



**Data Glacier**

Your Deep Learning Partner

# Exploratory Data Analysis

G2M Analysis for Investment in cab companies

**Sept 14, 2023**

# Agenda

Background

Approach

EDA

Recommendations

# Background

- **XYZ, a private equity firm, is looking to venture into the fast-growing cab industry**
- **Data is analyzed for 2 potential investments- Pink Cab & Yellow Cab**
- **Different customer aspects analyzed to gain insights on best investment choice**

Data Sources:

1. Cab Data
2. Customer ID
3. Transaction ID
4. City

Data Period:

Jan 31, 2016 to Dec 31, 2018

Data Exploration:

Each dataset was examined for the data structure, presence of missing (NaN) values and duplicates.

# Approach

Creating an applicable dataset:

- Imported necessary libraries
- Merged relevant datasets into a master dataset
- Removed missing values
- Converted TravelDate to a DateTime format
- Removed outliers

Data Preparation & Analysis:

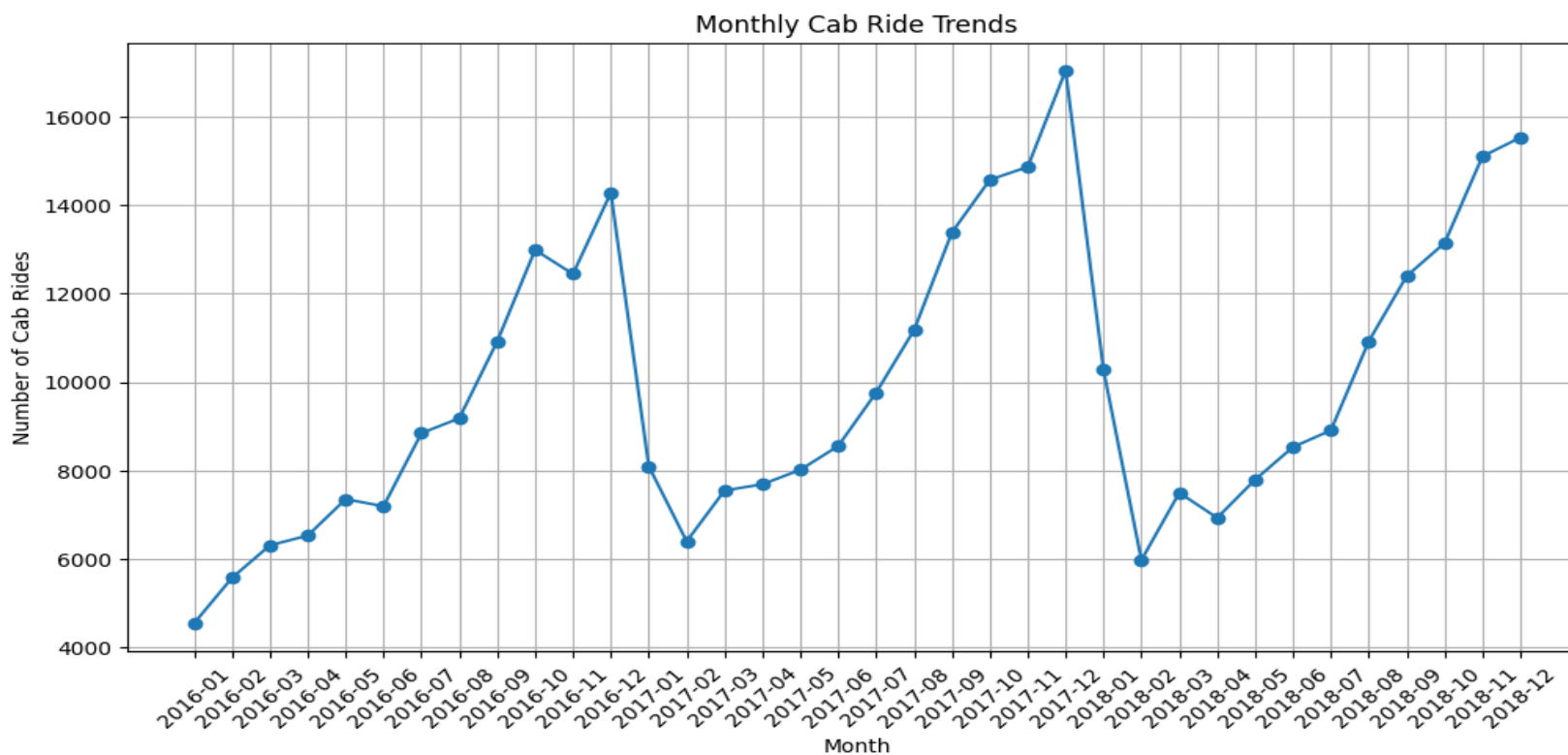
- Data Cleaning
- Exploratory Data Analysis
- Descriptive Data Visualization
- Comparative Analysis

