2023 6th International Conference on Recent Trends in Advance Computing (ICRTAC) | 979-8-3503-9470-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICRTAC59277.2023.10480792

Bike Sharing: An In-Depth Analysis on the Citi Bike Sharing System of Jersey City, NJ

Vijay Kumar Reddy Voddi
Data Science Institue
Saint Peter's University
Jersey City, USA
email address or ORCID

***Abstract*— Jersey City, New Jersey is transforming into business district and there's a large population of working professionals, students and general citizens. The success of CitiBike in NYC has created opportunity for the same in Jersey City and Jersey City had already has 52 stations as of now. The use of CityBike bicycles provides environment friendly and cost of cost friendly mode of transportation. Therefore, improving the efficiency of the Citibike system in Jersey City will lead to a better experience for the consumers in terms of time and availability. In this paper, we have analyzed the data provided by city bike using methodologies such as exploratory analysis, forecast modelling, and geo-spatial analysis. Our main contribution to this paper is based on suggesting ways to improve, optimizing and better in the CitiBike sharing system in Jersey City.**

***Keywords—component, formatting, style, styling, insert** (key words)*

I. Introduction (*Heading 1*)

The CitiBike sharing system of Jersey City allows registered members to pick a bike at any of the docking stations. A person can pick and drop the bike at any sharing station in the CitiBike system for convenience. Citi Bike, NYC bike share program allows members to access unlimited rides during their annual membership period (Gordon- Koven and Levenson 7). There are 52 Citi Bike stations in Jersey City and 750 stations across Queens, Jersey City, Manhattan, and Brooklyn. The annual membership enables individuals to access the entire system. For a $95 annual membership fee, individuals can ride any bicycle for as long as they want (Gordon-Koven and Levenson 7). The bike sharing system in Jersey City started in 2013, and it has significantly expanded to provide the services to a higher number of riders. Alternatively, the bike sharing system charges $9.95 per day for a 30-minute ride or $25 weekly. CitiBike offers special rates to Citi® cardholders, corporate partners, and Oxford® members. The bike sharing system encourages individuals to exercise, improves public health, and reduces the healthcare budget. Currently, the bike sharing system business model is successful, but there is a need to introduce new innovative ideas to increase the efficiency of the current system. The CitiBike system can optimize service delivery through effective distribution of bicycles in all sharing stations based on the expected demand.

The bike-sharing system is designed for quick trips and convenience and as an affordable way to get around by picking a bike from any station in Manhattan, Jersey City, Queens, or Brooklyn (Gordon-Koven and Levenson 7). Jersey City has a large population of working class, students, and general citizens who require a reliable transport system within Manhattan, Jersey City, Queens, or Brooklyn. Also, the bike sharing system is environmentally friendly and cost-effective. Thus, improving the current existing bike sharing system will enhance efficiency and create value for clients. The riders save time by using a more convenient and reliable mode of transport (O'Mahony and Shmoys 687-688). Citi Bike connects to diverse modes of transportation and forms an integral part of the community's transport system. Bike rides have reduced traveling time and increased the accessibility of many areas in the cities (Gordon-Koven and Levenson, 4-5). The analysis of the data will help to develop strategies and provide recommendations to enhance efficiency in the Citi Bike system.

Bike-sharing programs have transformed urban cities in the United States and other parts of the world since their establishment in Amsterdam in 1965 (Fritz 2). The initiative enhances public transport by reducing congestion, minimizing cost, and reducing environmental pollution. Bike sharing systems should focus on expanding their network to increase the number of customers. Fritz (31) found out that improving security can increase the number of female cyclists. Also, Citi Bike can attract more tourist cyclists by locating the sharing stations near motorized traffic.

According to Singhvi et al. (111-114), the population density influences the use of bikes for transformation. Previous studies analyze the influence of demographics and bicycle lanes on the demand for bike- sharing systems (Singhvi et al. 111-114). Also, the use of technology such as smart cards improves the collection of data for the study. The analysis of the data collected through technology provides the basis for generating innovative ideas to develop the effective bike-sharing system. For example, the existing bike-sharing systems can become more effective by enhancing safety and adding more bike lanes in the city.

The feasibility of the bike sharing such as CitiBike system depends on the area it serves, the location of the docking stations, and the population size (Rixey 1-15). Also, the author suggests that bike-sharing system operators can consider the relocation of underutilized docking stations closer to the central network, but they should address equity by offering the services to the underserved communities (Rixey 10-15). The study also found out that the demand for bike-sharing services depend on income, availability of alternative mode of transport, and level of education in populations near a docking station (12). Moreover, the non-white population is less likely to use bikes for transport. However, there is no difference in demand for ridership by considering race and level of income in the population.

"The arrival and departure rates at one docking station are related to the bikes flow rates in other stations within the bike sharing system" (Faghih-Imani and Eluru 218-227). Therefore, the bicycle sharing systems are interconnected and ignoring the effect may lead to biased results. To solve 2 the problem, Faghih-Imani and Eluru (218- 227) studied both observable and unobservable features on the bike's picking and return rates and found out that incorporation of observable and unobservable interactions between bicycle sharing stations improves the accuracy of the results and enhance the predictive capability of the system.

