The Implications of Citi Bike on Motor Vehicle-Bicycle Collisioins in New York City

Shi Fan
Center for Urban Science and Progress
New York University
New York, NY, United States
sf2632@nyu.edu

***Abstract*—In this paper, I attempt to analyze the effects that Citi Bike has had on motor vehicle-bicycle collisions in New York City. I approached the analysis both temporally and spatially, attempting to identify the change in the collision patterns before and after Citi Bike. As a result, Citi Bike does have a certain level of impact on the cyclist collisions in New York City.**

Keywords—motor vehicle collision; bicycle safety; public bicycle sharing system; urban transportation; urban spatial analytics

I. Introduction

OneNYC is one of the major movements in crowdsourcing potential solutions to civic and planning challenges in New York City. Within the four visions that OneNYC seeks to pursue, equity plays an important role in establishing a just and inclusive citywide economy. Security is one of the ultimate goals that the equity vision projects to achieve, and as a multi-nation movement in enhancing road traffic safety, Vision Zero has been one of the main players in the security piece of OneNYC since its inception in New York City in January, 2014.

Road traffic collisions constitute a major urban problem that affects a lot of big cities around the world. There are three primary types of motor vehicle collisions: vehicle with vehicle, vehicle with bicyclists, and vehicle with pedestrians. Historically in the United States, bicyclist fatalities in motor vehicle collisions have decreased drastically from 1,003 deaths in 1975 to 621 in 2010; however, from 2010 to 2012 the number has increased 16% to 722, while other types of motor vehicle fatalities almost remained unchanged [1].

The resurging trend is corroborated by a detailed report on bicyclist safety, which is prepared for Governors Highway Safety Associatioin (GHSA), in 2014. In an effort to unravel some of the baseline conditions for such an undesirable phenomenon, the report, written by Dr. Allan Williams, is able to shed some light on the key factors of crash patterns behind such a drastic increase. Demographically, the bicyclist fatalities involving adults age 20 or older have jumped from 21 percent in 1975 to 84 percent in 2012 [2]. Within the same interval, the bicyclist fatalities involving males have increased from 82% to 88 percent [2]. The demographic crash pattern reveals that adult males have become increasingly vulnerable to deadly incidents in motor vehicle-bicyclist collisioins. The main drivers behind this shifted demographic pattern of fatalities, according to the report, are helmet use and alcohol use. "Lack of helmet use is a major contributing factor in fatalities," as more than two thirds of fatally injured bicyslists in 2012 did not have helmets on [2]. On the other hand, starting from recent decades, "twenty-one states and the District of Colombia have helmet use laws for children," which partially accounts for the sharp decline in the percentage of teenagers involved in fatal incidents. The report also points out that "more than 1 in 4 adult bicyclists killed in 2012 were alcohol-impaired," which ironically downplays the healthy lifestyles and environmental benefits people typically associate biking with [2]. Geospatially, the proportion of incidents happened in urban areas have jumped from 50 percent to 69 percent from 1975 to 2012, and 37 percent of the incidents occurred at intersections in 2012 [2]. The implication that such statistics suggest is a certain degree of change in the environment setting of motor vehicle-bicyclist collisions over time. Particularly in New York State, the number of bicyclist death have increased by 25 percent from 36 in 2010 to 45 in 2012 [2]. It is tempting to relate such sharp increase to the expansion of urban areas or the increasing complexity in road infrastructure or transportation modes. In an effort to combat the issue, NYC Department of Transportation has established "slow zones" in residential areas across various neighborhoods [2].

Nevertheless, it is reasonable to assume that the increase in the number of incidents implies that there is an increase in the total number of biking trips in New York. Speaking of recent innovations in NYC transportation, Citi Bike has become quite a buzzword as it is the "official" public bike sharing system in the city. Since the inception of the system in July, 2013, more and more people have been using Citi Bike in lieu of other traditional transportation means, such as subway, buses and taxi. Meanwhile, it may have also increased the total number of bicycle trips in NYC. As Citi Bike is constantly expanding its stations across the city, it keeps track of all the trips from each subscriber and customer, which makes it convenient for civic hackers to dive into the data and analyze the mobility patterns that it triggers to the city. Besides, along with the spirit of the open data movement, NYPD has been sharing its motor vehicle collisions dataset with the public since July, 2012 on NYC Open Data. The combination of the two datasets pushes me to think of the following question: has Citi Bike had an impact on the motor vehicle-bicycle collision patterns in NYC? This paper attempts to answer the question by comparing the temporal and spatial patterns of motor vehicle-bicycle collisions before and after the initiation of the Citi Bike system. In addition, it attempts to identify certain correlation, temporal and spatial, between Citi Bike trips and cyclist collisions with motor vehicles.