**Reason for this approach:**

1. To gain meaningful insights from datasets
2. For investigating specific relevant questions
3. For comparison of metrics against various factors



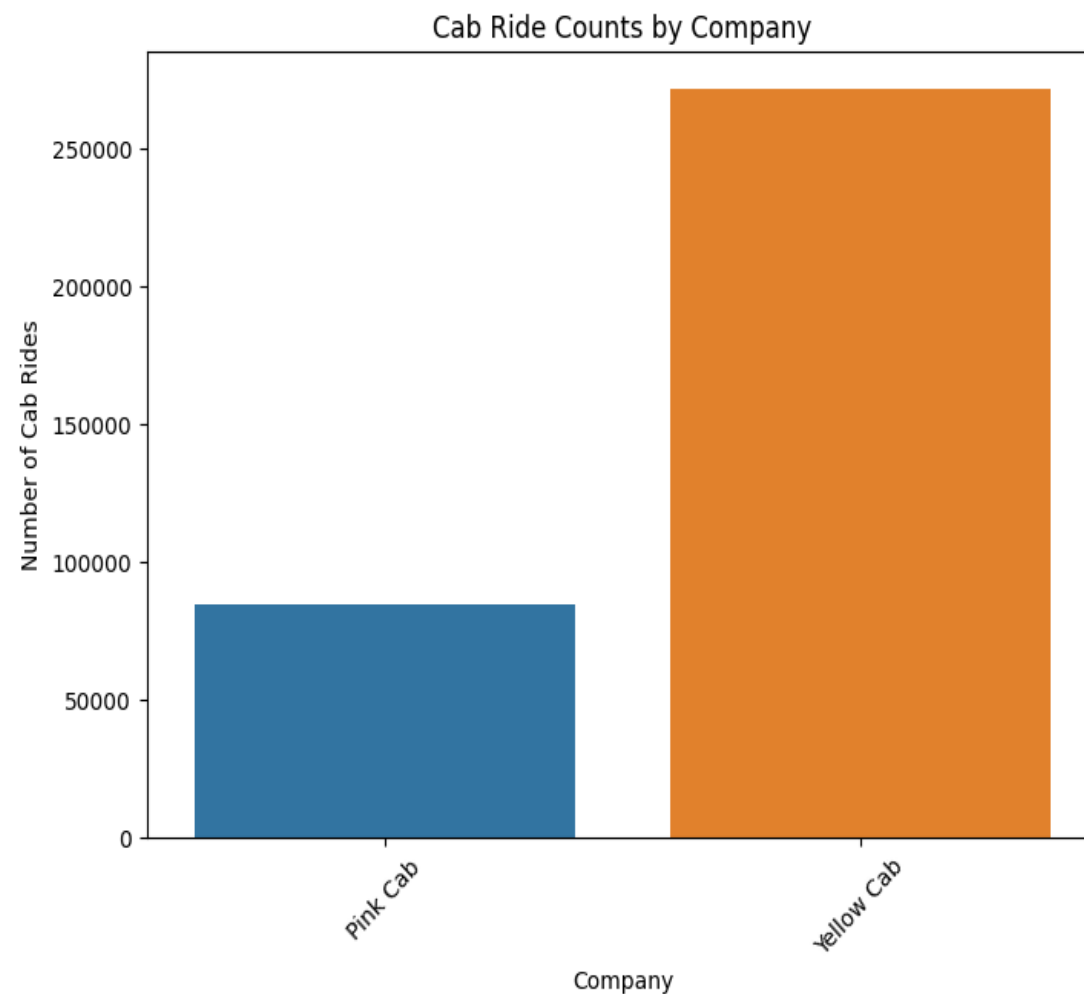
# EDA



Steady, somewhat predictable cycle & seasonality in the industry = less investment risk