Faghih-Imani and Eluru (218-227) also proposed a model to analyze the bicycle sharing stations to determine the impact of redistribution on the usage of docking stations. The model uses continuous variables and categorical variables. The continuous variables include hourly arrivals, hourly departures, temperature, station, capacity, population density, and humidity, while the categorical variables include rainy weather, weekends, substation, and path train station. Therefore, an in-depth analysis of data on bike usage can help in decision making regarding various issues, such as anomaly detection, relocation of bicycles, docking station location, and traffic prediction.

Chang et al. (1-16) conducted a research in Hangzhou, China, and Chicago, USA on bicycle data flow. The study considered the location of docking stations, trips per day, average hourly trips, and average weekly trips. According to Chang et al. (1-16), Chicago experiences high usage of bikes on rush hours and workdays. The results indicate that in Chicago, most users depend on the bike-sharing system to commute between home and workplace. However, Hangzhou reported a stable bikes usage over the week and only one day in a week has two peaks. The results provide valuable information for the repositioning of the bikes from one station to another to meet the demand during peak hours (Aviv et al. 2). Moreover, repositioning can be conducted at night when there is low usage or during the day when the demand for bike changes rapidly.

Shaheen argues that a high number of bike- sharing clients in Montreal, Toronto, Washington DC, and Minneapolis commute to school and work, and the second group of users include social entertainment and errands. Also, people use the bike sharing systems in cities as an extension of the existing public transport. Consequently, customers prefer bicycles to increase physical exercise and improve connectivity. Therefore, the use of bikes sharing systems has health benefits. Besides, it improves the quality of life and reduces health care costs through increased physical exercising.

Gordon-Koven and Levenson (4) argue that densely populated areas require more bikes and new sharing stations to reduce congestion in the existing system. Surprisingly, Rixey's study has similar results. The study shows that there is a positive correlation between the location of bus stops and colleges in the United States (Rixey 10-15). Therefore, new Citi Bike docking stations should be near institutions of higher learning and bus stops to improve their accessibility by college students and staff. However, there is an information gap on how to improve the current Citi Bike system to increase efficiency. Zhang et al. (3-5) argue that only a few studies have been conducted to determine the impact of the expansion of the system to its efficiency.

Examining the trend and improving efficiency through the strategic location of the sharing stations can enhance Cite Bike's performance. Gordon-Koven 3 and Levenson (5) also found that Citi Bike can use a smartphone and web-based applications to make easier for riders to access information about the availability of bikes in different dock stations at any given time. The ability to access real-time information makes the bike-sharing system more convenient for riders.

A study conducted by O'Mahony and Shmoys (687) shows that bike sharing systems face optimization problems. One of the primary challenges is rebalancing the bicycles to meet the changing demand during peak hours. In this case, rebalancing the demand during the mid-rush hour and overnight requires a different approach to enhance efficiency. Citi Bike can use the data to plan routes and the number of bikes to move between the sharing stations (688). The bike sharing system achieves optimal performance by making data-driven decisions to reduce asymmetry in the number of bicycles in each dock station at any given time.

The CitiBike sharing system of Jersey City allows registered members to pick a bike at any of the docking stations. A person can pick and drop the bike at any sharing station in the Citi Bike System for convenience. Citi Bike, NYC bike share program allows members to access unlimited rides during their annual membership period (Gordon- Koven and Levenson 7). There are 52 Citi Bike stations in Jersey City and 750 stations across Queens, Jersey City, Manhattan, and Brooklyn. The annual membership enables individuals to access the entire system. For a $95 annual membership fee, individuals can ride any bicycle for as long as they want (Gordon-Koven and Levenson 7). The bike sharing system in Jersey City started in 2013, and it has significantly expanded to provide the services to a higher number of riders.

Due to the increased cases of traffic congestion, air pollution, and transport cost and the rapid growing population in Jersey City. There was a need to come up with a bike sharing program in the city so that the general population may benefit from the program (Chen, Zhang, Pan, Ma, Yang, Kushlev & Li, 2015). Bike sharing concept dates back to the 1920s. however, the concept was slow on development since the level of technology was very low at the moment. Therefore, this research was aimed at analyzing the data from one of the bike firms so that they may determine how the number of trips may be determined in future. But with the advancement of technology, bike sharing has gain popularity significantly at a higher rate. This might be due to the fact that the programs have a low implementation cost and are easy to implement in comparison to other means of transport. Also, this is an easy win for urban societies and the government since the scheme is a design which is environmental friendly.

II. DATA

Since the research as focusing on the assessment of the trips number and duration, the study used a secondary data called CitiBike Trip Data 'provided by Motivate, a bike share data firm. The study preferred the use of secondary data because it is less time consuming, there is high level of accuracy and reliability, and it cut on cost (cost effective). The dataset was based on the months of January to December 2017 and January to September 2018. The firm had kept the data separately in 21 different 4 sites and thus for the ease of analysis, we combined those files. The data comprised of 564,567 records where the number of the variable was 15. This implies that thee total sample size for the study was given by 564,567. Therefore, the sample size was adequate and representative. A higher sample size ensures that the margin of error is reduced and this leads to accuracy results and valid conclusions concerning the data are drawn. In order to enhance data analysis and apply different methods, we added the number of 9 new variables in the study. The table below indicates the list of the added variables.

