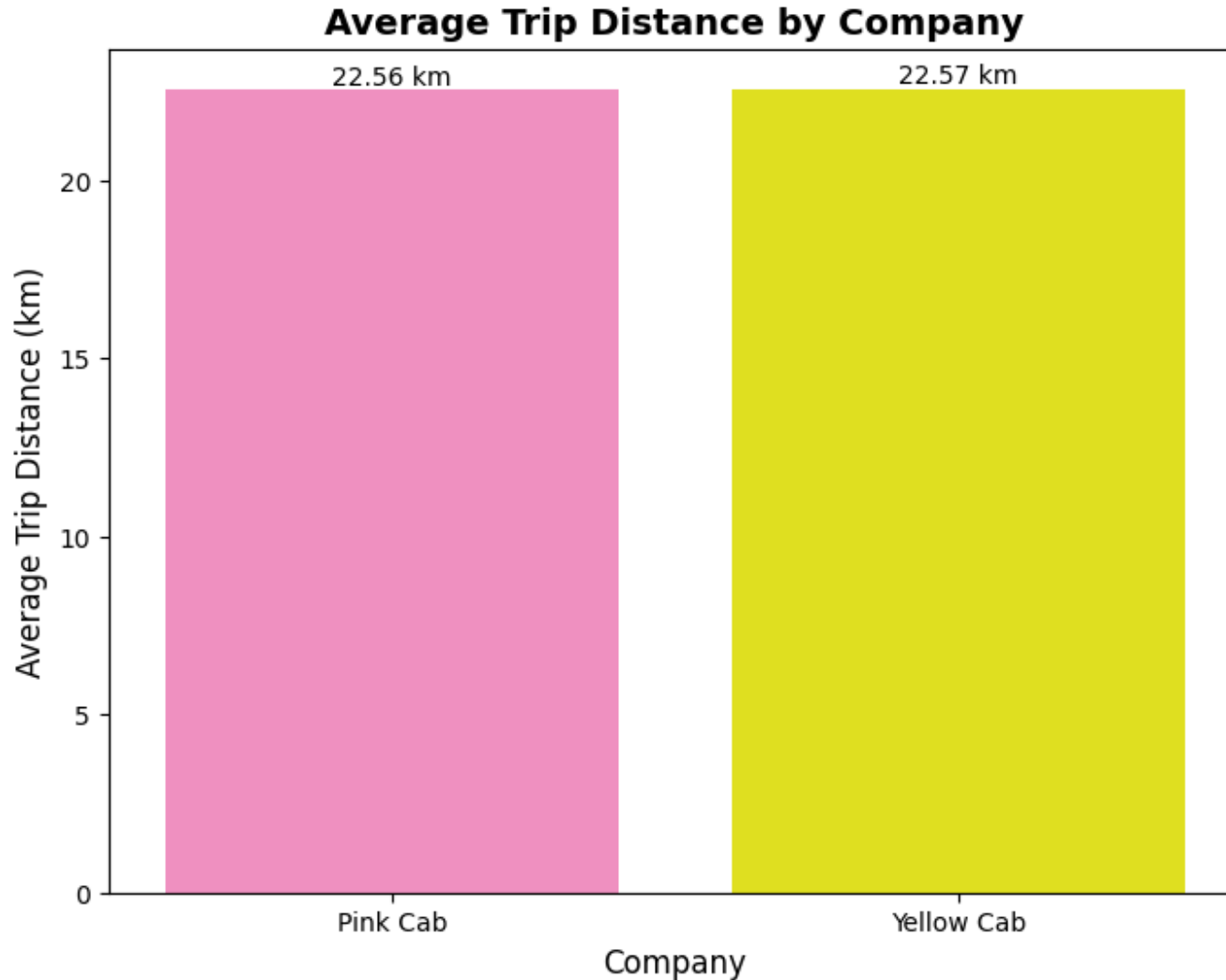
Company

The average trip distance for both companies is nearly identical.