II. Literature review

A. New York City Baselines

In order to get a realistic handle on the issue, it is worthwhile looking into some previous works related to New York City that have been done. As mentioned earlier, lack of helmet use is a major culprit for the severity level of injuries in motor vehicle-cyclist collisions. A paper published in the Journal of Community Health in February, 2014 presents a detailed and elaborate analysis on the helmet usage pattern among cyclists in NYC. Researchers recorded two hours of video at five different locations in Manhattan and "coded the observations on slow speed to enumerate the total number of cyclists, gender, helmet use and type of cycle" [3]. Notably, type of cycle is broken down into three categories: Citi Bike, other rental bike, and non-rental. The overall result turned out to be that about 50 percent of cyclists observed were wearing helmets [3]. Demographically, male bikers are more likely to wear helmets than female bikers. By the type of users, recreational bikers are less likely to wear helmets than commuters [3]. By the type of bicycles, rental bikes were found much less likely than non-rental ones to be riden by bikers with helmets on [3]. "Specifically, only 21.7 percent of Citi Bike users were found wearing helmets," while the number is even lower (15.3 percent) for users of other rentals [3]. The study is able to shed some light on the helmet usage pattern among Citi Bike users, being that they tend to wear helmets much less often than other bikers. The implication of such findings suggests that Citi Bike users can be potentially more susceptible to severe injuries than other users. However, such findings were summarized from a two-hour observation period, which may create a certain degree of bias in the conclusions they draw from the issue.

Another study, named "evaluating the safety effects of bicycle lanes in New York City," offers some insights on how the increase of bicycle lanes have impacted the safety environment of transportation in the city [4]. The major findings from the analysis are that "the installation of bicycle lanes does not lead to an increase in the number of bicyclists after the addition of such lanes," and that "crashes at intersections appeared to increase, although not significantly" [4]. The latter argument echoes the increase of accidents at intersections which was pointed out in the Bicyclist Safety report. The causational factor in generating such a seemingly unusual phenomenon can be that there is a lack of safety treatment at the intersections corresponding to such additions of bicycle lanes. The study proposed more detailed treatment for each lane added, and planners and engineers should be particularly careful when dealing with intersections [4].

Feeding from creating corresponding facilities to accommodate the additions of bicycle lanes, a study done by the same group of researchers illustrates the types of effective safety countermeasures in reducing the number of crashes in New York City [5]. The research selected the 13 types of important countermeasures and evaluated their effectiveness based on real-time data. The key finding is that it is the combination of reducing the number of travel lanes and installing bike lanes, rather than simply installing bike lanes, that is highly effective in reducing the number of crashes. It correlates to the assumption of the previous study that detailed road treatment should come with bike lanes together as a package, especially in New York City. Ultimately, well-designed bike lines are able to bring significant amount of public health benefits by reducing motor vehicle-bicycle crashes in the city.

B. Built Environment

After an overview of the baseline conditions in New York City, it is necessary to look into certain contextual frameworks that are used to illustrate the generally complex, case-by-case scenarios in the issue of motor vehicle-bicycle collisions. A study done by two civil engineering researchers at the University of Utah attempts to explore the causation between multimodal transportation infrastructure and traffic safety in an urban context [6]. There has been a growing popularity of multimodal transportation infrastructure across the majors cities in the US—New York City is no different. The researchers used Chicago as a case study to measure the "expected crash frequencies of different types and severities involving mortorized and non-motorized users," reaching a conclusion that areas of higher density tend to have a higher crash rate due to an influx of multimodal trips, such as transit and biking facilities [6]. If such relationship can be drawn from the situation in Chicago, a metropolitan with a variety of transportation means just as New York, it is possible that the higher density areas in NYC, equipped with even more complex transportation infrastructure, may also suffer in terms of vehicle collisions. Therefore, additional safety countermeasures should be provided in order to offset the potential risks that come along with the benefits of multimodal transportation. Particularly in the case of Citi Bike, the increasing numbers of stations should be carefully planned and structured by taking safty countermeasures into consideration, and the trips should also be studied to reveal some insights on biking density across the city.