Variable	Notes
Trip Duration	In seconds.
Start Time	With date.
Stop Time	With date.
Start Station ID	
Start Station Name	
Start Station Latitude	
Start Station Longitude	
End Station ID	
End Station Name	
End Station Latitude	
End Station Longitude	
Bike ID	
User Type	Subscriber or Customer.
Birth Year	

Table 1: List of Variables

Variable	Notes
Age	
Age Group	Age buckets.
Month	
Day Year	Day of the year.
Day Week	Day of the week.
Day Month	Day of the month.
Hour	Hour of the day.
Date	
Distance	Distance between stations.

Table 2: List of Variables added.

The 9 variables added in the study include: Age (continuous variable), Age Group (categorical variable), Month (qualitative variable), Day of the year (continuous variable), Day of the week (qualitative variable), Day of the month (discrete variable), Hour of the day (discrete variable), Date, and Distance between stations (continuous).

III. METHODOLOGY AND RESULTS

A. Processing

First, we added a unique variable to the dataset called Trip ID and renamed all the variables. Then we checked for NA's and the data type to ensure that there were no missing observations and that the variables format entered was correct. Then we checked for NA's and the data type where we found that the dataset had 0 NA values and all the data types were correct except for Start Time, Stop Time, and Gender. Start Time and Stop Time were converted to datetime while Gender was coded as a dummy variable i.e. 1=male, 2=female, and 0=unknown. There was a need to create and 5 add new variables in the dataset so that more information could be added to the dataset. This would in turn help in data analysis, interpretation and making of valid conclusion. This is shown in the table below.

Variable	Notes
Trip ID	Unique.
Age	
Age Group	Age buckets.
Month	
Day Year	Day of the year.
Day Week	Day of the week.
Day Month	Day of the month.
Hour	Hour of the day.
Date	
Distance	Distance between stations.

Table 3: List of variables added.

This is how the variables are measured; The End Station ID is measure on a continuous scale, End Station Name is a qualitative variable, End Station Latitude and Longitude are continuous variable, Bike ID is a continuous variable, User Type is a qualitative variable, Birth Year is a continuous variable, and Gender is a categorical variable.

B. Exploratory Data Analysis

The below graphs show a visual exploratory analysis of total Trips by Day, Hour and Time

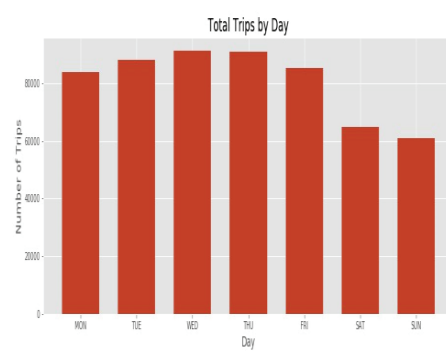

Figure 1: Total trips by day

The total trips by day graph indicates that Wednesday had the highest number of trips followed by Thursday. Also, Sunday has the lowest number of trips per day followed by Saturday. The number of trips increase from Monday through to Wednesday and starts declining all the way from Thursday to Sunday. Therefore, we can conclude by saying that weekdays have a larger number of trips in comparison to the weekends. This may be due to the fact that on weekends, people may be preoccupied with other activities such as spending time with family. The higher number of trips on the weekdays may be attributed to the busy nature of people and thus they are in need of bikes so that they ride to workplaces or other events that they may be having in the course of the week (Chen et al, 2015). The total trips per day seems to follow a uniform distribution. This is because the number of trips is almost equal in all the days of the week.

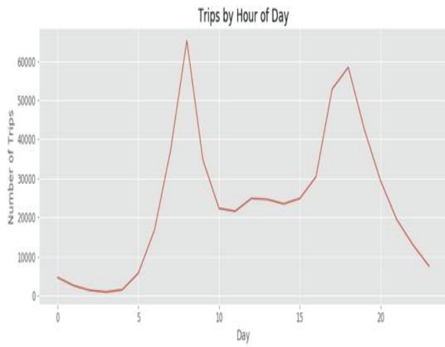

Figure 2: Trips by hour of Day.

The graph above indicates that between hour 0-5, there is a lowest number of trips. From hour 5-8, the number of trips increases significantly but stats declining from hour 8- 16. Hence, between 9.00am and 10.00am and between 4:00pm and 7:00pm there is higher trips per hour. Thus, the number of trips in the afternoon hours is low i.e. 10:00am to 3: 00pm. We can therefore, conclude that between 9:00am and 10:00am most people are going to work and thus the demand for bikes is high at that particular time while between 4:00pm and 7:00pm people are through with their daily work and thus they need to get back to their homes and families and this translates to a higher demand for bikes. In the afternoon hours, the demand for bikes is low and this may be due to the fact most of the riders are at work or are busy somewhere else and this translates to low bike demand.

