Bikesharing

Overview:

In this project, we used Pandas and Tableau to analyze data from the Citi Bike program in New York City, for the month of August, in order to create a bike trip analysis. The ultimate goal is to present a business proposal for a bike-sharing company, using Tableau

The objective of this project are:

* Create effective visualizations from the Citi Bike data in order to impress potential investors with the purpose of convincing them that a bike-sharing program in Des Moines is a solid business proposal.

Resources:

* Data Source: 201908-citibike-tripdata.csv.zip from Citi Bike System Data page
* Software: Python, Jupyter Notebook, Tableau.
* Python code: [NYC\_Citibike\_Tableau \_Challenge\_code](file:///C:\Users\Seanm\OneDrive\Desktop\HYPERION\SCHOOL%20AND%20LIFE\Columbia%20Bootcamp\Module%2015\Tableau%20Bikesharing%20Challenge\NYC_CitiBike_Challenge_Tableau%20Challenge.ipynb)

Results:

Using Tableau, we have created the following visualizations:

1 - NYC Citi Bike Dashboard

[Graphical user interface, timeline

Description automatically generated](file:///C:\Users\Seanm\OneDrive\Desktop\HYPERION\SCHOOL%20AND%20LIFE\Columbia%20Bootcamp\Module%2015)

Based on the picture we can conclude:

* Number of trips recorded: 2,344,224
* Type of users: 1,900,359 trips were checkout by subscribers users and 443,865 by customers.
* Gender of users: 1,530,272 Male - 588,431 Female - 225,521 Unknown. It’s interesting to notice that the number of male users is almost 3 times that of female.
* We can see that ride duration increases when age decreases.

2 - How long bikes are checked out for all riders and by gender?

Graphical user interface, chart

Description automatically generated

* It is worth noting that the checkout time for all users is mainly under 60min with a peak of 5min use for 146,752 users.
* As the time increase, the number of users decreases.

Chart

Description automatically generated

From the above plot, we can conclude:

* The number of male riders is the highest among all genders.
* The average of checkout time is approximately the same for both male and female users: male riders have a peak at 5min while female have a peak at 6min.

By unchecking the first Hour Trip Duration and checking larger amount of time, we can notice:

* The number of riders decreases as the time increases.
* When the amount of time for renting bike increases, the number of rides by male remains higher than that by female. However, the difference in the number of rides between sex becomes relatively smaller.

3- The number of bike trips by weekday for each hour of the day as a heatmap

Graphical user interface, PowerPoint

Description automatically generated

The above heatmap shows:

* The highest number of bike trips is on Thursday.
* The highest number of trips is between 5pm and 7pm of each day of the week. This number decreases as we go after 7pm to reach the the lowest valuesbetween 12am and 4 am.

4- The number of bike trips by gender for each hour of each day of the week as a heatmap

Graphical user interface, application, PowerPoint

Description automatically generatedEven though the numbers are higher for male users, the highest number of trips for each gender remains between 5pm and 7pm.

5- The number of bike trips broken down by gender for each day of the week by each Usertype

Treemap chart

Description automatically generated with medium confidence

* In all days of the week, the probability of male checking a bike is higher than that of a female or unknown gender.
* For Subscribers riders, there's no exact trend of which days are busier than others, however we can notice that the biggest number of renting bikes is on Thursday for all the genders.
* While for Customer riders, one can notice that the number of rented bikes is larger during the weekend (Friday to Saturday), for all gender types. Moreover, Saturday is the day with the biggest number of customers' users. Which is logical, considering customer riders must manly be visitors or tourists while subscribers should mainly be local.

6- The number of bike trips by days of the week

In order to make sure that Wednesday is the day of the week with the lowest numbers of trips, we have created a heat map to show the number of bike trips in each day of the week. As expected, Thursday is the day with most trips while Wednesday is the day with the lowest number of trips.

Chart, waterfall chart

Description automatically generated

7- Suggested Days and Times for Repair

Chart

Description automatically generated with low confidence

* Based on Bike Utilization bubble created before, one could check the bikes ID with most utilization time, in order to check them for repairs.
* By drawing a horizontal bar chart showing the number of bike trips according to each hour of the day and by referring to the heat maps created before we can confirm that the day with the lowest number of rides is Wednesday which could be the best option for bike repairs. While, the hours with the least bike rides are from 9pm to 4am which make these duratin the best one for bike repairs.

Conclusions:

Investing in the bike-sharing program in Des-Moines is a sound business proposal. However, in order to increase the company’s benefits I would suggest:

* Drawing maps in orders to find any correlation between the number of bike trips and household income.
* Finding data about bikes types (city bike, mountain bike, kids bike, tandem bikes, tricycle ...) and sizes; this will provide answers to refine the product available, such questions should include:
  + Which types and sizes are most rented in general?
  + Which types and sizes are most rented by gender and by weekdays and hours?
* Finding out the reason behind the low number of female renting bikes compared to male, in order to find a solution and increase the number of female users.