Unsurprisingly, built environment plays an important role in cyclist injury severity during crashes. Zoning should indeed "consider the safety effects of diversifying land use" in terms of both collision rate and severity [7]. While nowadays the discussion of bicycle safety tends to be often accociated with the urban built environment that has a mixture of land use, it is also important to consider how it is affected by the built envirnoment in those less-urban areas which have less convenient access to multimodal transportation. Chia-Yuan Yu, an urban planning researcher from Texas A&M University, reveals some insights on the bicycle safety in less-urban areas, claiming that "people living in poor areas are more likely to bike to work but their environments do not provide enough safety for them" [8]. Applying this argument to the motor vechile-bicycle collision scenario in New York City, I expect to find certain remote areas, relatively far away from those Central Business Districts, that have a high frequency of cyclists accidents due to insufficient biking infrastructure. Those areas may not usually equipped with Citi Bike stations but definitely cannot be ignored.

Nevertheless, it is important to consider how people actually perceive the safety level in the routes they take, and how they leverage the subjective safety measurement with the efficiency-based path finding. This is usually an issue in urban areas where people are faced with multiple routes and most concerned about efficient routing. Another researcher from Texas A&M, Shailesh Chandra, applied the routing model that takes both safety and efficiency into account to the street network system in the City of College Station, Texas. The main idea is to prioritize safety and suggest the fasted route among those that are marked "safe," based on certain metrics and thresholds developed by the model [9]. However, applying such a model to New York City will be particularly challenging, because the city has an relatively complex road network. It will be crucial to mitigate the biases between safety and efficiency and make scalable suggestions in real-time. Such insights can be valuable for Citi Bike to inform their bikers on the real-time traffic situation in terms of both safety and efficiency.

C. *Research Objective*

The overarching goal of the research is to inform actionable insights on the operations of Citi Bike and city agencies in charge of transportation planning (i.e. DoT) by providing an analysis based on the motor vehicle-bicycle collisions and Citi Bike trips. The first step is to identify whether or not Citi Bike has had an impact, temporally and spatially, on the patterns of motor vehicle-bicycle collisions in New York City. Then it is critical to assess such an impact by relating it to the specific contexts involved, ultimately drawing potential conclusions from those case-by-case scenarios.

III. Data

A. *Sources*

1) *NYPD Motor Vehicle Collisions Data*

This dataset prepared by New York Police Department contains detailed information on all the real-time motor vehicle collisions in New York City. Temporal components of the data include exact date and time. Spatial component of the data include specific location denoted by latitude and longitude, zip code, borough, on street name and cross street name. Description of each incident include number of persons injured and killed, number of pedestrians injured and killed, number of cyclists injured and killed, number of motorists injured and killed, contributing factors of each vehicle involved in the collision, and type of each vehicles involved in the collision.

In this project, I primarily focus on analyzing the incidents involving cyclists injured and killed based on both temporal and spatial variables.

2) *Citi Bike Trip Data*

This dataset prepared by Citi Bike system contains detailed information regarding each Citi Bike trip by both customers and subscribers. Temporal components of the data include exact date and time, for both start and end of each trip, and trip duration. Spatial components of the data include station ID, specific locations denoted by latitude and longitude, and station names denoted by on street and cross street, for both start and end of each trip. More detailed description of each trip include bike ID, user type, gender and birth year.

In this project, I primarily focus on analyzing the number of trips in each month, the number of trips in each hour, and the most popular stations in terms of the volume of trips for both start and end stations.

B. *Units*

- Since the point is to compare the number of collisions and the number of Citi Bike trips, there is no such scientific units of measure specifically applied here.
- I created a variable, however, for motor vehicle collisions, named the weighted number of cyclist accidents. The purpose of establishing such a variable is to get a reasonable measurement on the severity of accidents. Details are explained in the Equations section.

C. *Equations*

As mentioned, the new variable is called the weighted number of cyclist accidents, which is denoted by "w" as in (1). "w" is calculated by adding the number of injuries to ten times the number of deaths, respectively denoted by "i" and "k" as in (1). The purpose of introducing this variable is to leverage the weights that each typical injury and fatal incident has on the overall severity measurement of vehicle collisions.

$$w = \iota + 10 * \kappa \tag{1}$$

I applied the weighted number of cyclist accidents to the aggregated number of cyclist accidents over time, so that an interval that has more cyclist deaths tends to be more severe than its result calculated by simply aggregating the number of injury accidents over time.