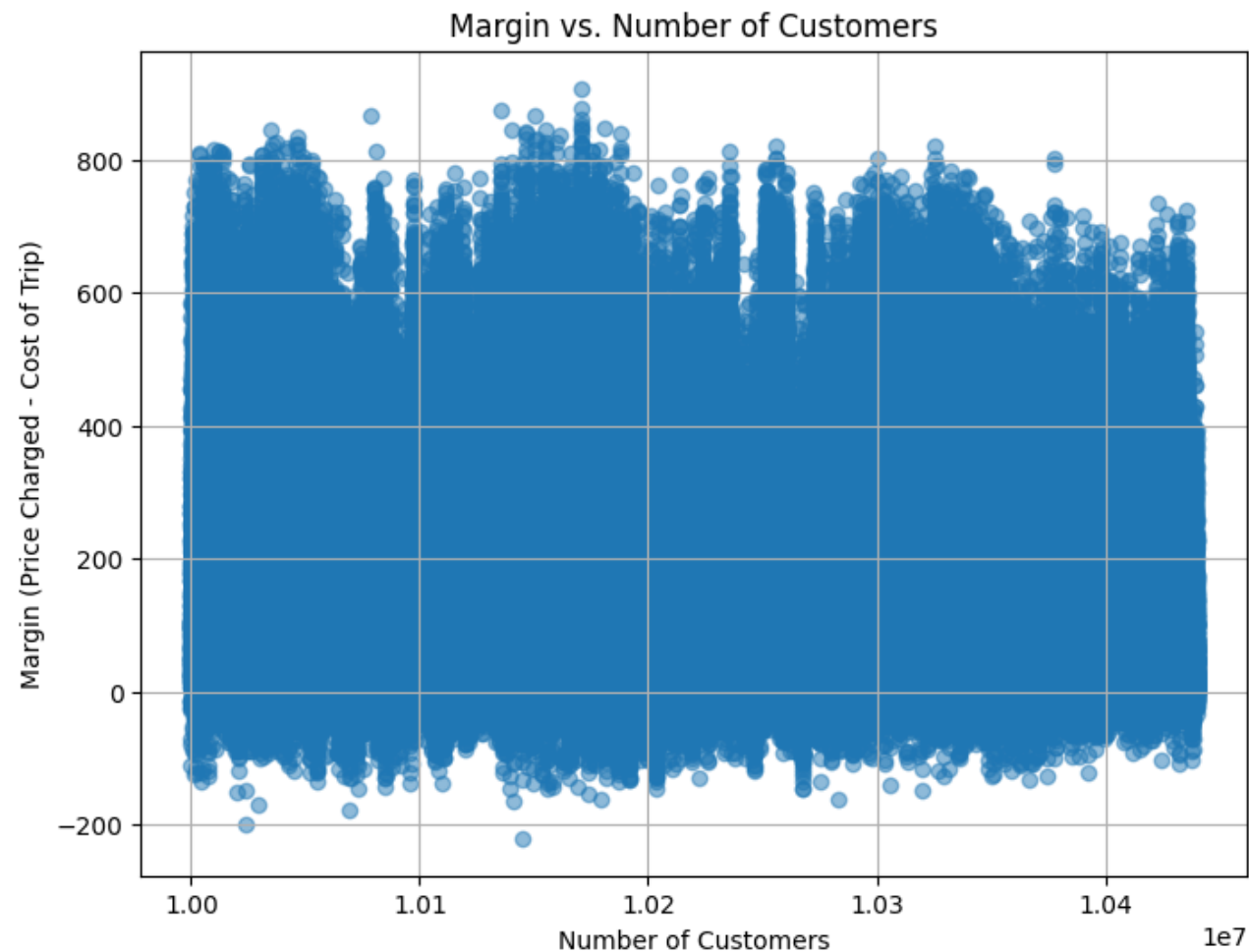


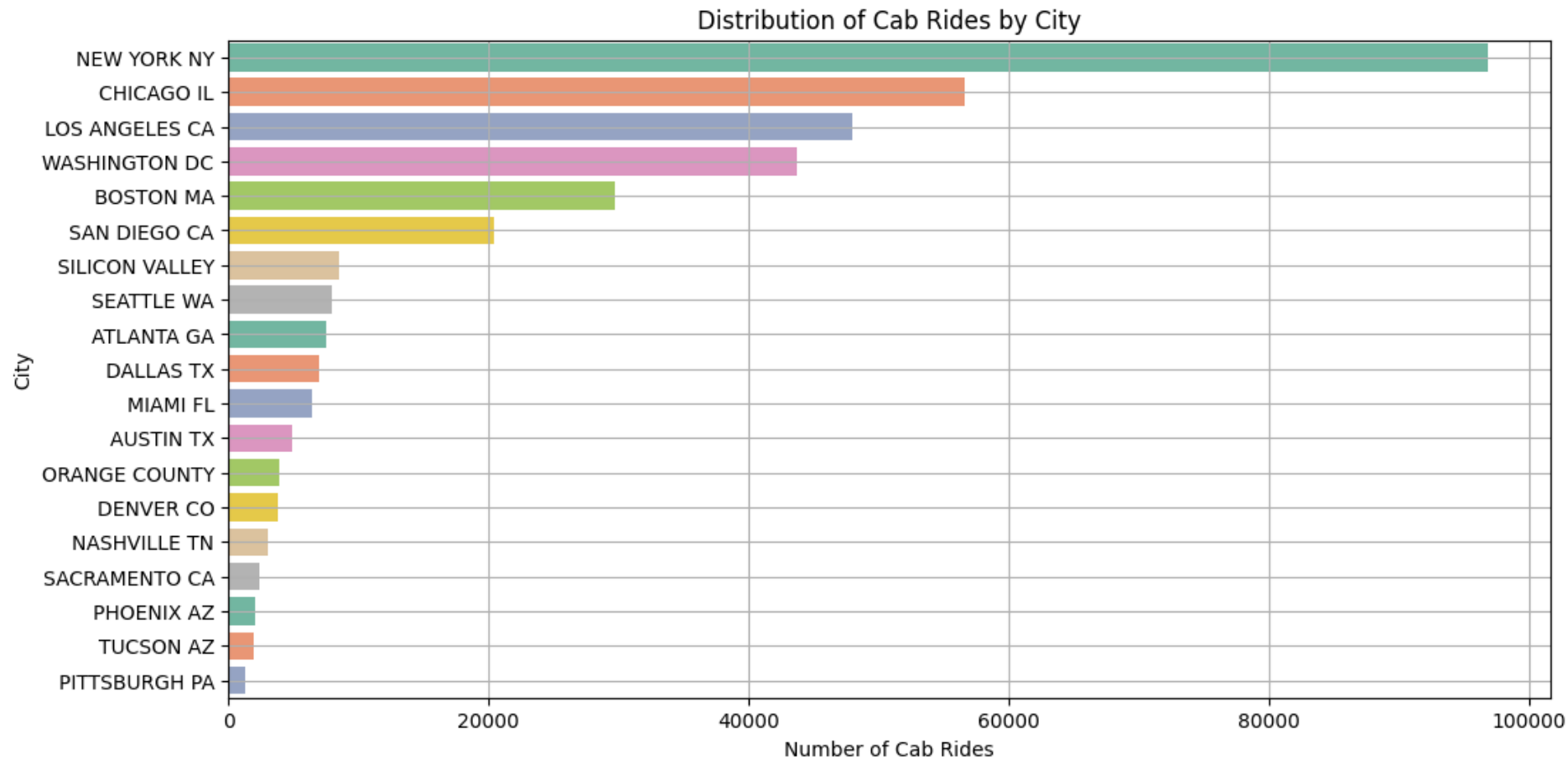
Yellow Cab seems to be dominating the market with over  
**2x Market Capitalization**  
compared to Pink Cab





