Data Glacier

Your Deep Learning Partner

G2M insight for Cab Investment firm

Bora Engin Deniz

20 July 2023

What Is The 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.

- Our main purpose here is to analyze the data and to offer appropriate recommendations to the company in the light of the data we have obtained.

Data

- **Cab_Data.csv –** this file includes details of transaction for 2 cab companies

- **Customer_ID.csv** – this is a mapping table that contains a unique identifier which links the customer's demographic details

- **Transaction_ID.csv –** this is a mapping table that contains transaction to customer mapping and payment mode

- **City.csv –** this file contains list of US cities, their population and number of cab users

Transformations

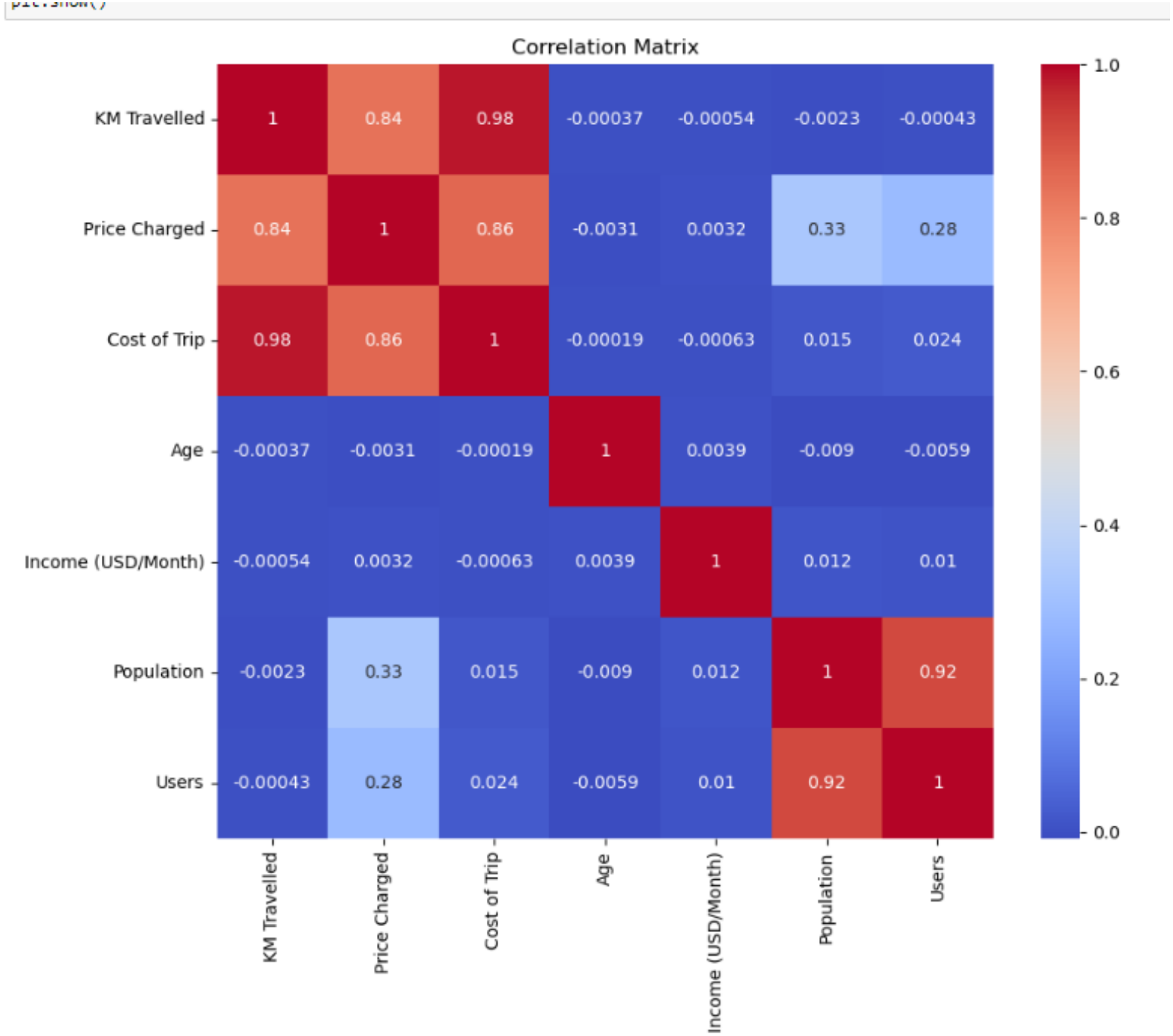
- There were no blank rows or columns in any of the datasets, I didn't need to fill in in general.
- I've implemented a process of converting various required data to the types they should be (like converting string values to integer values).

Relations

You can find the relation table from the "stats.ipynb" file , since seeing from here is really hard.

Relations

Corelations



Data Glacier

Your Deep Learning Partner

Relations

Corelations

From the last slide , we can say that:

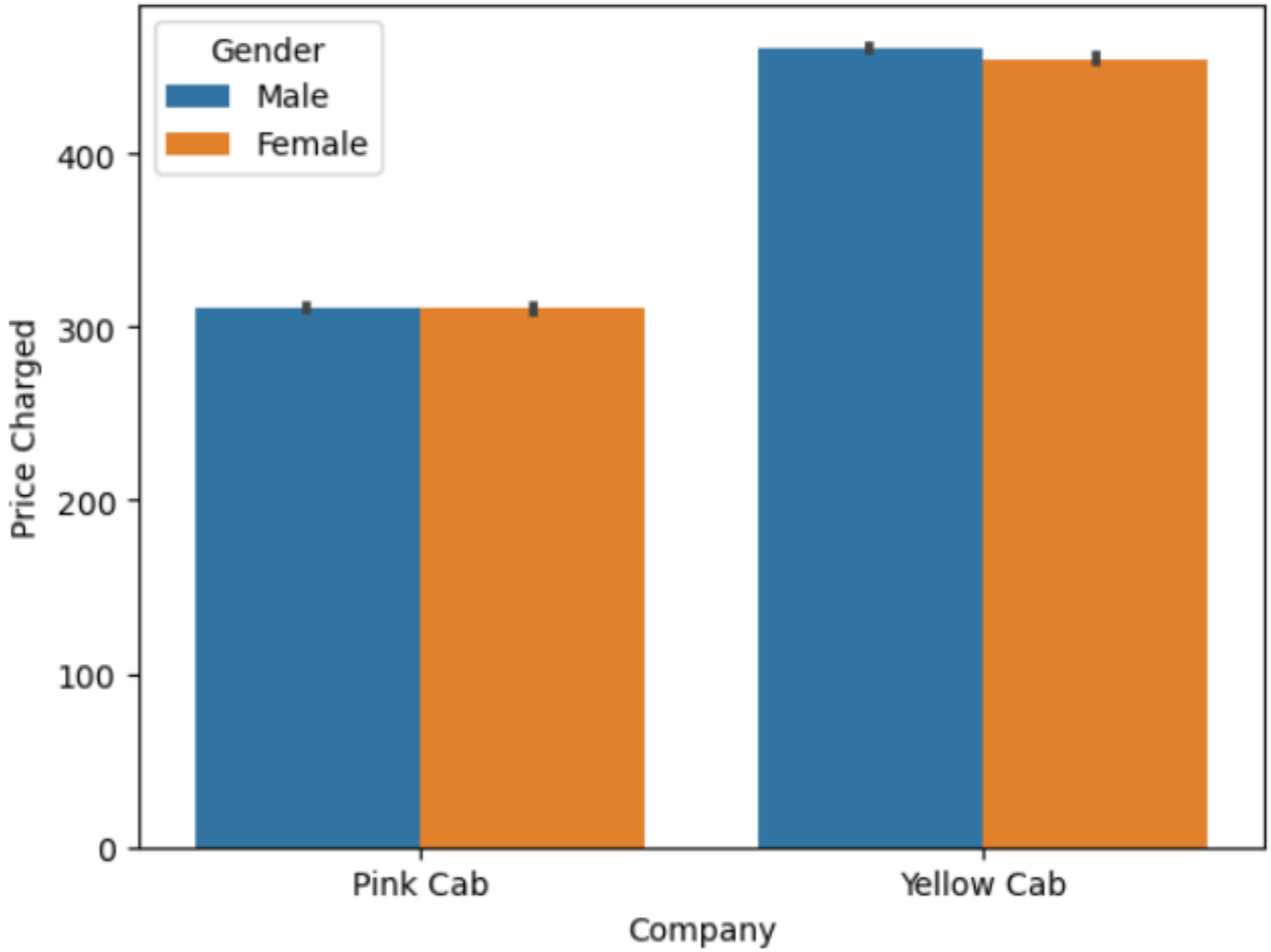
- there is a strong correlation between Population and Users , as we expected.
- We can say that there are very strong connections between Km travelled , Price Charged and Cost of Trip (Especially Km and Cost).
- Price charged is not that depends on User or Population numbers , and there is no such correlation for Age and Income values between other values.