IV. Methodology

As mentioned earlier, I intend to approach the analysis temporally and spatially, because the two datasets both have temporal components as well as spatial components. The main idea is to aggregate the number of motor vehicle-bicycle collisions and the number of Citi Bike trips by a few selected metrics, in order to identify whether or not there exists such a correlation between the two variables. More importantly, it is tempting to consider the causation in addition to the correlatioin, because our research question is on the causal effects that Citi Bike trips may have on motor vehicle-bicycle collisions.

A. *Temporal Analysis*

I based my temporal analysis on three different temporal metrics. The first metric that I used was every six months. This metric was selected due to a few reasons. The first reason is that I wanted to get a sense of the general trend in cyclist collisions, so it is appropriate to aggregated by a relatively long period. A six-month period provides a handle that lies somewhere in between the yearly trend and the seasonal trend. More importantly, some of the key events that happened along the way made it convenient to analyze the number of collisions based on a six-month period. The motor vehicle collision data has been recorded since July, 2012. Citi Bike public bike sharing system started to be implement in July, 2013. Using the 6-month interval makes it possible to perceive how the patterns in the number of accidents correspond to these series of events, particularly before and after the usage of Citi Bike. Basically I aggregated the number of collisions involving cyclist injuries over each 6 months from July, 2012 to July, 2015.

The other two metrics are month and hour. The justification of choosing them is quite straightforward, because we are interested in discovering trends and correlations within a year and a day. The way of handling the data is similar, just aggregating the number of collisions involving cyclist injuries over each month and each hour.

1) Collisions by Every Six Months

The purpose of analyzing the number of collisions by every six months is to get a rough idea on the overall image of the issue. It should be noted that I used the weighted number of collisions introduced in the Data—Equations section early on, for the purpose of taking the severity level of each accident into account. Some interesting patterns are illustrated by Figure 1 in section C. Figures and Tables. First, for each year, the second half (July to December) tend to have an obviously higher weighted number of bicycle accidents than the first half (January to June). Second, for the second half years the weighted number of accidents have been steadily increasing over the past three years, from about 2,200 in 2012 to about 2,600 in 2015, while for the first half years the trend seems to be exactly the opposite but less severe, from about 1,800 in 2013 to about 1,600 in 2015. As a result, the gaps between the number of accidents in the first half and that in the second half of each year have been constantly increasing over the past three years. Considering the fact that the number of accidents in the second halves are increasing at a faster speed than those in the first halves are decreasing, it is okay to conclude that overall number of accidents is increasing each year.

2) Collisions by Month (Before and After Citi Bike)

A look into the number of collisions by month reaffirms my previous finding on the seasonal patterns of bicycle collisions. I used a categorical model to predict the weighted number of bicycle accidents over months, as shown in the top portion of Figure 2. The model considers month as a categorical variable, so that each month is assigned a particular coefficient corresponding to its performance on the number of bicycle collisions in the fitted model. As the plot suggests, at the beginning of the time series, our model tends to over-predict the number of collisions within those collision-heavy months, while it tends to under-predict the number of collisions within those collision-light months. However, at the end of the time series, our model tends to do something on the contrary: under-predicting in collision-heavy months and over-predicting collision-light months. The fitted model by month reassures *our* previous finding based on the general temporal param*eter*. number of accidents is increasing each year.

In addition, I fitted a linear regression model on the number of bicycle accidents in each month, separately for before and after the Citi Bike system started to be implemented. The result turned out to be that before Citi Bike, the general trend in the number of monthly bicycle accidents was downward, and after Citi Bike, it has become upward. It may be tempted to think that Citi Bike has generated more bicycle trips in New York City, which leads to an increase in the number of accidents.

3) Temporal Correlation

Having seen an increase in the number of bicycle accidents after the implementation of the major public bike sharing system in the city, I wanted to compare the number of Citi Bike trips with the number of bicycle accidents on two different temporal scales—month and hour. I aggregated the number of accidents and the number of Citi Bike trips by both month and hour. I expect to find certain correlation between the two variables, and the result shown in Figure 3 confirms my hypothesis. From the trend lines, we can tell that the weighted number of bicycle accidents is indeed highly correlated to the number of Citi Bike trips in each month, with a value of 0.93. Such correlation also exists yet is less intense in the hour case, which has a value of 0.83.

B. Spatial Analysis

The general steps in approaching the spatial analysis is similar to what was done in approaching the temporal analysis. I started with comparing the spatial patterns in cyclist collisions before the use of Citi Bike with those after. Then I tried to figure if such a change in spatial patterns was due to Citi Bike, more from a causation standpoint rather than correlation. I looked into the Citi Bike trip data to find out the most popular stations and analyzed how they could have triggered the change in the spatial patterns of cyclist collisions. Ultimately, such causation factors result in a spatial correlation between cyclist collisions and Citi Bike trips.