The relationship between the margin (the profit earned from cab rides) and the number of customers. Each point represents a cab ride transaction



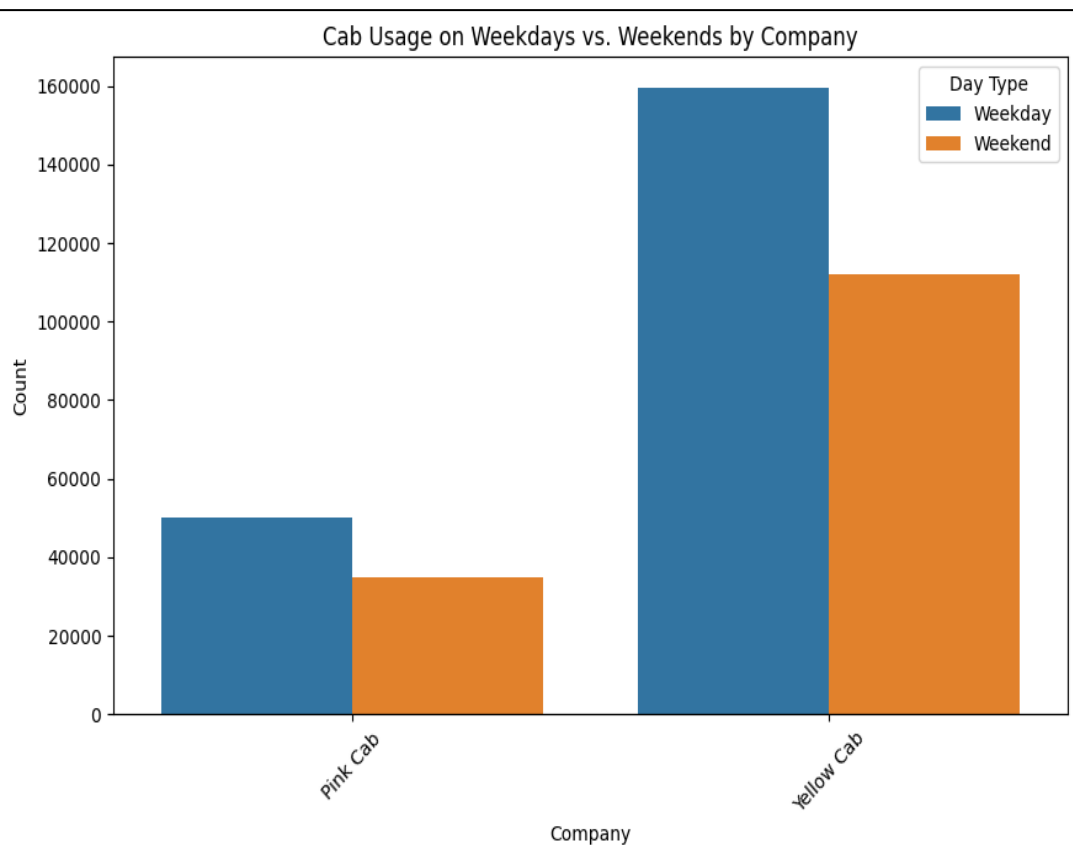


Two densely populated major markets on each (East & West) coast.