Data Glacier

Your Deep Learning Partner

Analysis

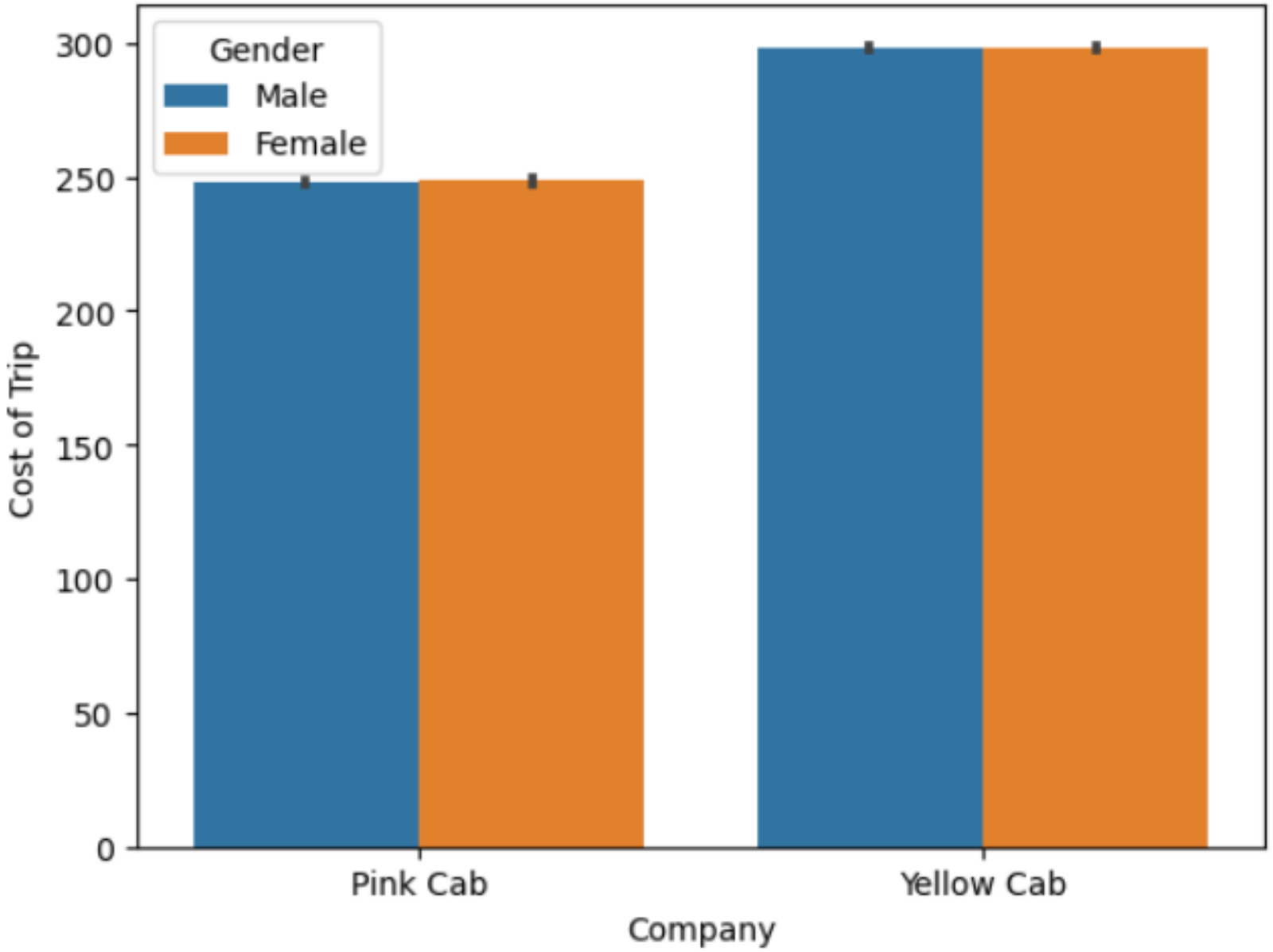
Company-Price Charged



Data Glacier
Your Deep Learning Partner

Analysis

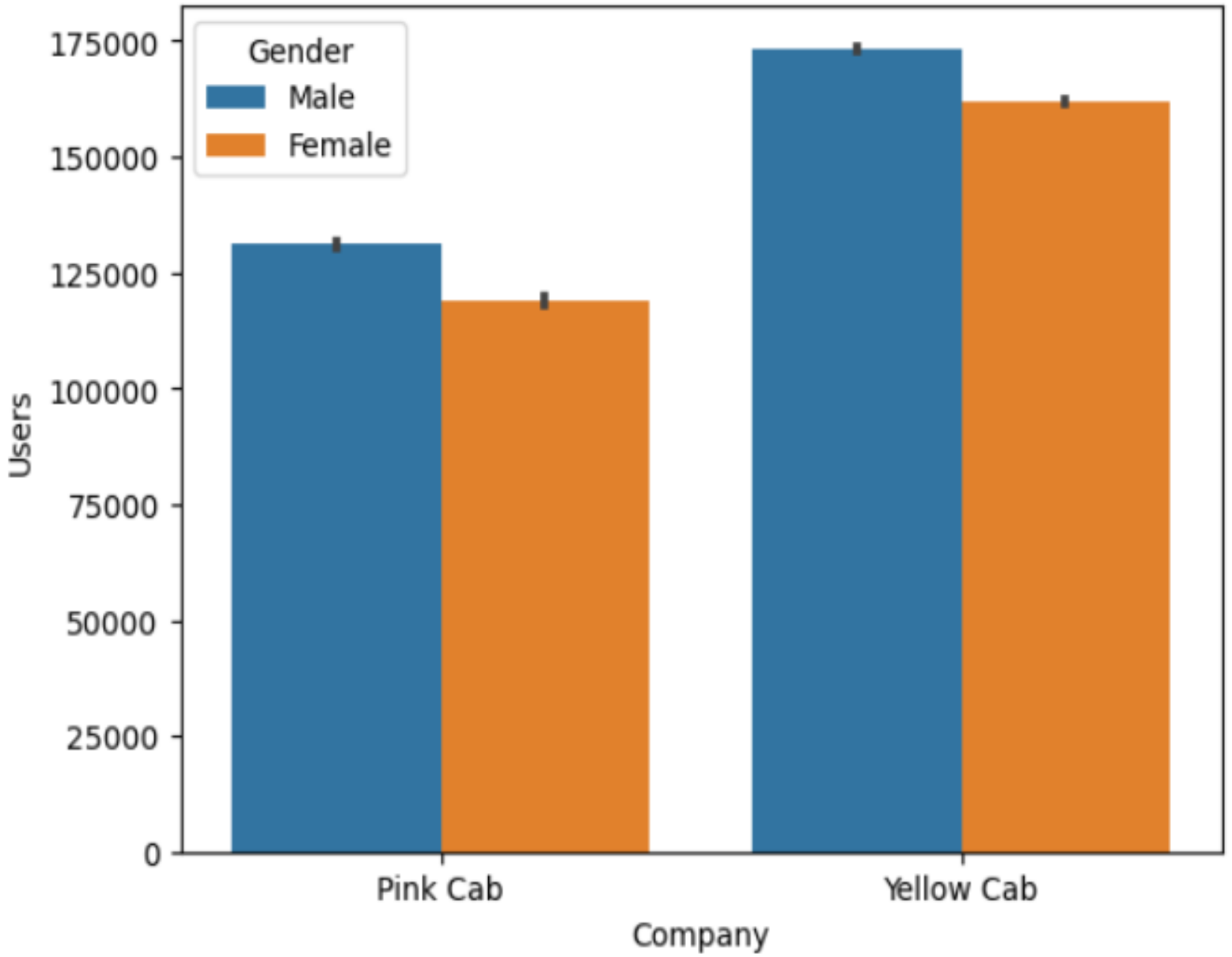
Company-Cost of Trip



Data Glacier
Your Deep Learning Partner

Analysis

Company-Users

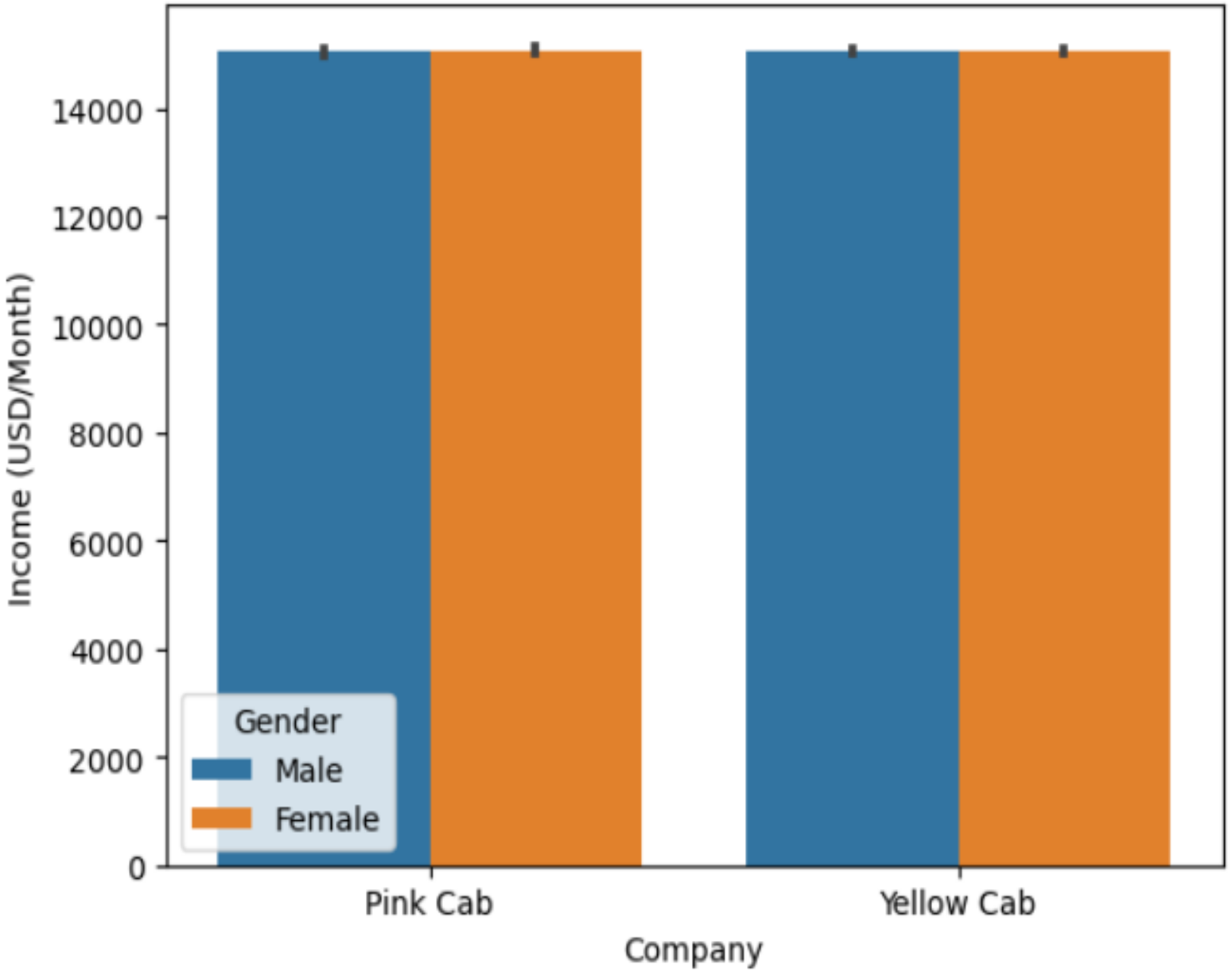


Data Glacier

Your Deep Learning Partner

Analysis

Company-Incomes



Data Glacier

Your Deep Learning Partner

Analysis

From the observations we got from the graphs:

- Yellow cab is a bit expensive to Pink cab even that , Yellow Cab is preferred more than Pink cab so maybe we can say that Yellow cab a safer Company to get a service like this.