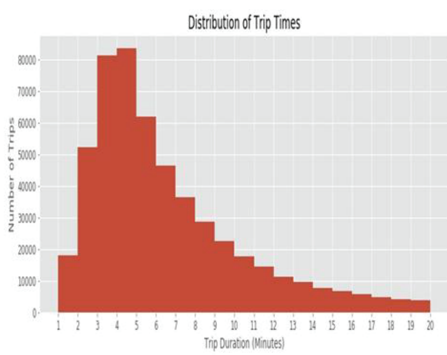

Figure 3: Trip Time by Minutes

The number of trips is higher between the 3rd and 5th minute and lowest between the 18th and 20th minute. From the 5th minute, the number of trips declines significantly up to the last minute i.e. 20th. Trip times per minute is exponentially distributed. Therefore, it is clear that the number of trips declines with the increase in trip duration. In other words, there is an inverse relationship between the number of trips and the trip duration. This can be attributed to the fact that at lower trip duration, people still have the energy and momentum to participate in bike riding but as the number of trips increases, riders become more and more tired and thus the number of trips declines with the increase in trip duration. Also, in the morning hours, many people would prefer bike riding rather than using public or private transport. This is because at the morning hours, there may be a presence of traffic congestion and this may hinder people from reaching at workplace on time. This would in turn make them prefer to use bikes so that they may escape traffic congestion (Zhang, Duan, & Bryde, 2015).

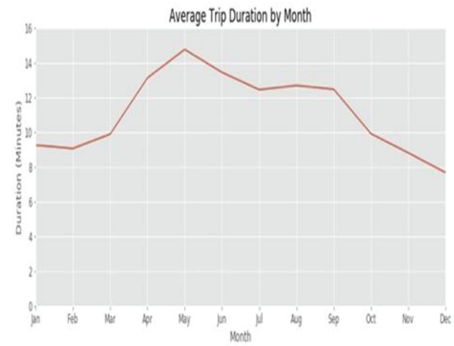

Figure 4: Average trip duration by month

The graph above shows that average trip duration is higher between March and October. The average trip duration is lower between January to March and October to December. From this, we can say that the winter months have lower trip duration as compared to other months. This is because during winter, people tend to use other means of transport in commuting since the temperatures outside do not favor bike riding. Thus, people prefer bike riding activities during warm days and months. Thus, the rate of turnover is very high in warm months as compared to colder months. This is quite normal because technically one would refer to using a means of transport that is comfortable and warm during the cold months. Conversely, during the hot months i.e., summer, most people would prefer bike riding because it would seem more comfortable since one needs a cool breeze at that particular time.

C. *Forcast Modelling*

It is important to make future predictions based on the given data through the analysis of trend. This will help the researcher to know how the number of trips will be related with time and therefore valid conclusion may be drawn from the including custom seasonality and holidays. This makes it easy for those people without forecasting expertise use this to make meaningful predictions for various problems in business scenarios (Chiru & Posea, 2018). In our case, we focused on the number of trips on a daily basis. Forecasting was done for 180 days i.e. 3 months. The model has three main components i.e. trends, seasonality and holidays. The following equation gives their combination. y(t)=g(t)+s(t)+h(t)+ɛ́_t Where, g(t) is the piecewise linear or logistic growth curve for non-periodic changes modelling in time series. S(t) is periodic changes (in week, years or seasonality) H(t) irregular schedule holiday effects ɛ́t is the error term FBProphet tries to fit several linear and non-linear time functions as components. Adding seasonality in the model is the similar approach that is used in exponential smoothing. In our case, we were able to train the model to forecast 6 months from October, 2018 to April, 2019 and so the model is working properly since we already removed the noise factors by fitting it on monthly basis (Chiru & Posea, 2018). The model is shown below:

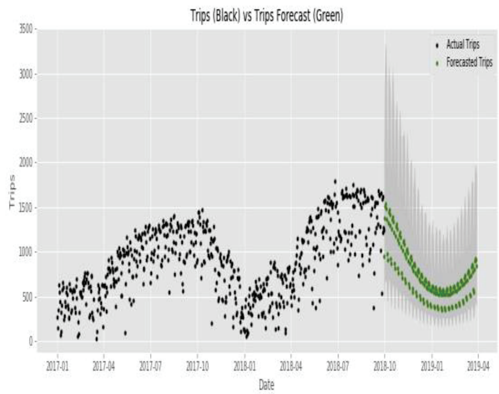

Figure 5: Forecast model by Day.

From the graph above, the trips data exhibits seasonality and trends. For instance, the number of trips is higher between April and January each year i.e. 2017-4 to 2018-1 and also 2018-4 to 2019- 1. Also, the number of trips are lower between January and April each year. Therefore, we can forecast the trips number in any month in future. For instance, if we are interested to know the trips number in June, 2030, we could use the above model and conclude that in June, the number of trips would be at peak/ higher. Thus, through the use of this model, the bike firm may make sure that during the peak season, resources are provided to the riders in order to counter the excess demand. Seasonality in this data would help a bike station to develop business objectives and goals. Using the given historical data, the bike station firm may predict the future through the use of data analysis and time series models. When a data exhibits a broader time period, time series are the best models that may be used to in the forecasting (Chiru & Posea, 2018). When we take a weekday (Monday) compared to the weekend day Sunday), the weekdays have a higher trip as shown in the figure below.