Relationship between Fare and Distance

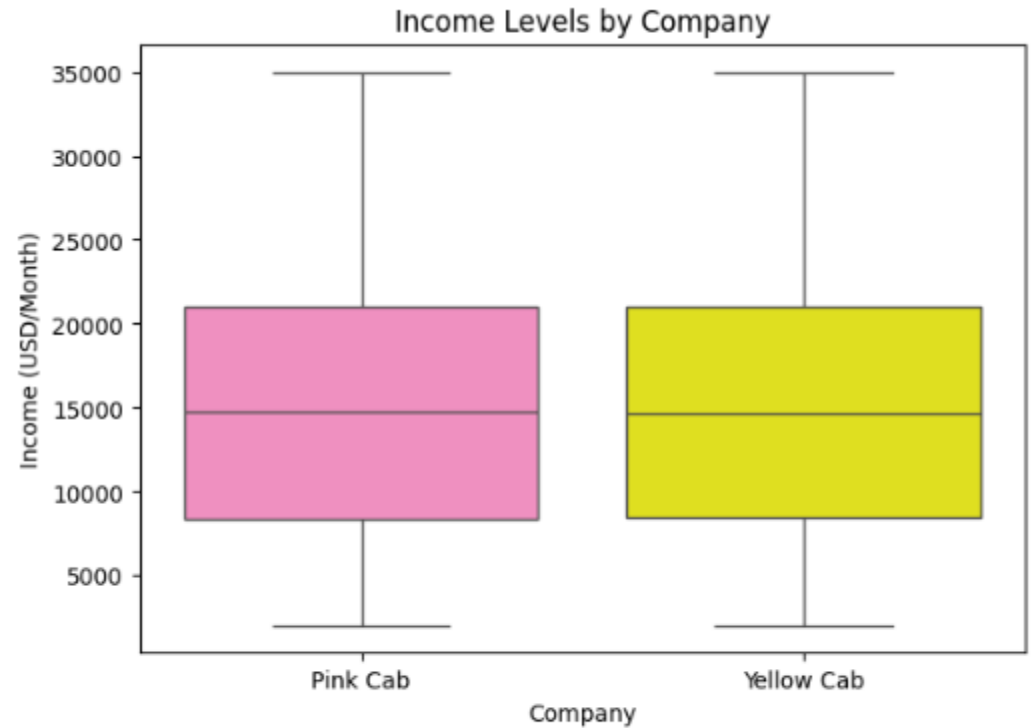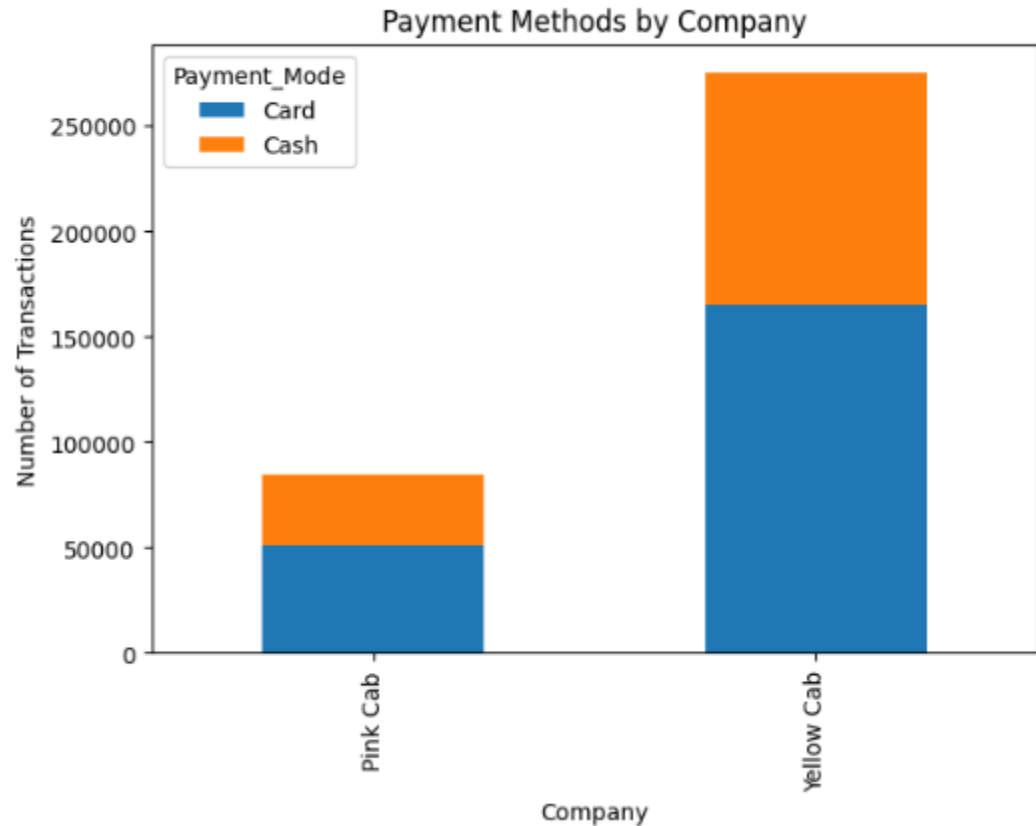Price Charged vs. KM Travelled