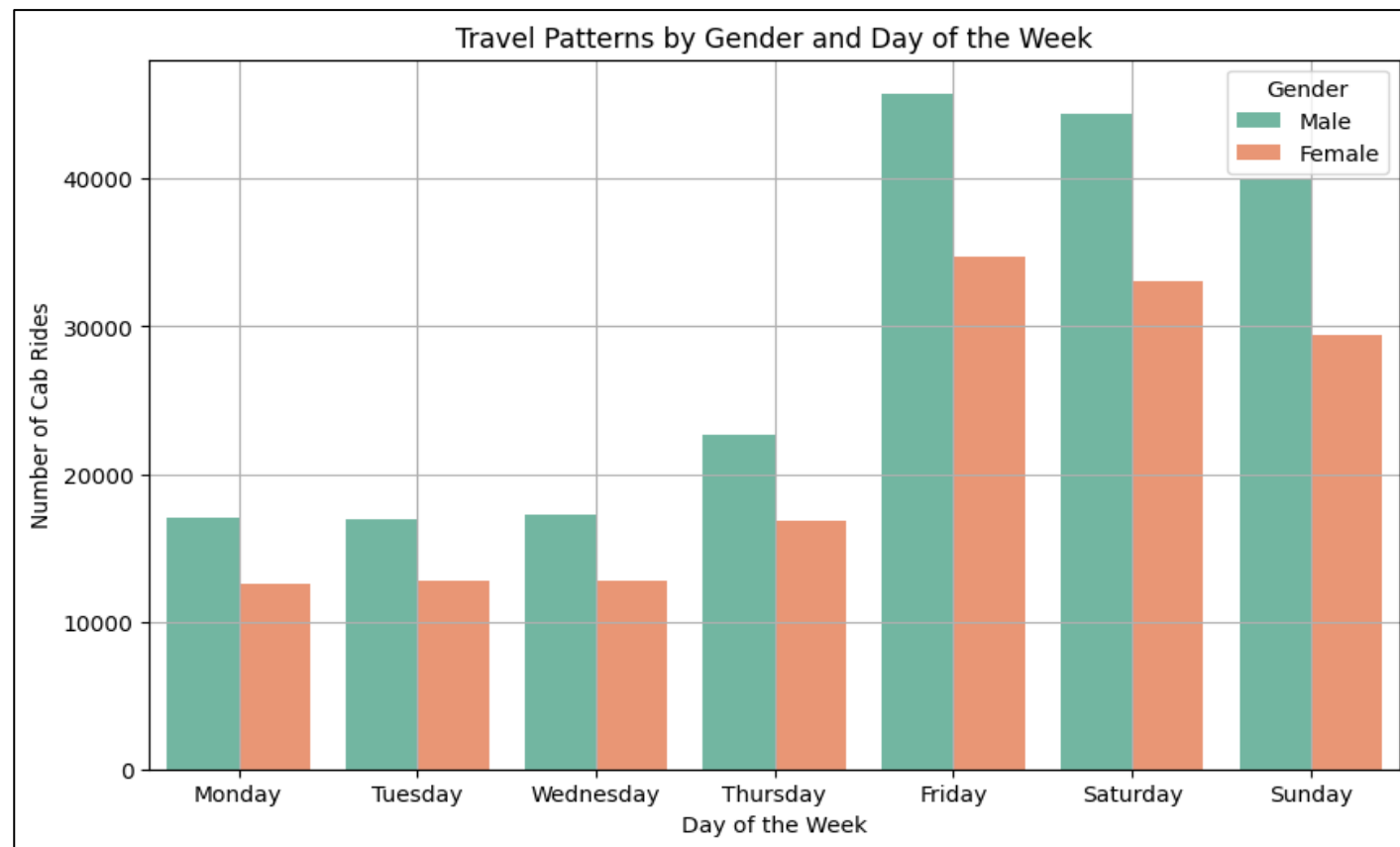




# EDA



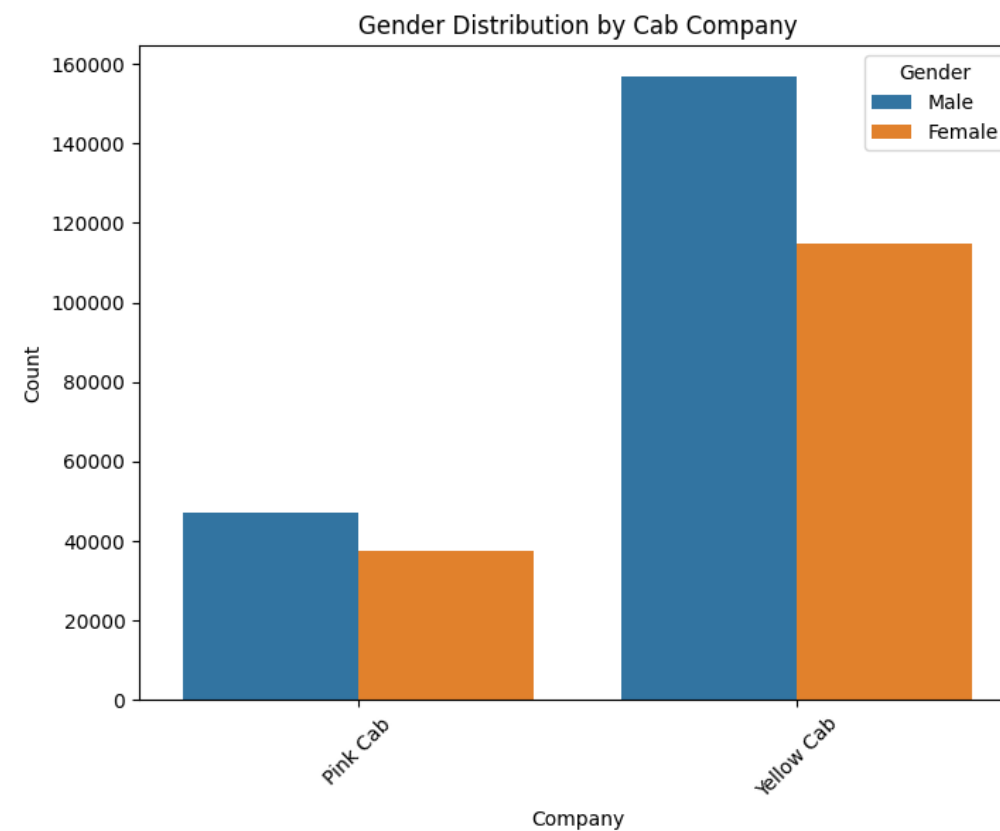
Cab usage patterns between weekdays and weekends for both companies, which can be valuable for making business decisions