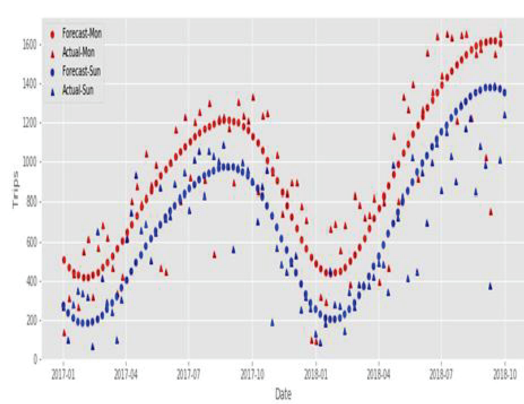

Figure 6: Forecasting Model Performance Weekday/ Weekend Comparison

The figure indicates that weekdays have a higher number of trips as compared to weekends. The model above shows seasonality and trend. For instance, the number of trips increases with the weather changes. Therefore, when the weather is cold i.e. December and January, the number of trips declines significantly. From this, we can depict that the number of trips is affected by weather/seasonality. Thus, riders prefer to engage in bike riding activities on warm weather months rather than cold weather months. This seasonality and trends should be taken an advantage of by the bike firm by considering the macro weather and realize that in the mid-year, the level of sales are very high and thus they should take the advantage of the situation and maximize the returns (Chiru & Posea, 2018). Also, they should be able to adjust bids or modify budgets so that they may counter the competition during the peak seasons. They should also make sure that they stay focus and relevant by being able to plan ahead to enable creation of new ad copy. 9 Hence, the firm should realize that the seasonal menus increase the overall sales/returns and this means that more money is being invested in the firm and not at the firm’s competitors. they should ensure that the bikes are available daily on 24/7 so that one can access transport each and every time they are in need. The management of the bikes need also to specify whether the bikes should be paid annually or time based fee. This will enable riders to be flexible and adjust the user pricing model as the system matures. Seasonality and trend also increases the business bottom-line, allows staff to show creativity, keep the business on trend and fresh, makes sure that menu favorites are kept, and most importantly the firm should try offering new items (Chiru & Posea, 2018).

D. Geo-Spatial Analysis:

Geospatial analysis may be defined as the collection, presentation and manipulation of GPS, imagery, past data, and satellite photography which is described clearly in terms of geographic coordinates or indirectly, in terms of street address, postal code or forest stand identifiers as applied in geographic models. In Canada, Geospatial analysis was used alongside geographic information system (GIS). Geographical information systems are used in prediction, management and learning of the phenomena that affect the earth, its system and inhabitants. Geospatial analysis may find its application in climate change modelling, management of crisis, sales analysis, weather monitoring, forecasting of human population and management of animal population. We used Maptitude to analysis the data by showing in a map for visualization about the area where the stations are located, and we selected an area based on radius where we felt that needs of a new station. The area selected was suitable because its favorite climate. As we have seen earlier, the number of trips is dependent on the weather i.e. in cold months, the number of trip is lower while in warm/hot months, there is a higher number of trips. Thus, this area was considered the most appropriate because most of the months are warm and thus this will attract a large number of bike riding activities. The area has got also a scenic view and thus people may enjoy the beautiful view while they are on their adventure. Also, we noted the area has a very good infrastructure in terms of the roads and this would make the use of bikes in commuting much easier (El-Assi, Mahmoud & Habib, 2017). The area is also recommendable due to its population. Setting up a bike station requires that the population of that given area is large enough, so that each time the bikes are on the road and this would help boosting the firm's profits. Therefore, demographic factors are very important in setting up such station. Also, the age group was considered because the business was targeting young people. Young people would likely to use a bike more than the old population. also, gender is also an important factor that was put in consideration. For instance, male population is likely to use a bicycle frequently when commuting as compared to their female counterparts. We recommend this are because of it is 0.25miles radius from other station and this means the level of completion for the services will be quite low and thus this would enable the business to thrive. The area also has a good neighborhood and the 10 level of security is high and thus this puts confidence on investors of the project. Therefore, this is the perfect location for the station due to its demographic and environmental factors. Setting up this station on the area will enable the business to thrive and this will help reduce road congestion, air pollution, economic competitiveness, safety risks, sustainable growth and social cohesion.

E. Geograpgic Info(Emersion Radius)

The proposed site for the station has a total population of 15,550 and the number of households is 5,266. Thus the average number of people per household is approximately 3. The average household income is given by $64,902. Thus, from this figure, we can deduce that the living standards of this population is quite high. We can understand that households from Hamilton Park earn an income of $131,586 spend 35,709 while those from exchange place earn $189,189 and spend 33,985. The two groups are the highest earners and spenders in Jersey City. The households from Sip Ave, Bayside Park, Jackson Square, Bethune center earn an income of $60,809, $48633, $47,560, 56,334 respectively and spend 29414, 216, 203,85 respectively on the number of trips taken. Thus, in Jersey City, the high income earners are the big spenders as shown from the graph above.