Data Glacier

Your Deep Learning Partner

Analysis

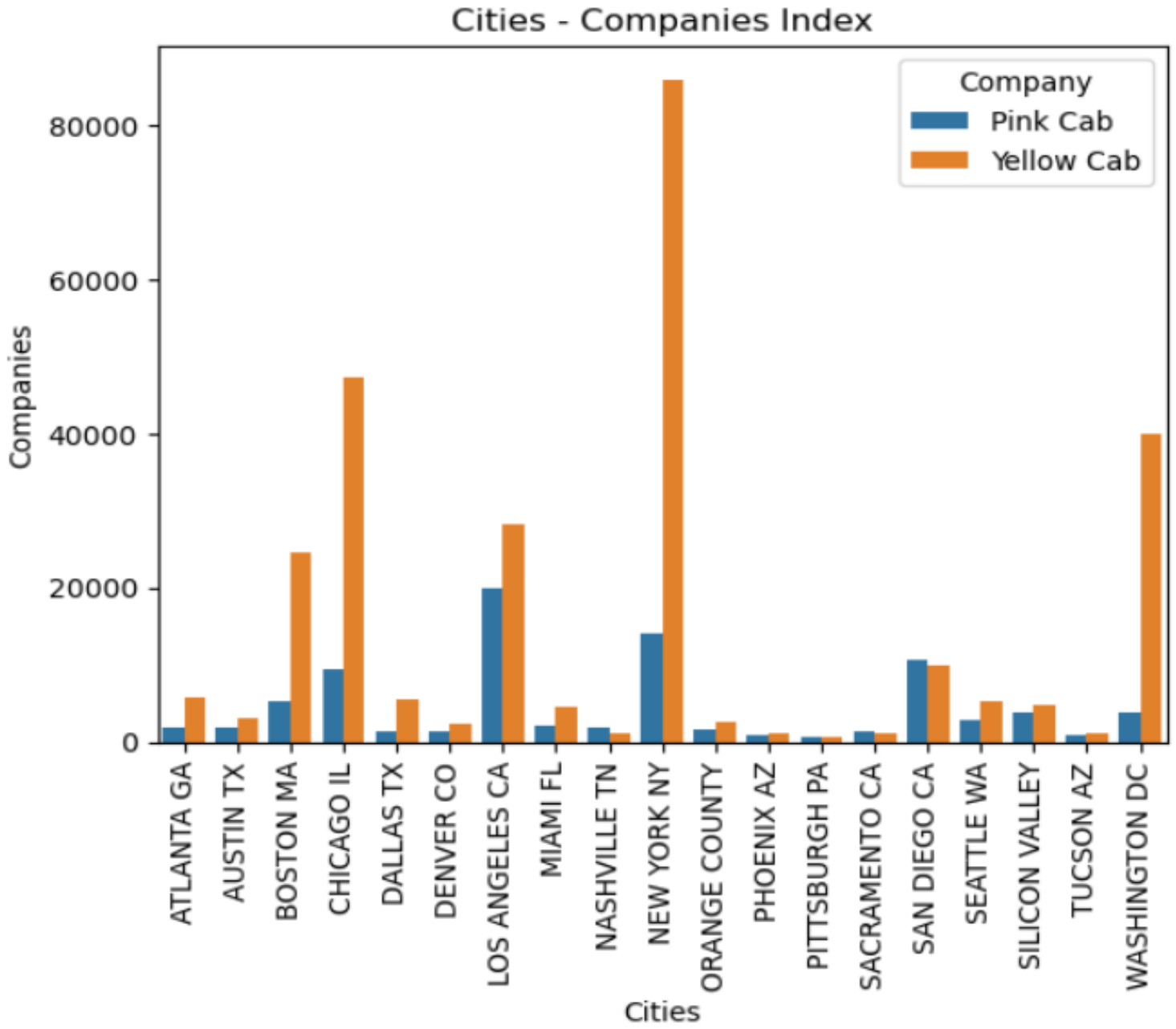
Cities – Price Charged

Data Glacier

Your Deep Learning Partner

Analysis

Cities – Companies

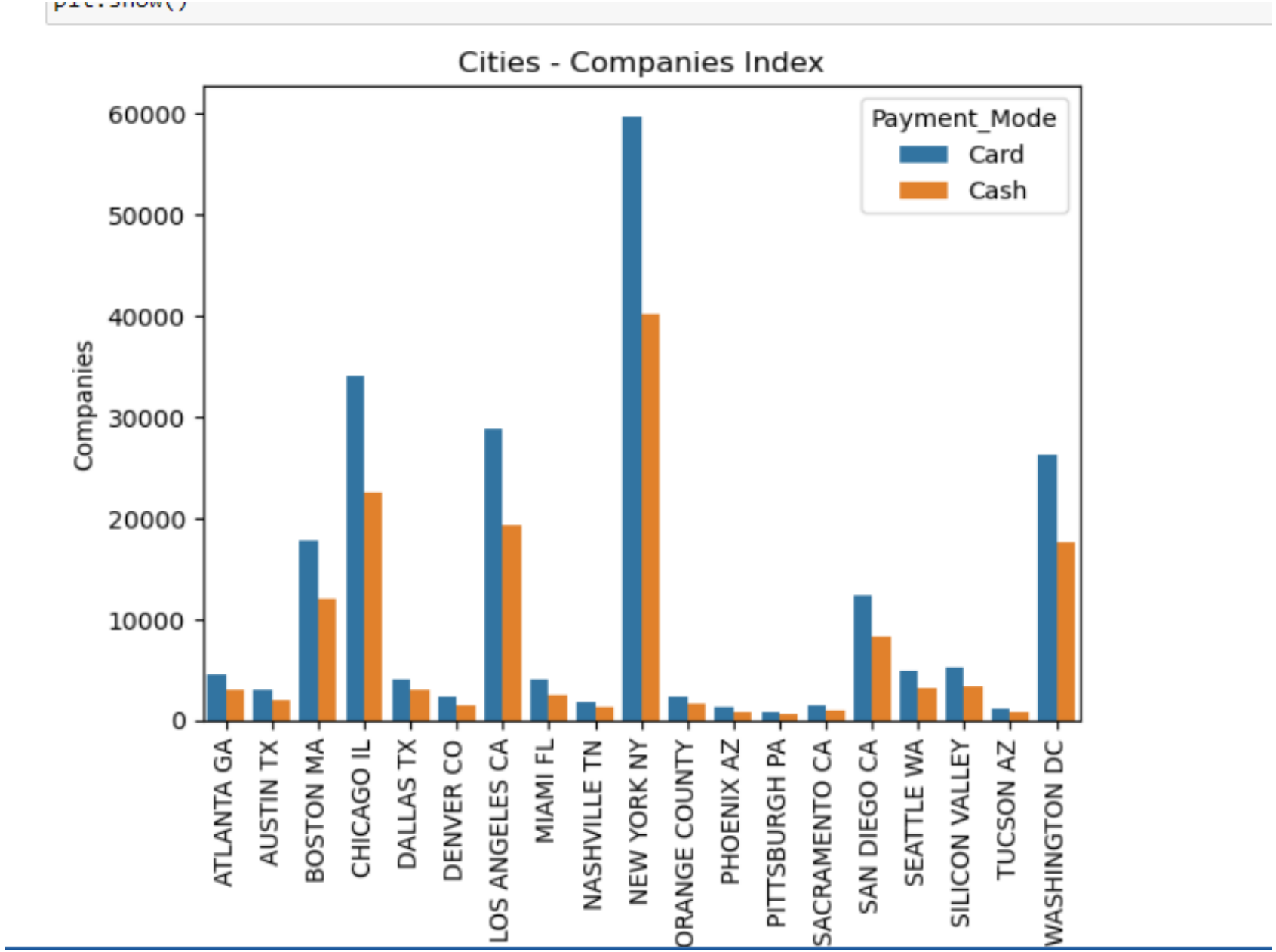


Data Glacier

Your Deep Learning Partner

Analysis

Cities – Companies (Payment Method)