A positive correlation between kilometers traveled and price charged for both

Analysis of Customer Demographics
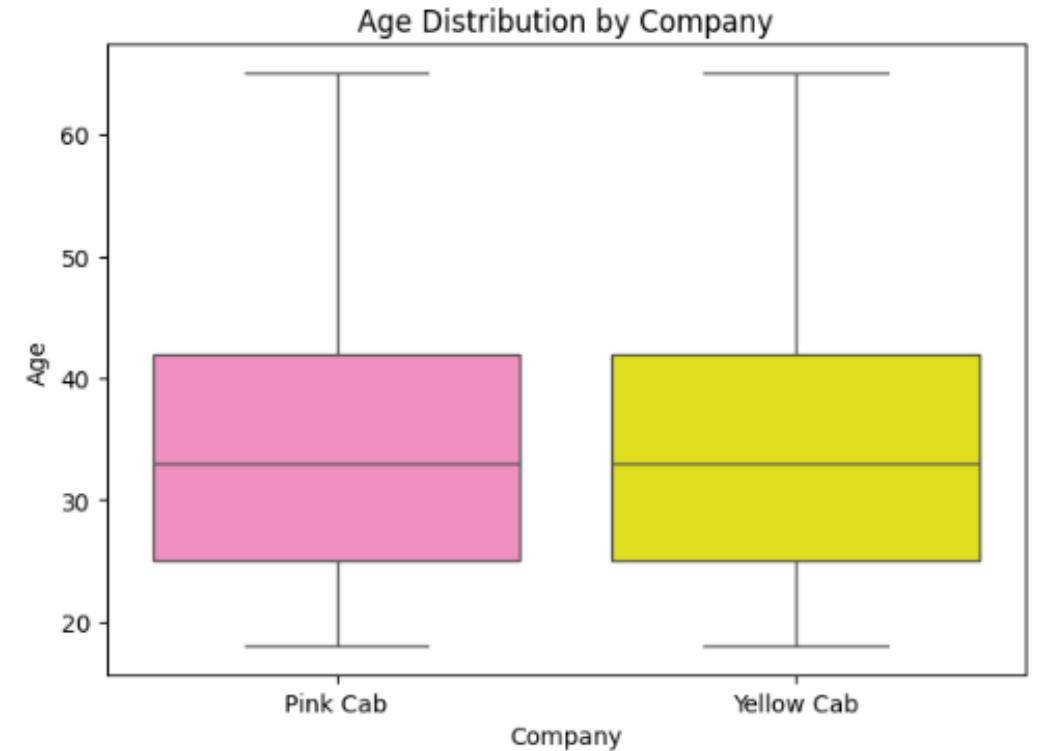
The customer gender & age distribution is similar for both companies, with a majority of male customers.
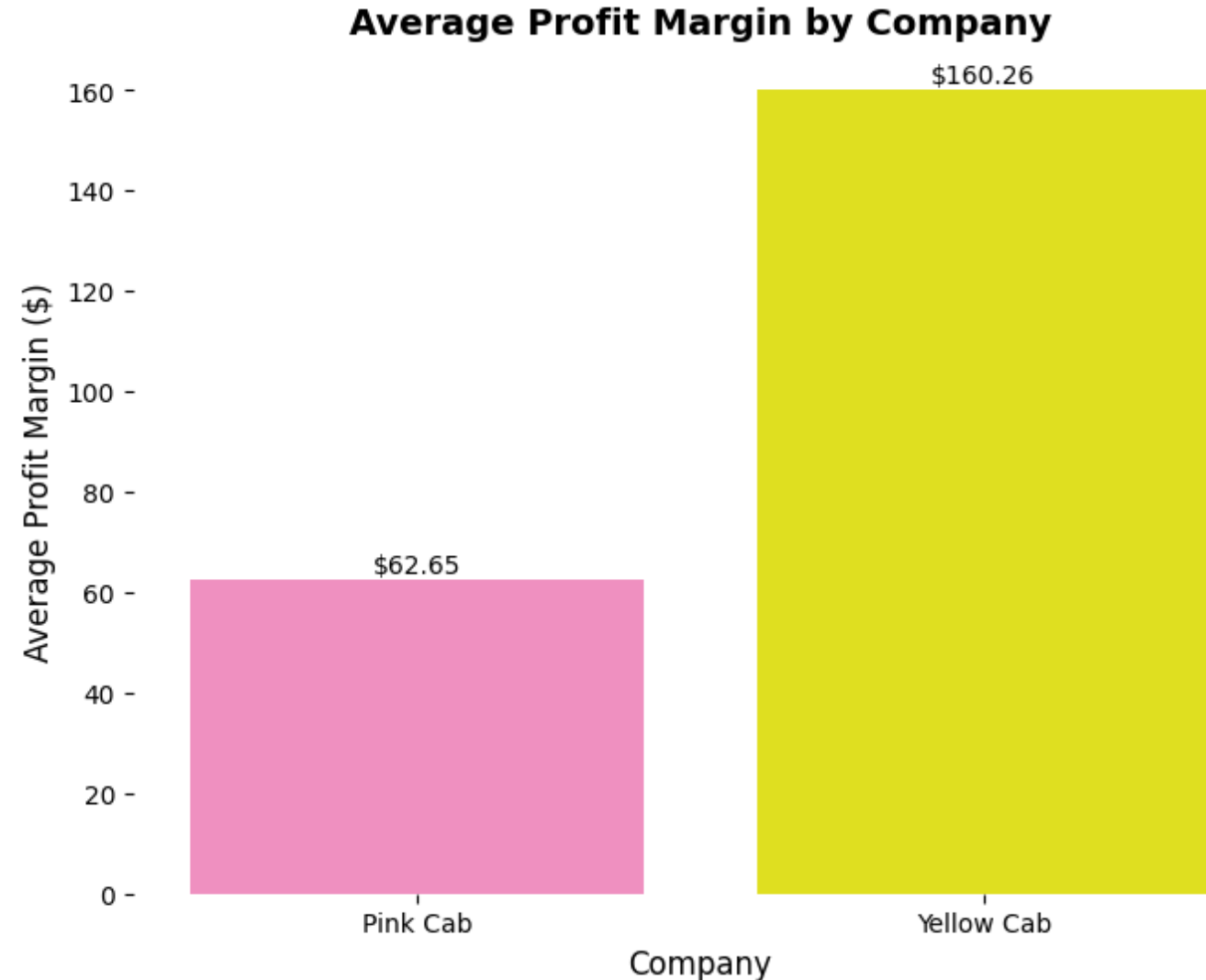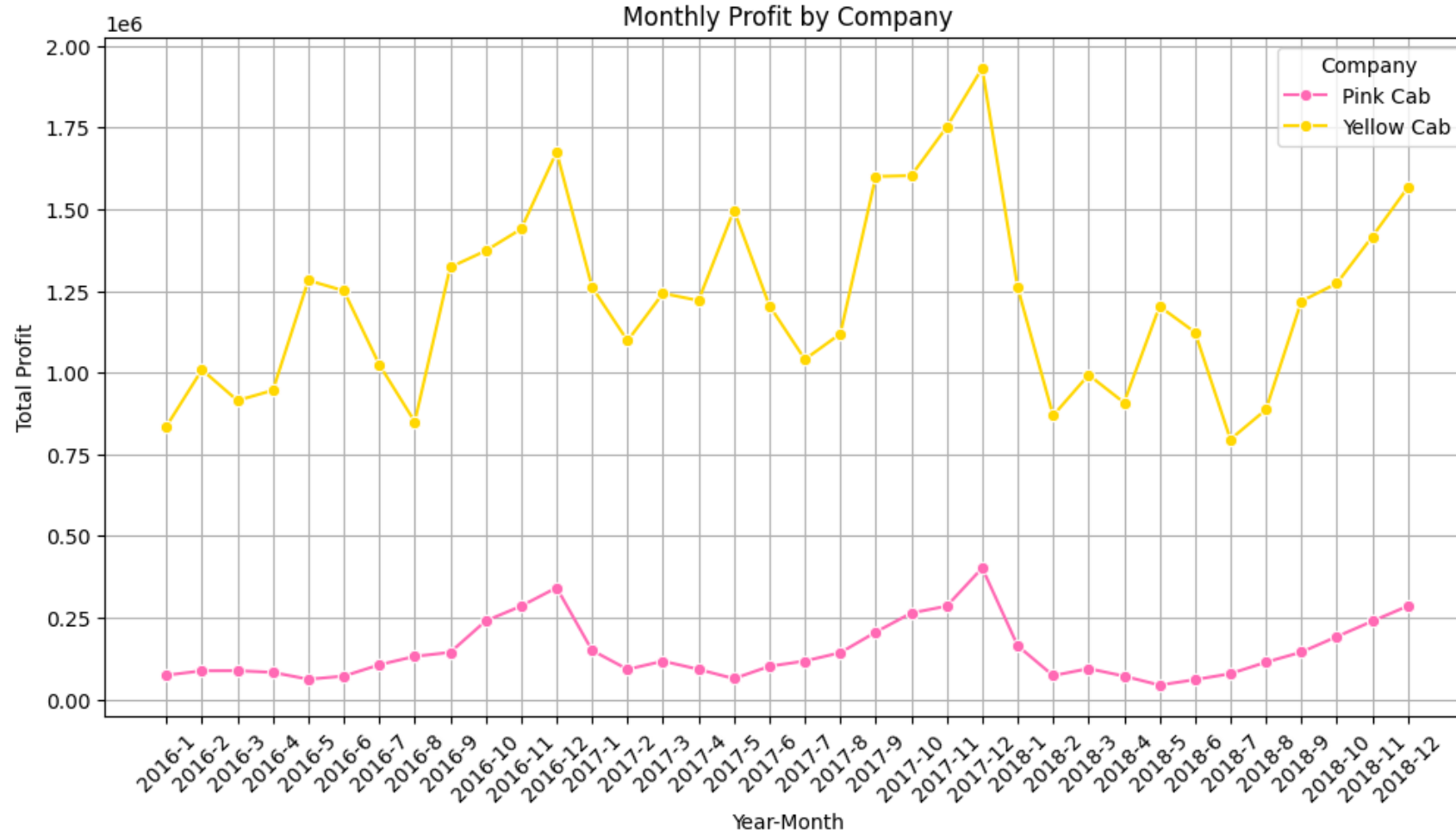
Analysis of Payment and Income Metrics