F. Route Optimization and Recommendation

Due to the rapid population growth in Jersey City, there is a need to put up another station so that the general population of jersey does not strain on the available resources. Thus, putting up a 7th station in Jersey City will be of much help to the residents of Emerson Ave as well as other people who may be interested to use the resources for the proposed station. The growing population is in need of transport services and thus if each person bought a vehicle, that would lead to traffic congestion in Jersey City. Therefore, this project may as well help reduce traffic congestion in major public roads. Also Emerson Ave & 7 Sip Ave (Intersection) is an ideal location for a new station in Jersey City based on the area analysis. There is no station within .50 miles of this location and it has a major connecting station is the Sip Ave Station and this makes the proposed station a stopover. This translates to a higher rate of turnover. In order to facilitate transport services in the Emersion station, there is a bus stop at this intersection which makes it a good placement option. Thus, people coming from longer distances will commute easily and at a lower cost. Additionally, setting up a new station in Jersey City will supplement the public transport system and will further reduce traffic congestion in the main public roads since most people will prefer commuting with a bike (Zhang et al, 2015). Below is a table showing the different stations and the number of trips.

Top 5	
Station	Trips
Grove St PATH	66,287
Hamilton Park	35,709
Exchange Place	33,985
Sip Ave	29,414
Newport PATH	24,620

Table 4: Top 5 Start Station usage

Bottom 5	
Station	Trips
JCBS Depot	41
Bethune Center	85
Jackson Square	203
Bayside Park	216
Danforth Light Rail	257

Table 5: Bottom 5 Start Station usage

The table above indicates that Grove St. Path has the highest number of trips i.e. 66, 287 followed by Hamilton park which is at 35,709. Exchange place has a 33,987 trips while Sip Ave and Newport Path have 29,414, 24,620 trips respectively. JCBS Depot and Bethune Center have 41 and 85 trips respectively. Jackson Square, Bayside park and Danforth Light Rail have 203, 216 and 257 trips respectively. therefore, we can depict from the table above that Grove St. PATH has the highest number of trips while JCBS Depot has the lowest number of trips. The low number of trips in this station could be due reasons such as low population, low income by the households around those areas, lack of proper marketing and advertising, inadequate number of bikes thereby losing trust and reliability by customers, the stations may also be located in places where the riders do not see them. All this factors could contribute to low number of trips recorded by those stations. We have also optimized the top performing unique routes in order to increase the efficiency of the system.

Route #	Start Station	End Station	Trips
Route 1	Hamilton Park	Grove St PATH	23,416
Route 2	Morris Canal	Exchange Place	15,958
Route 3	Brunswick St	Grove St PATH	6,778
Route 4	Jersey & 6th St	Grove St PATH	6,428
Route 5	Brunswick & 6th	Grove St PATH	6,061

Table 6: Top 5 Unique Routes

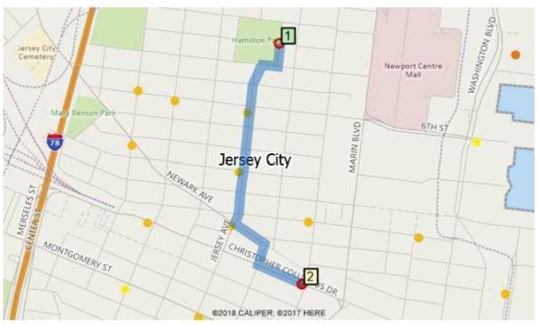

Figure 7: Route 1 (Optimized)

Figure 8: Route 2 (Optimized)

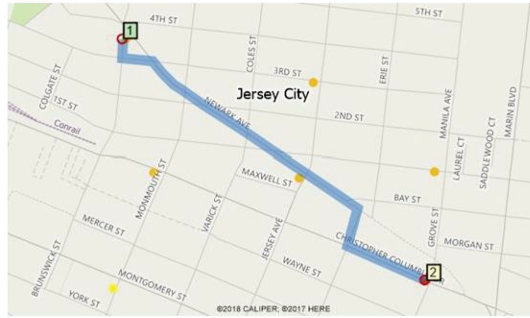

Figure 11: Route 5 (Optimized)

CitiBike deploys a rebalancing system where their employees manually remove and move bikes around based on activity and usage. We have optimized the route to follow while rebalancing based on the Top 13 Stations used. All the route optimizations are done based on the shortest distance between two specific stations.