The number of cab rides taken by males and females on each day of the week



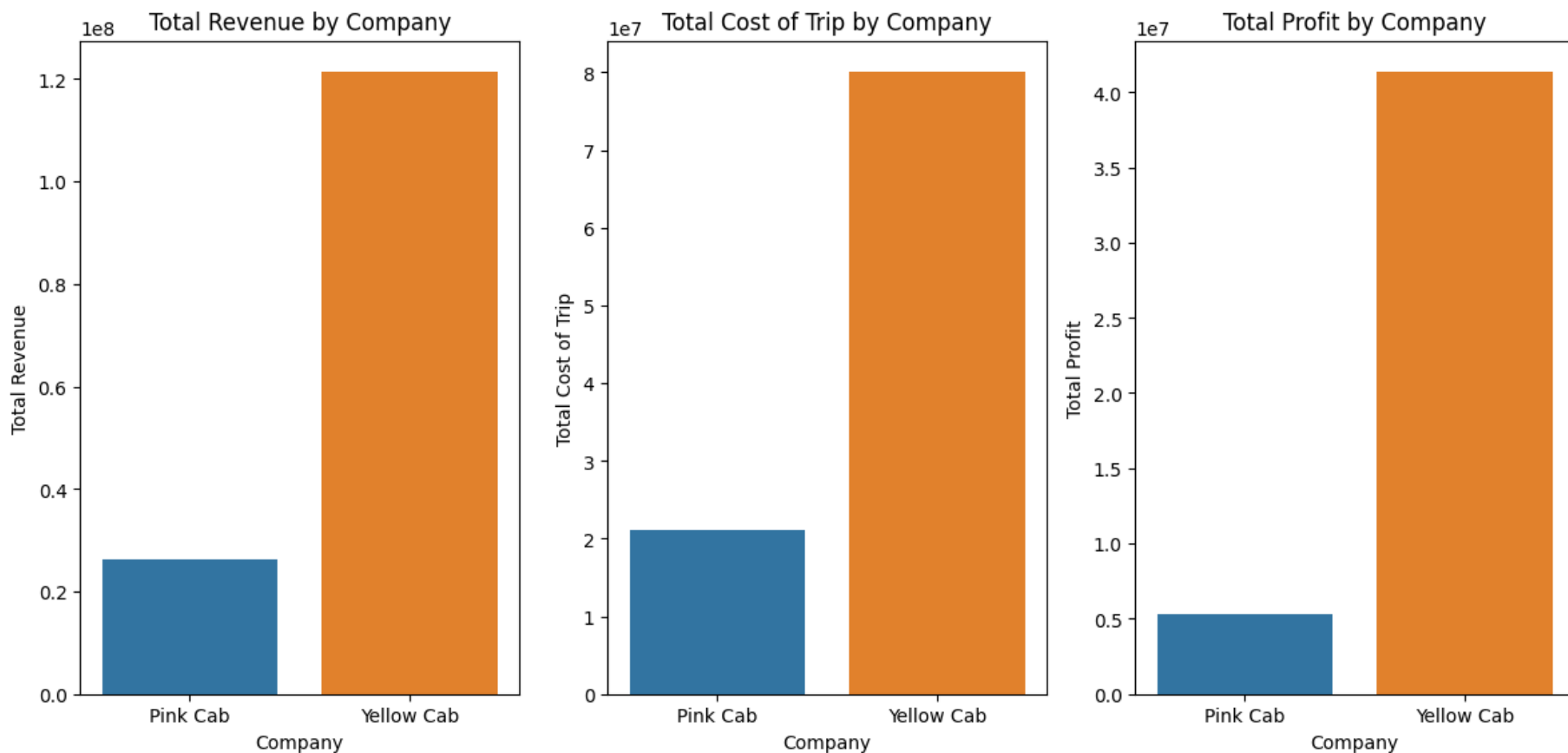
# EDA



The proportion of male and female customers for Pink Cab and Yellow Cab



# EDA



Yellow cab is generating higher revenue & profits while Pink cab has higher costs resulting in lower total profits



# Recommendations

Based on the extensive analysis conducted, including seasonal trends, company performance, and customer segments, it is recommended that XYZ consider investing in Yellow Cab.

- **Company Performance:** Yellow Cab has consistently outperformed Pink Cab in terms of total revenue and profit. This indicates better investment potential for XYZ.
- **Market Dominance:** Yellow Cab's market capitalization is over 2 times that of Pink Cab, indicating a dominant presence in the market.
- **Profitability:** Despite higher costs, Yellow Cab manages to generate higher total profits compared to Pink Cab.
- **Market Insights:** The analysis provides valuable insights into seasonal variations in cab usage and customer behavior, which can be leveraged for strategic decisions.
- **Risk Mitigation:** The somewhat predictable cycle and seasonality in the industry suggest lower investment risk in Yellow Cab.

Overall, the data-driven analysis supports the hypothesis that Yellow Cab presents a stronger investment opportunity for XYZ in the cab industry.

# Thank You