Data Glacier

Your Deep Learning Partner

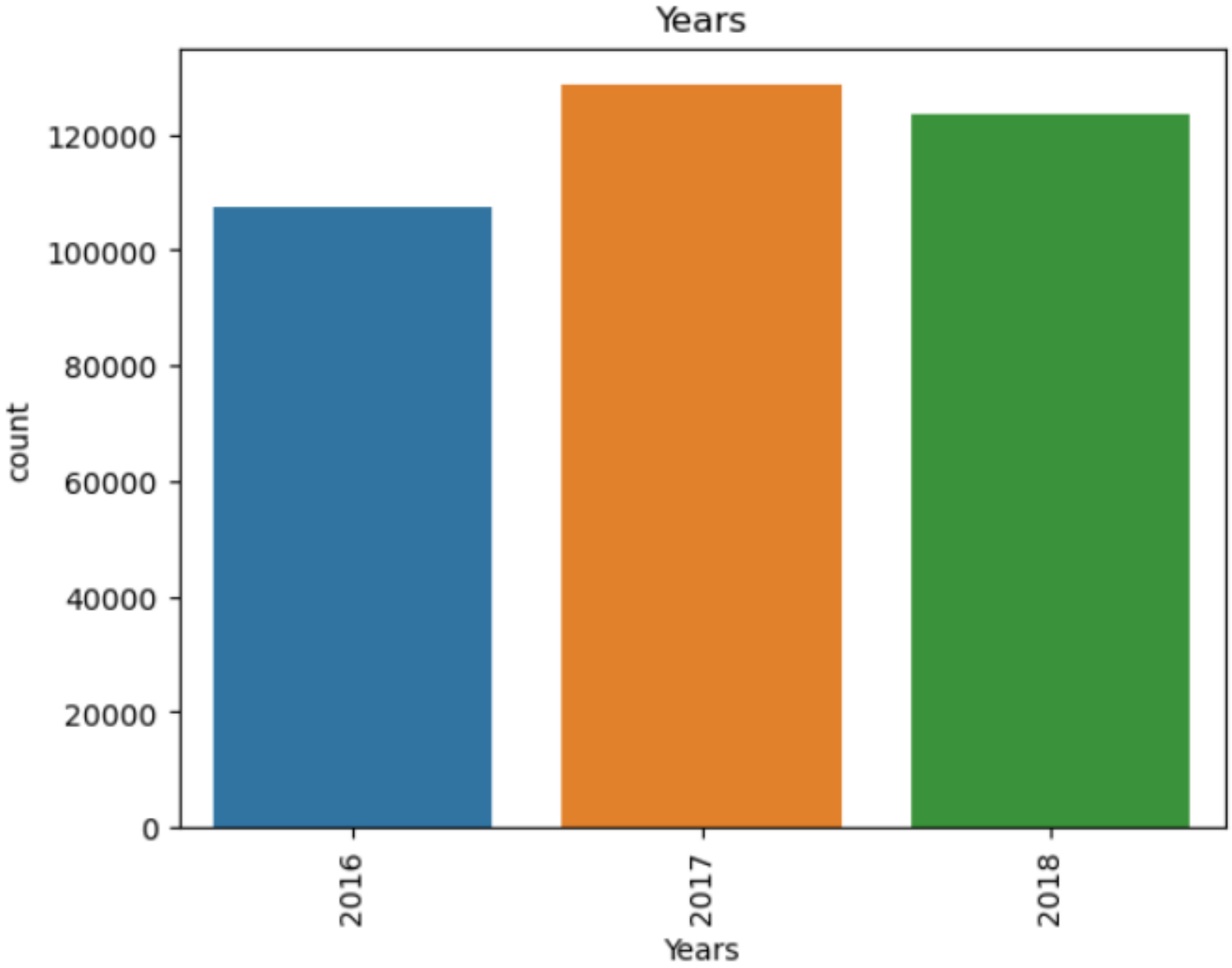
Analysis

From the observations we got from the graphs:

- We can see here that at NYC , The charging value becomes the most , rest of the cities are not that changing. Also we can clearly see that Yellow cab charged more than Pink Cab.
- Clearly Yellow cab is preferred more than Pink Cab , except Nashville TN and San Diego CA.
- Also We can clearly see that Card is preferred more than Cash.