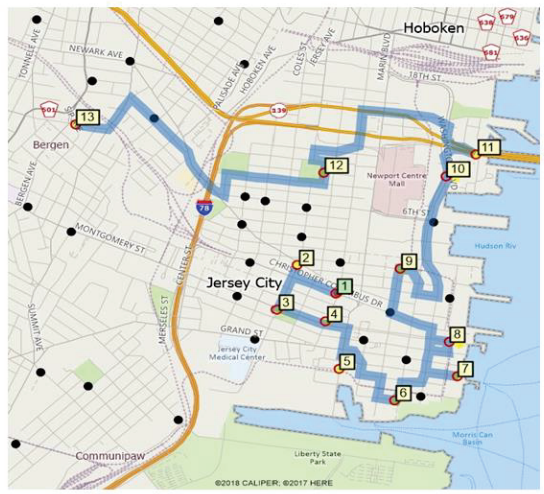

Figure 12: Rebalance Route (Optimized)

G. Predicted Hot-spot

Below is the predicted Hot-Spots for the stations based on the October, 2018 - April, 2019 predictions. The below table and map shows the Top 10 Predicted Stations along with the predicted Hot-Spots based on station use. A specific area/radius of the CitiBike sharing system in Jersey City makes up the most in terms of station use.

Predicted Top 10	
Station	**Trips**
Grove St PATH	18,231
Hamilton Park	9,157
Newport PATH	7,500
Exchange Place	6,620
Marin Light Rail	6,215
Harborside	5,405
Washington St	4,489
Newport Pkwy	4,374
Brunswick & 6^{th}	4,013
Pershing Field	3,826

Table 9: Top 10 Predicted Station Usage

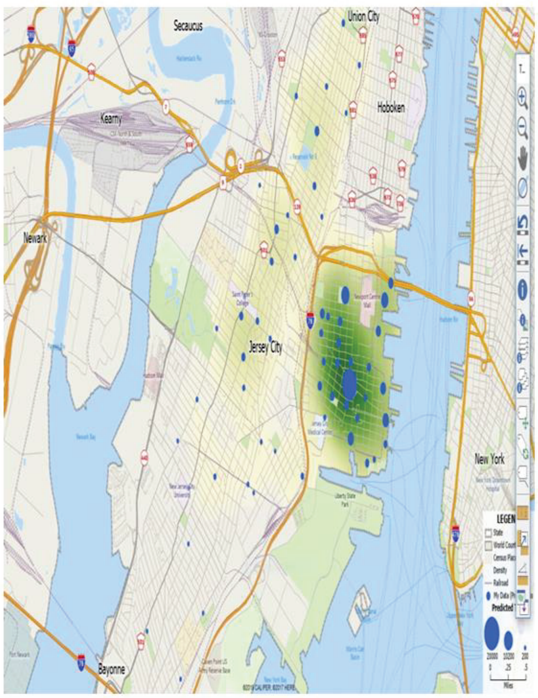

Figure 8: Predicted Hot-Spots for the Stations

IV. Discussions, Conclusion and Recommendations

An analysis of the CitiBike Trip Data has revealed that in certain months, the number of trips are high while in other months the number of trips is low. It is clear that this behavior is due to seasonality and trend. For instance, in cold months, the number of trips reduce significantly as compared to other months which are warm. Therefore, in setting the new station in Jersey town, one should consider seasonality and take advantage of the situation. For instance, in the peak season, one should make sure that a greater rider experience is designed so that riders may be attracted from different parts of the City. Since Jersey City Citi- Bike sharing system is an ongoing project and improvements are constantly being made. We can say that by analyzing these are in need. The management of the bikes need also to specify whether the bikes should be paid annually or time based fee. This will enable riders to be flexible and adjust the user pricing model as the system matures. In the same way, they need to specify the kind of bikes they are offering so that the riders may know if they are offering mixed bike types. The new bike station should also determine the type of bike sharing system that would work best on the basis of rider type and experience. Thus they need to make a decision on the different types bike share program (El-Assi, Mahmoud & Habib, 2017).

Since the best location for the bike station has been identified and studied, the project is good to go from there. Location is very critical in every business. This is because location enhances visibility. The station should be conveniently located to enable visibility by riders or located near to roads for biking. The management should also make sure that the stations bikes are distinguished from other personal bikes. It is also the duty of the management/shareholders to calculate the amount of bikes needed in the area although this might be a very hard task. This is because too few bikes would have riders lose trust in the system reliability. On the other hand, too much bikes would leave many unused bikes sitting alone in the station. When considering the amount of bicycles, the station need, they have to consider factors such as total population of the area, percentage population thot would use bicycles on a regular basis, the average number of trips taken per week by a regular rider, and the average trip time taken by a regular rider.

Finally, the station needs to create a strategy for the launch, marketing and recreational activities. This is because the most crucial moment of bike share system is the launch. Strong user adoption will create confidence on the riders and this would eventually lead to the success of the station. Therefore, the management/shareholders should not underestimate the launching and running a new station complexity. Thus, on planning the bike share program, the management should be willing and able to pick the right equipment and partners for successful launching and the long run growth and success of the new station (El-Assi et al, 2017). we came up with a recommended station in the Emerson Ave and Sip Ave (Intersection) is an ideal location for a new station in Jersey City based on the area analysis. There is no station within 0.50 miles of this Location and it has a major connecting station in the Sip Ave Station. There is a bus stop at this intersection which makes it good placement option. When we went in-depth about the location we found that the Population is 15,550 in which 5,266 Households with an Avg. Household Income is $64,902. It is a good location to start a new station because the people in this area need it and they can even affordable.