Both companies have similar payment distributions, with card payments 1.5x more frequent, and comparable income levels.

Compare Profit Margins

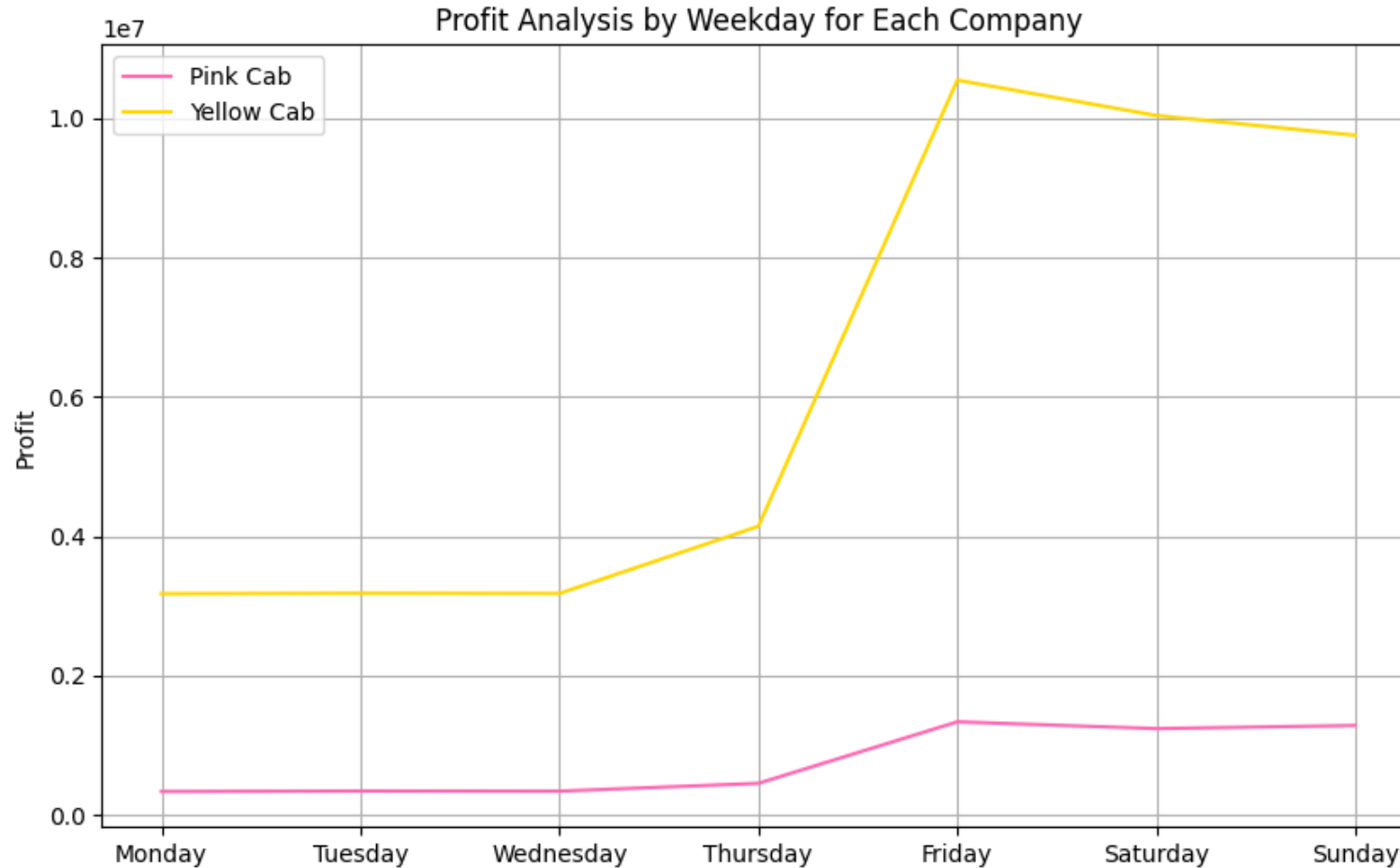Yellow Cab's average profit margins are 2.5 times higher than those of its competitor

