**Module 14 Challenge: NYC CitiBike Analysis: (Tableau)**

**Overview:**

The purpose of this analysis is to analyze CityBike usage trends in New York City for the month of August 2019 to determine if it would be a successful business in Des Moines Iowa. The data was obtained from the Citybike website, and then modified in Python to obtain additional features, such as changing the dates to include date time so we could drill down to month, days, and hours in analyzing the trip duration information. We then created a tableau story to visualize the analysis based on trip duration and usage.

[Tableau Workbook CitiBike\_Challenge.twb](file:///C:\Users\13854\Desktop\GITHub_Repo\bikesharing\Challenge\CitiBike_Challenge.twb)

**Analysis Results:**

**Chart, line chart

Description automatically generated**

**Checkout Times by User:** This graph uses the count of trips and trip duration to visualize the most common trip lengths. Most trips were under thirty minutes and the highest frequency are trips of 5 to 6 minutes. Further review of this data might be needed.

**Chart, line chart

Description automatically generated  
Checkout Times by Gender:** The trip durations were also reviewed based on gender, which indicated most trips were taken by males.

**Graphical user interface, PowerPoint

Description automatically generated**

**Trips by Weekday per Hour:** This heatmap indicates most trips were taken during the 7:00-9:00 morning hours and the 4:00-7:00 evening hours during the weekdays, with an even spread on the weekends.

Graphical user interface, application, PowerPoint

Description automatically generated

**Trips By Gender(Weekday per Hour):** Here we can see the most common start times in the previous graph were true for both male and females.

Chart

Description automatically generated with low confidence

**User Trips by Gender by Weekday:** Citi Bike provides a monthly subscription plan for frequent users (Subscribers) and then pay as you use plan for (Customers). This visualization shows the trips by the user type per weekday and gender. Thursday shows the highest usage for Subscribers, while Saturday was more popular for Customers.

Chart, pie chart

Description automatically generatedGraphical user interface, application

Description automatically generated with medium confidence

**User% by Gender:** This confirms most users are male. I changed this graph from numbers to percentages since the number in Des Moines will not be as high as New York due to population differences. I also removed the Unknow gender since that just added noise to the analysis.

Chart, histogram

Description automatically generated

**Average Trip Duration by Age:** In the module we did this graph by birthyear. I thought calculating the age would be a better view. While reviewing the data I found dubious data, such as users over 100 years old. So, I filtered for ages to only include those under 70 thinking it would be more informative. I also exclude the “unknown” gender type as I did in the previous chart to be consistent.

### **Summary:**

From our analysis of the New York Citi data for a single month (Aug 2019) we could identify interesting usage trends and make some conclusion about how the CitiBike business operates in New York.

**Results of NYC Analysis:**

Most users have a monthly subscription for usage, indicating probably that they are frequent if not daily users.

Males between the ages of 31-40 are the most frequent users with the second highest being Males between the ages of 26-30.

Usage is highest on weekdays between the hours of 7:00 & 9:00 in the morning and 4:00 & 7:00 in the afternoon/evening. High probability that these are commuters to and from work.

Trip Durations were relatively short averaging less than a half hour with most of them being less than 10 minutes in duration. This could also indicate that these are commuter related trips verses sightseeing or leisure activity trips.

However, before we can decide to move forward with the idea of opening a similar business in Des Moines, we should do additional analysis of the data and gather additional data. We should look at several months to see if the trends in August are reoccurring throughout the year. It would also be helpful to understand the geography of where the CitiBike locations are located and what is nearby.

**Suggested Additional Analysis:**

Use the Latitude and Longitude information from the NYC data to determine, origin destination and distances of each trip.

To better understand usage, determine the highest frequency by location and then use some geographical data to understand why they locations were popular. (i.e., close to a metro station, nearby bike paths, etc.)

All this information is only valid it we can determine how comparable New York is to Des Moines. The idea of a Citibike business in Des Moines could be successful but it might need to be based on a different business model, perhaps catering to more of the leisure user verses the commuter usage that appears to be common in New York.