Analysis

Years – Usage Count

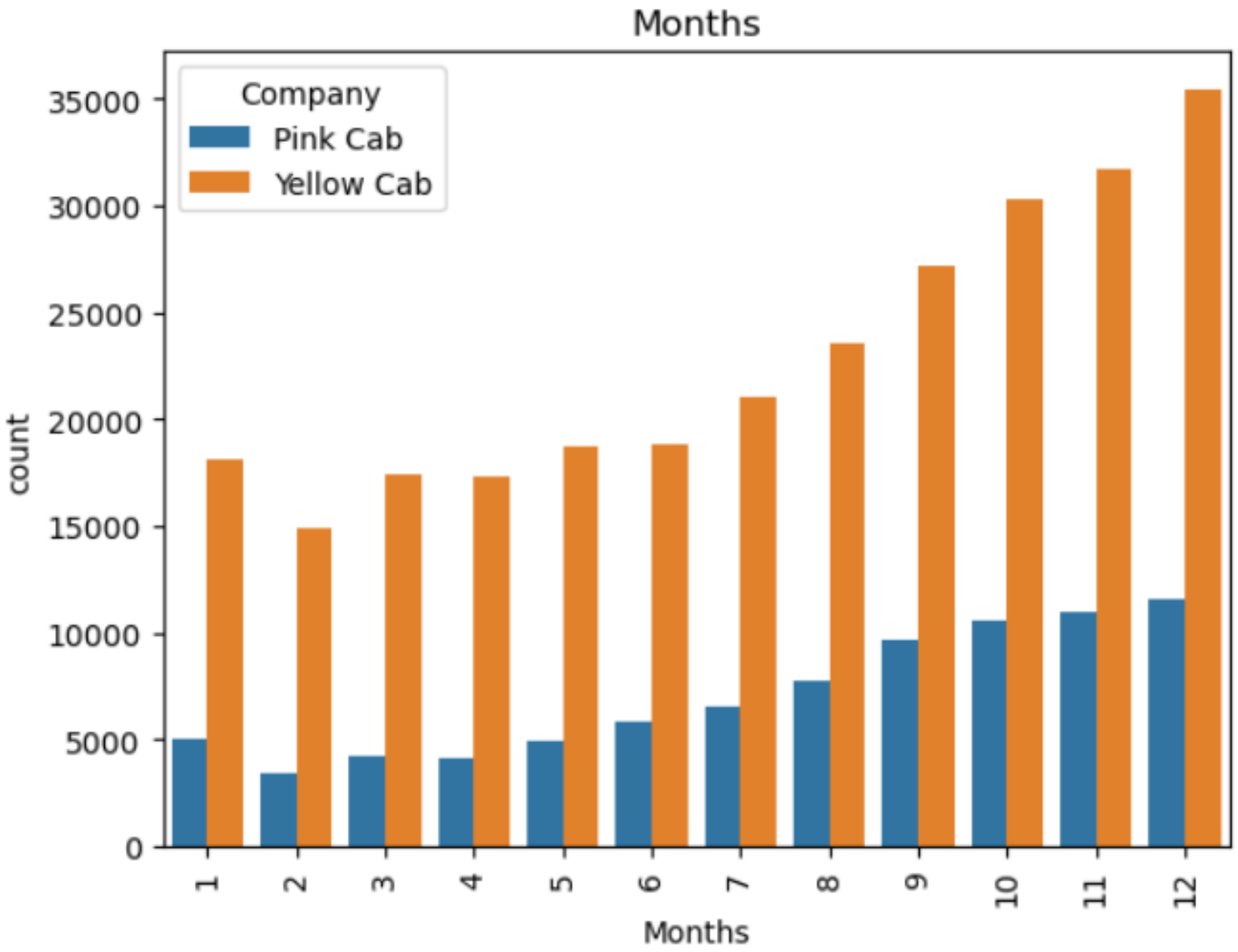


Data Glacier

Your Deep Learning Partner

Analysis

Months – Usage Count

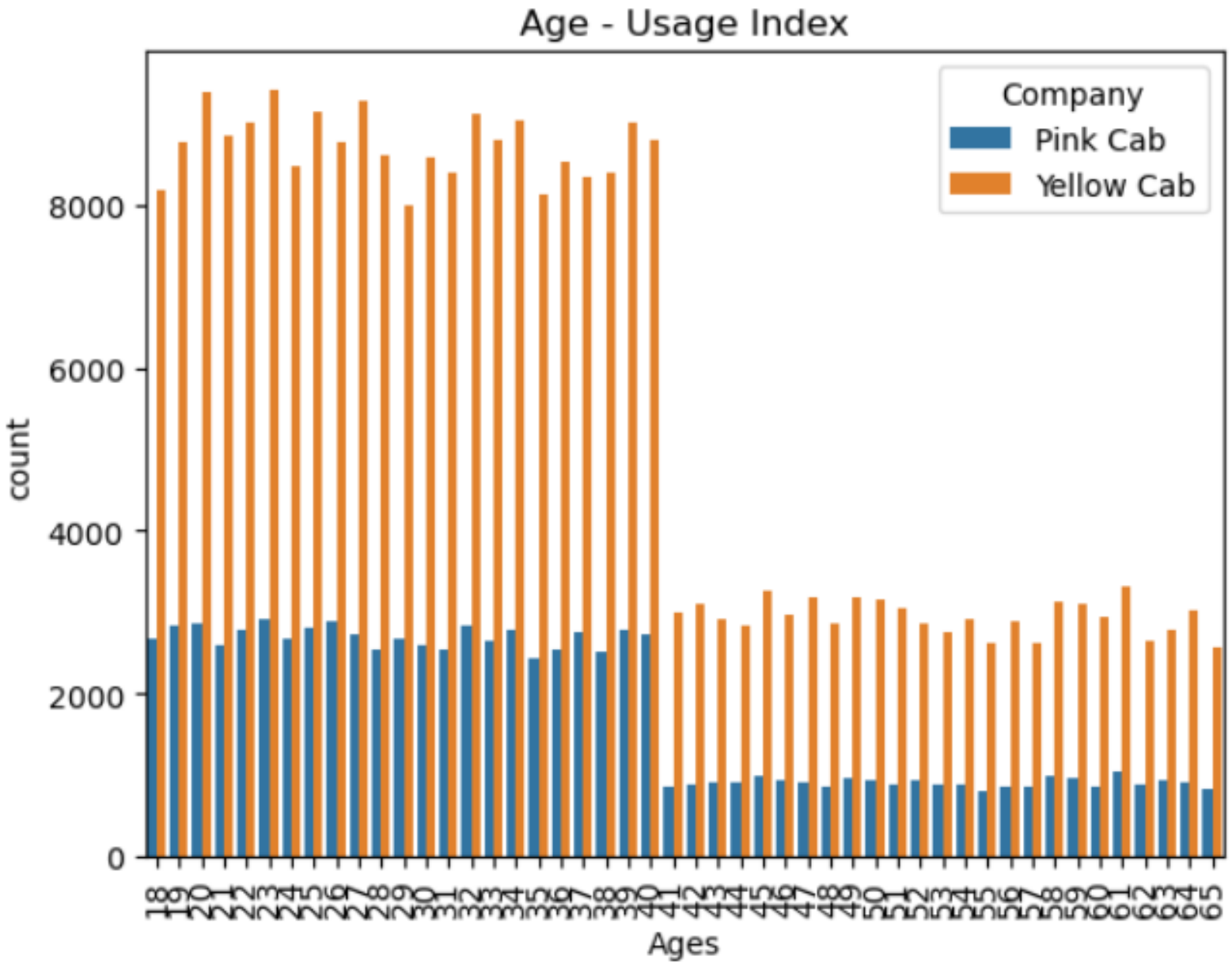


Data Glacier

Your Deep Learning Partner

Analysis

Age – Usage Count

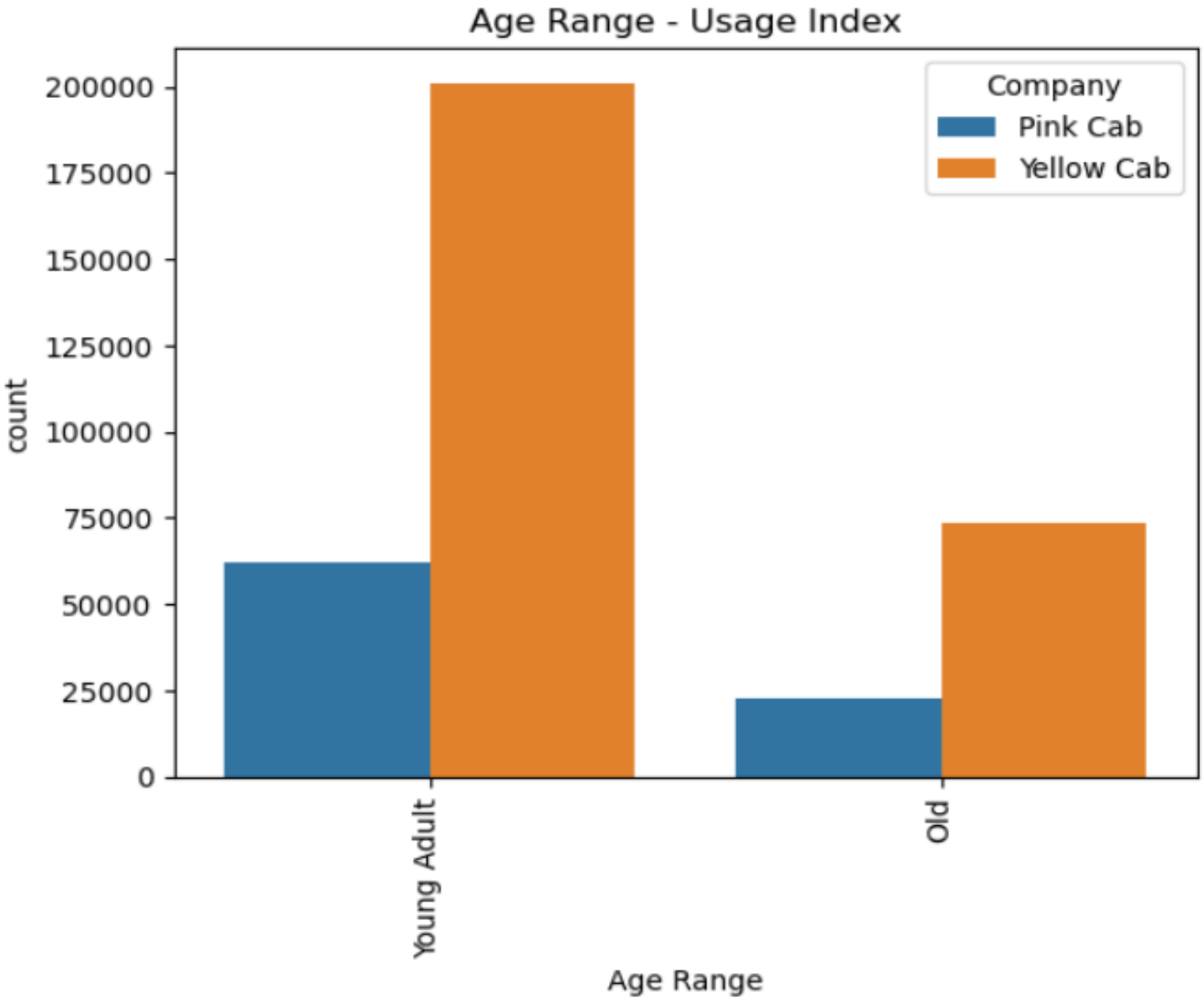


Data Glacier

Your Deep Learning Partner

Analysis

Age Range – Usage Count



Data Glacier

Your Deep Learning Partner

Analysis

From the observations we got from the graphs:

- In 2017 , there was the most usage of cab companies.
- We can clearly see that , especially in Autumn and Winter , there was the most usage .
- We can clearly see that there is a HUGE change when the age value becomes bigger than 40 , mostly young people using this companies.

Data Glacier

Your Deep Learning Partner