Average Profit Margin by Company

Monthly Profit by Company

Yellow cab consistently outperforms and both companies exhibit seasonal spikes in performance during December.
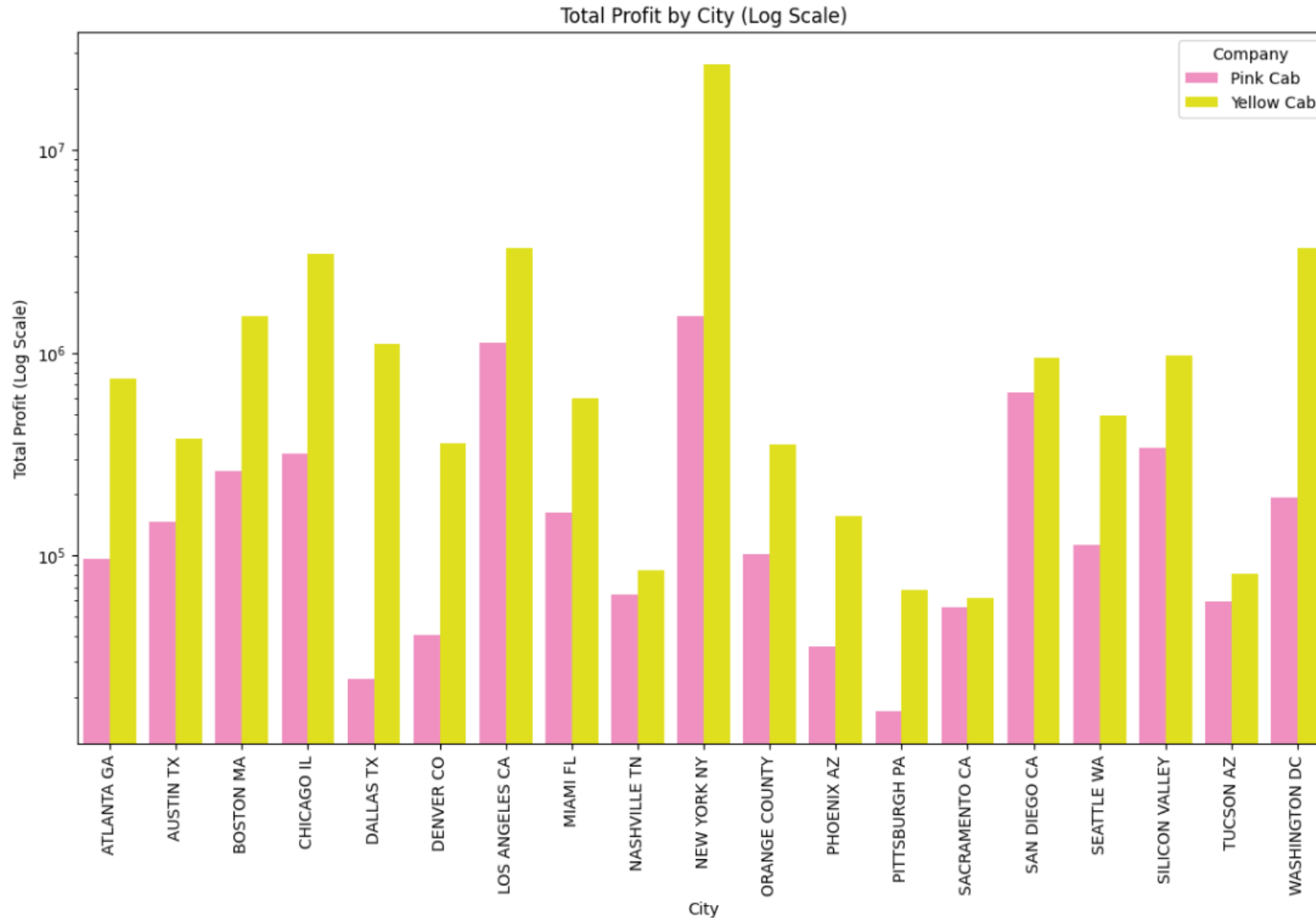
Profit Analysis by Weekday

Profit Analysis by Weekday for Each Company

Friday was the most profitable day for both companies, followed closely by the weekend.