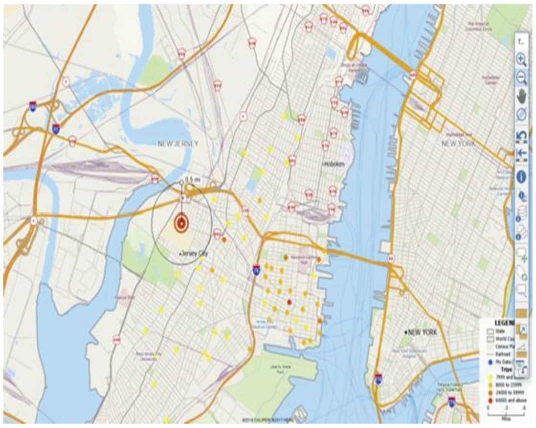

Figure 9: New Station Recommended

In conclusion, setting up this program will help in many ways. For instance, the program will help reduce road congestion, air pollution, economic competitiveness, safety risks, sustainable growth and social cohesion. Therefore, to maintain a healthy city, we must adjust to this program. However, there are certain measures that may help address this problem. One of the measures is the prevention of the conventional car use or rather providing an alternative. cycling has been viewed as the simplest and realistic solution to the problem in cases of shorter trips. Thus, bike sharing combines both innovations and comes out with the best solution. Therefore, bike sharing program helps in transport flexibility, health benefits, reduction of air pollutions, fuel consumption reduction, reduction in congestion, and individual's financial savings. But the most interesting idea in this program is the quality of sharing with others. This is because there is no cost and responsibilities associated with the ownership.

REFERENCES

[1] Aviv, T. (2011). Static Repositioning in a Bike-Sharing System: Models and Solution Approaches. Tel Aviv University, 1-40. Retrieved from https://link.springer.com/article/10. 1007/s13676-012-0017-6

[2] Chang, X. (2018). Statistical patterns of human mobility in emerging Bicycle Sharing Systems. PLOS One. doi:10.1371/journal.pone.0193795

[3] CiteBike. (2018, 2018). Experience NYC in a whole new way. Retrieved October 10, 2018, from https://www.citibikenyc.com/how-it-works

[4] Eluru, A. F.-I. (2016). Incorporating the Impact of Spatio-Temporal Interactions on Bicycle Sharing System Demand: A Case Study of New York CitiBike System. Journal of Transport Geography, 54(2016), 218-227. doi:org/10.1016/j.jtrangeo.2016.06. 008

[5] FRI, F. R. (2017). Bike Sharing in New York ci ty: How the Citi bike System Serves Points of interest. University of Twente. Retrieved from https://essay.utwente.nl/73652/1/Fr itz_BA_TNW.pdf

[6] Levenson, L. G.-K. (2014). Citi Bike Takes New York. Rudin Center for Transportation Management and Policy, 1-31. Retrieved from https://wagner.nyu.edu/rudincenter

[7] Rixey, R. A. (2013). Station-Level Forecasting of Bike Sharing Ridership: Station Network Effects in Three U.S. Systems. TRB 2013 16 Annual Meeting, 1-15. Retrieved from http://journals.sagepub.com/doi/10. 3141/2387-06

[8] Shaheen, S. A. (2012). Public Bikesharing in North America: Early Operator and User Understanding, MTI Report 11-19. Mineta Transportation Institute Publications. Retrieved from https://scholarworks.sjsu.edu/mti_p ublications/73/

[9] Shmoys, E. O. (2015). Data Analysis and Optimization for (Citi)Bike Sharing. Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, 678-694. Retrieved from http://mobility-workspace.eu/wp-content/uploads/Analysis_CitiBike Sharing.pdf

[10] Singhvi, D. (2015). Predicting Bike Usage for New York City's Bike Sharing System. Association for the Advancement of Artificial, 111-114. Retrieved from https://aaai.org/ocs/index.php/WS/ AAAIW15/paper/download/10115/ 10185

[11] Zhang, Y., Thomas, T., Brussel, M. J., & M. F. A. M. van Maarseveen. (2016). Expanding Bicycle-Sharing Systems: Lessons Learnt from an Analysis of Usage. Retrieved from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0168 604

[12] Chen, L., Zhang, D., Pan, G., Ma, X., Yang, D., Kushlev, K., ... & Li, S. (2015, September). Bike sharing station placement leveraging heterogeneous urban open data. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 571-575). ACM.

[13] El-Assi, W., Mahmoud, M. S., & Habib, K. N. (2017). Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. Transportation, 44(3), 589-613.

[14] Zhang, L., Zhang, J., Duan, Z. Y., & Bryde, D. (2015). Sustainable bike-sharing systems: characteristics and commonalities across cases in urban China. Journal of Cleaner Production, 97, 124-133.

[15] Chiru, C. G., & Posea, V. V. (2018, September). Time Series Analysis for Sales Prediction. In International Conference on Artificial Intelligence: Methodology, Systems, and Applications (pp. 163-172). Springer, Cham.