Yearly Profit Analysis

Yearly Profit Analysis by Company
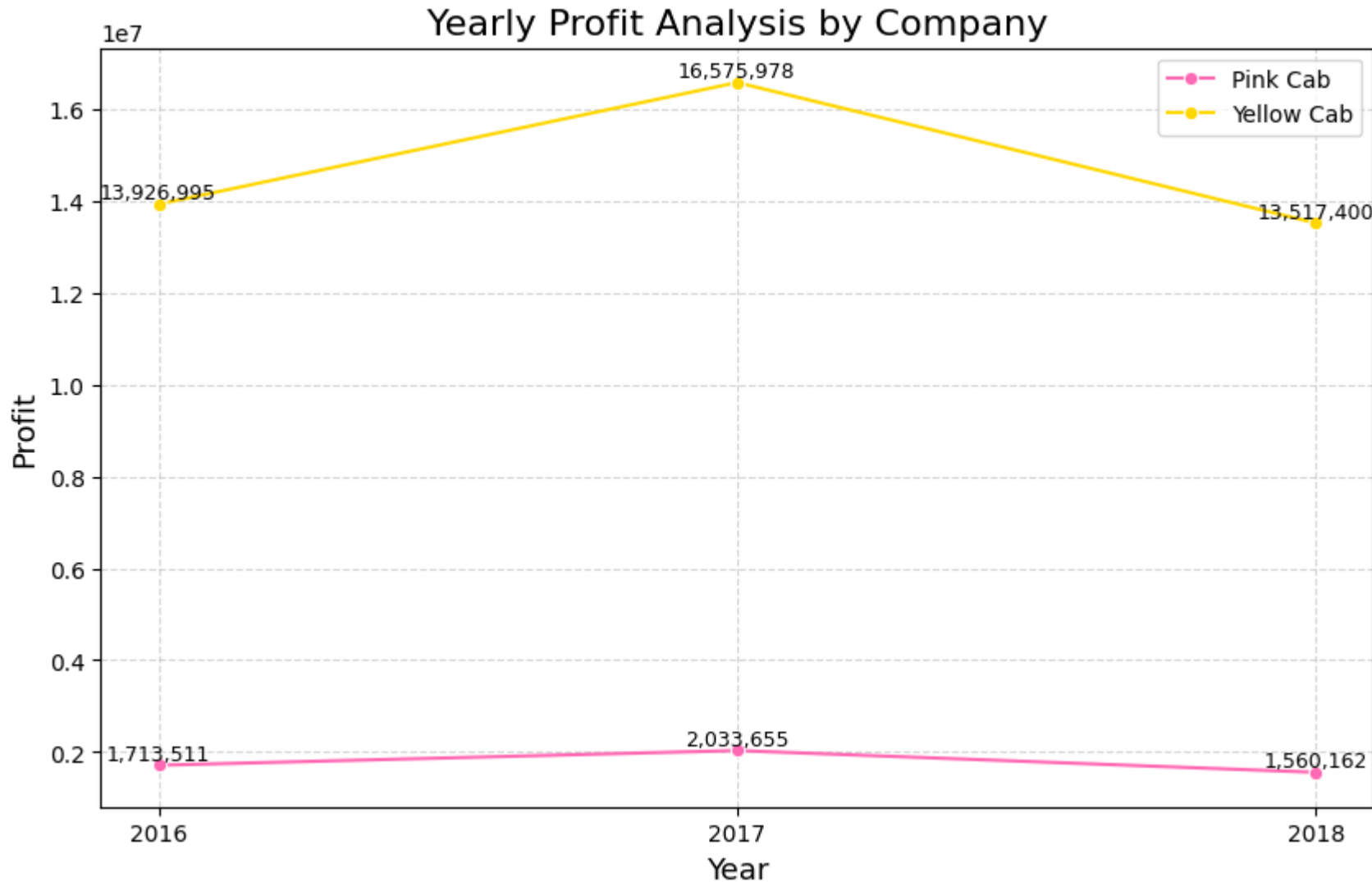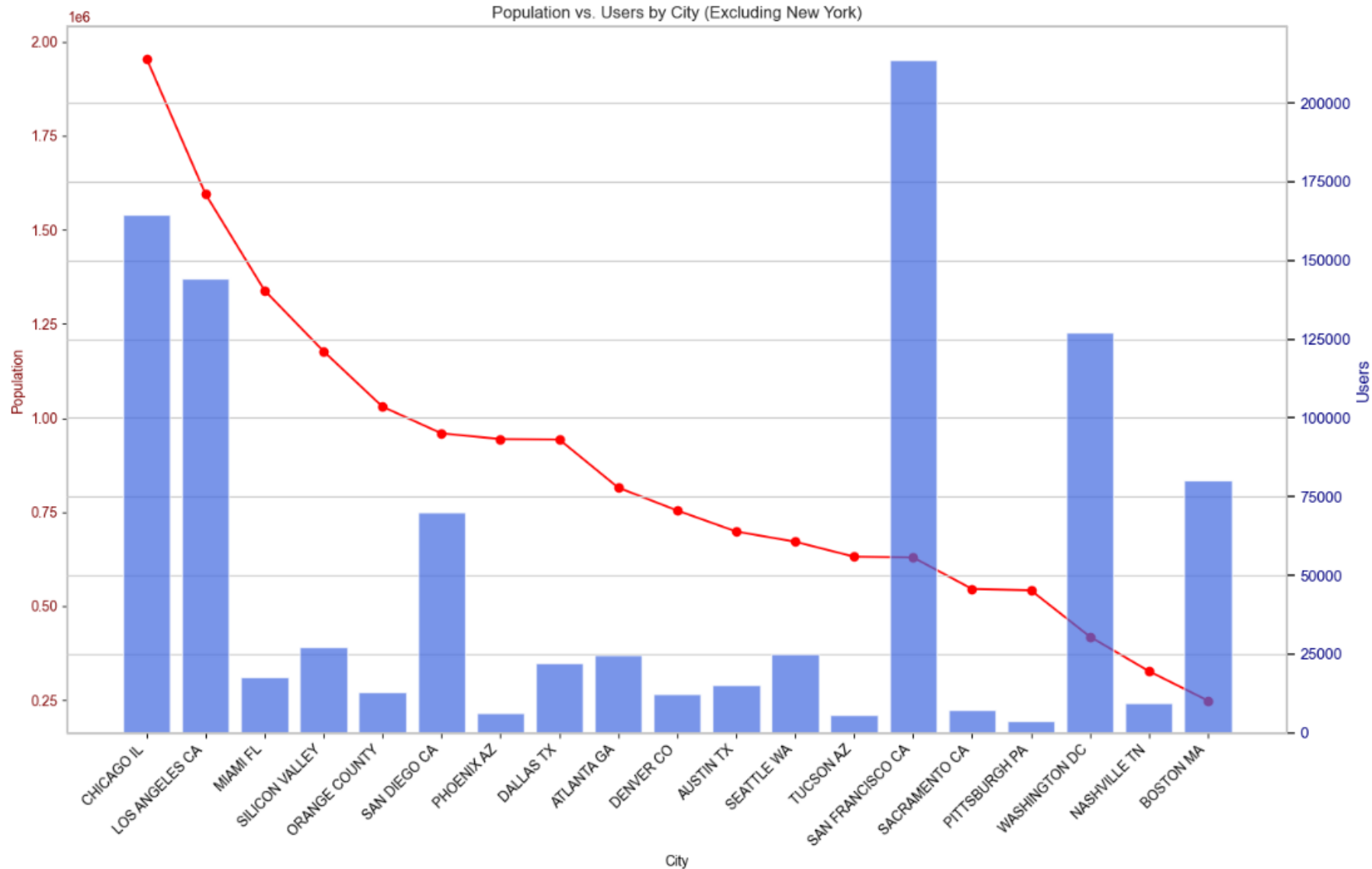
Revenue for both companies peaked in 2017 and this was followed by a decline in 2018
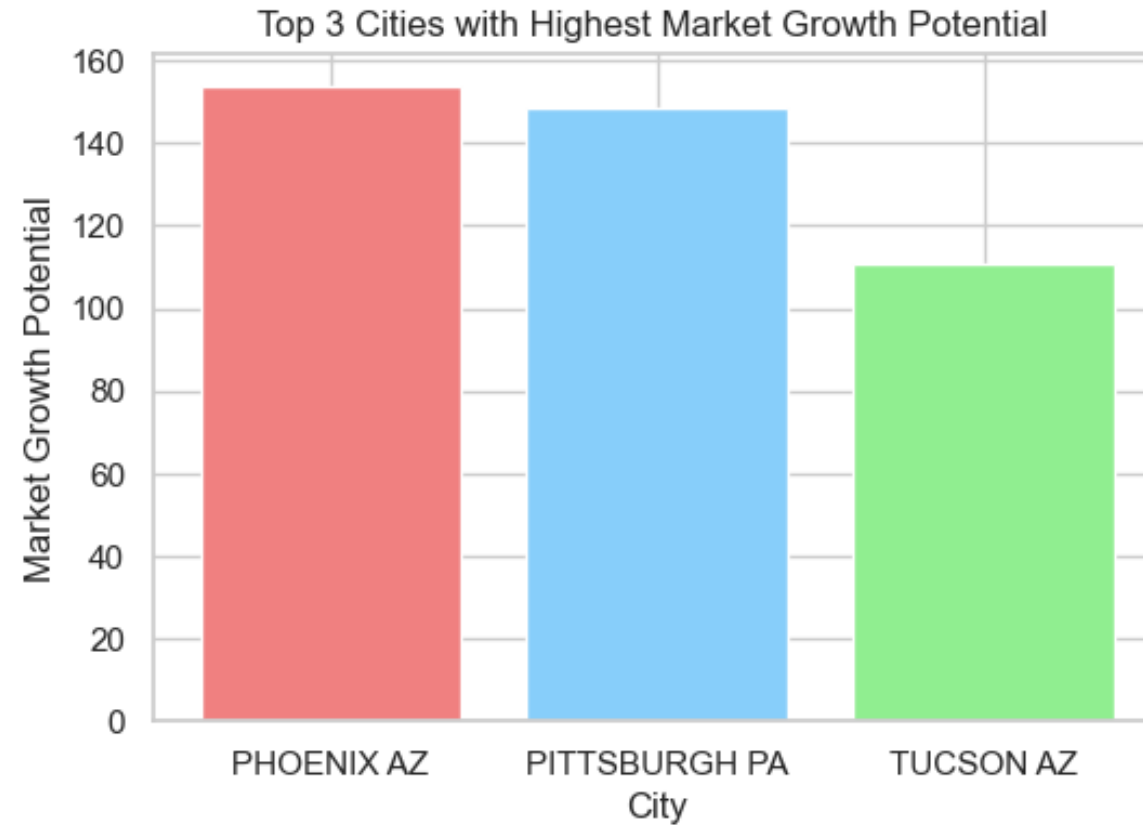
Total Profit by City

Total Profit by City (Log Scale)

New York City demonstrates significantly higher profitability for both companies

City and Population Analysis

Population vs. Users by City (Excluding New York)
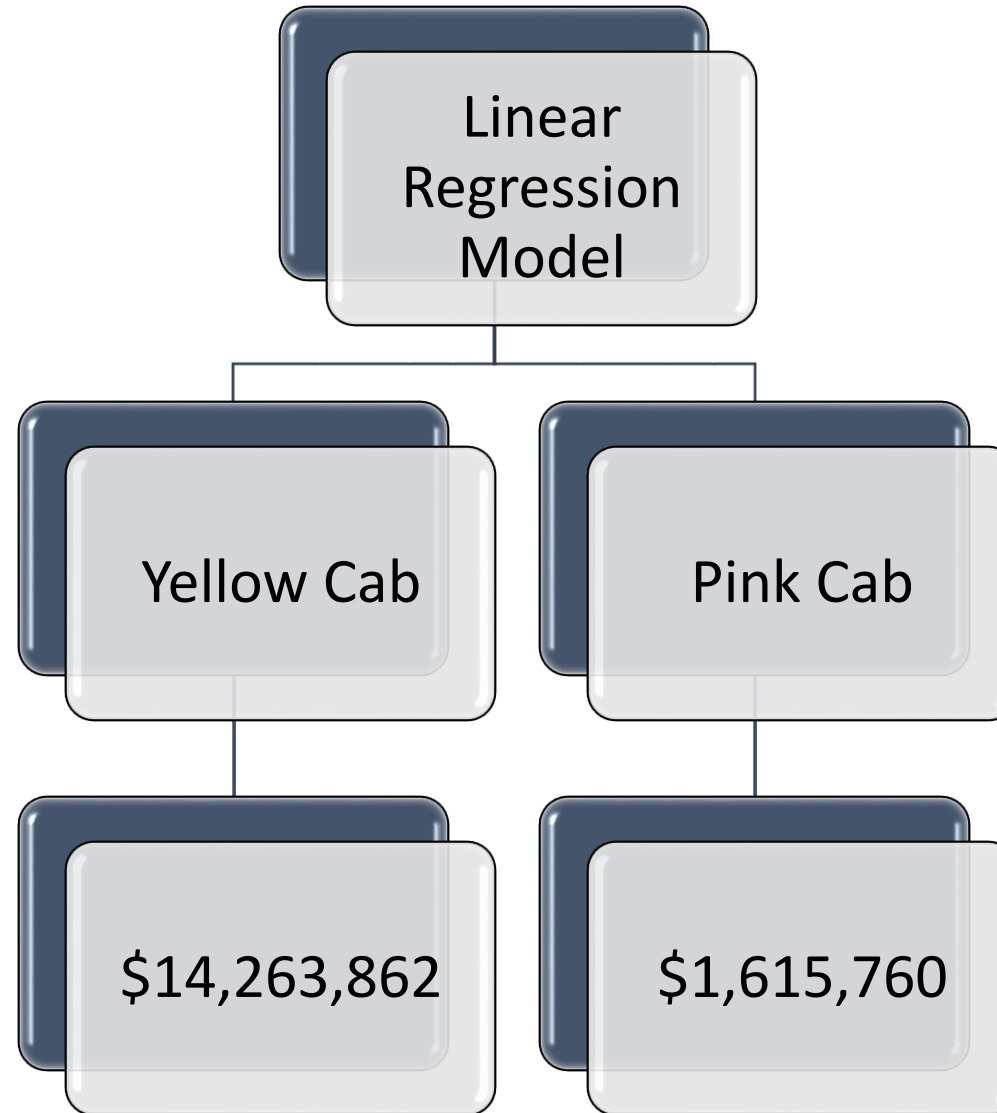
Market Growth Potential

Top 3 Cities with Highest Market Growth Potential

PHOENIX AZ, PITTSBURGH PA, and **TUCSON AZ** have large populations
but few users, highlighting market expansion potential

Future Forecasting for 2019 Profit

EDA Summary

- **Revenue & Profitability:** Yellow Cab significantly outperforms Pink Cab in both revenue and profit.

- **Customer Behavior:** Similar demographics, trip distances, and payment preferences.

- **Seasonality:** Higher profits during colder months and peak demand in Dec.

- **Geographical Insights:** NYC is the most profitable market, with expansion opportunities in Phoenix, Pittsburgh, and Tucson.

- **Forecast:** Yellow Cab projected to maintain higher profitability in 2019, while Pink Cab's underperformance needs further analysis.

Recommendations

- **Invest in Yellow Cab** – Strong market position with 5x higher revenue & 2.5x greater profits.

- **Optimize Pink Cab** – Needs cost efficiency & market expansion before consideration.

- **Expand to High-Potential Cities** – Target Phoenix, Pittsburgh, Tucson for growth.

- **Leverage Seasonal Demand** – Maximize fall, winter, and December peak profits.

- **Enhance Digital Payments & Loyalty** – Boost retention with seamless payments & rewards.

- **Monitor Financial Trends** – Address declining profits with data-driven strategies.

Final Verdict: Yellow Cab is the best investment choice; Pink Cab needs improvement.

THANK YOU

Data Glacier

Your Deep Learning Partner