1) Collisions by Zip Code (Before and After Citi Bike)

Similarly for the spatial analysis, I tried to compare and contrast the patterns in the number of cyclist collisions before and after the use of Citi Bike. I wanted to see where the most collision-heavy areas are located, and how they have changed accordingly after Citi Bike has gained widespread popularity. I used zip code as the area metric to aggregate the number of cyclist collisions. Then I chose the top ten zip codes that have the highest number of motor vehicle-bicycle collisions. As illustrated in Figure 4, the spatial patterns of cyclist collisions have become quite different after the use of Citi Bike. Noticeably, Manhattan has become more susceptible to cyclist collisioins than before, and there has been less cyclist collisions in the lower Downtown Brooklyn area.

2) Citi Bike Spatial Pattern

It is tempting to think that Citi Bike has had an effect on the biking flow in New York City, which has triggered a locational change in term of the number of cyclist collisions. Therefore, I wanted to see where the most popular stations are located, and compare them with the locational change in the collision patterns. I aggregated the number of trips by each station, both start and end, over four months—January, April, July and October—to mitigate seasonal biases. The result turned out to be that all the hotspots are located in Manhattan, mostly lower Manhattan and a few in Midtown, as illustrated in Figure 5. Another observation is that 8 out of the 10 most popular start stations and end stations are identical.

3) Spatial Correlation

It is reasonable to infer that the overwhelming influx of biking trips in Manhattan has labeled two additional zip codes within the borough as cyclist colllision-heavy zones. On the other hand, such influx may have shifted the biking density in Brooklyn to some extent, resulting in cyclist collisions moving from one area to another. Due to the complexities involved in quantifying such spatial correlation between cyclist collisions and Citi Bike trips, I developed my spatial analysis mainly from visually comparing the spatial trends of the two and draw potential connections from such a comparison. More details are discussed in the next chapter.

C. Figures and Tables

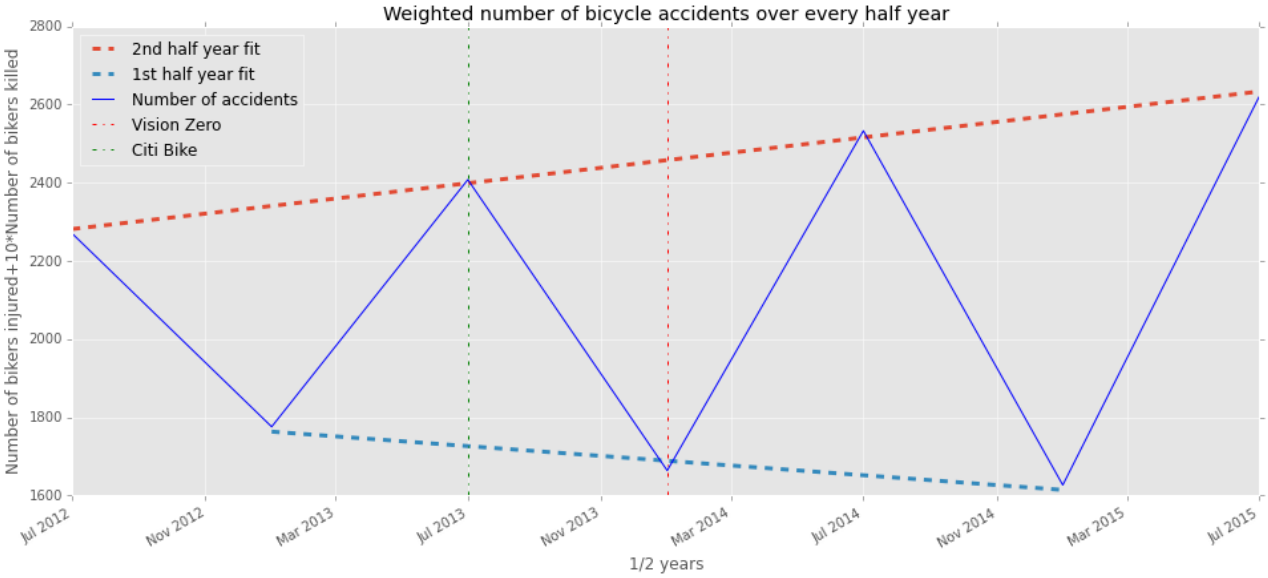

Figure 1 Weighted number of bicycle accidents over every 1/2 year

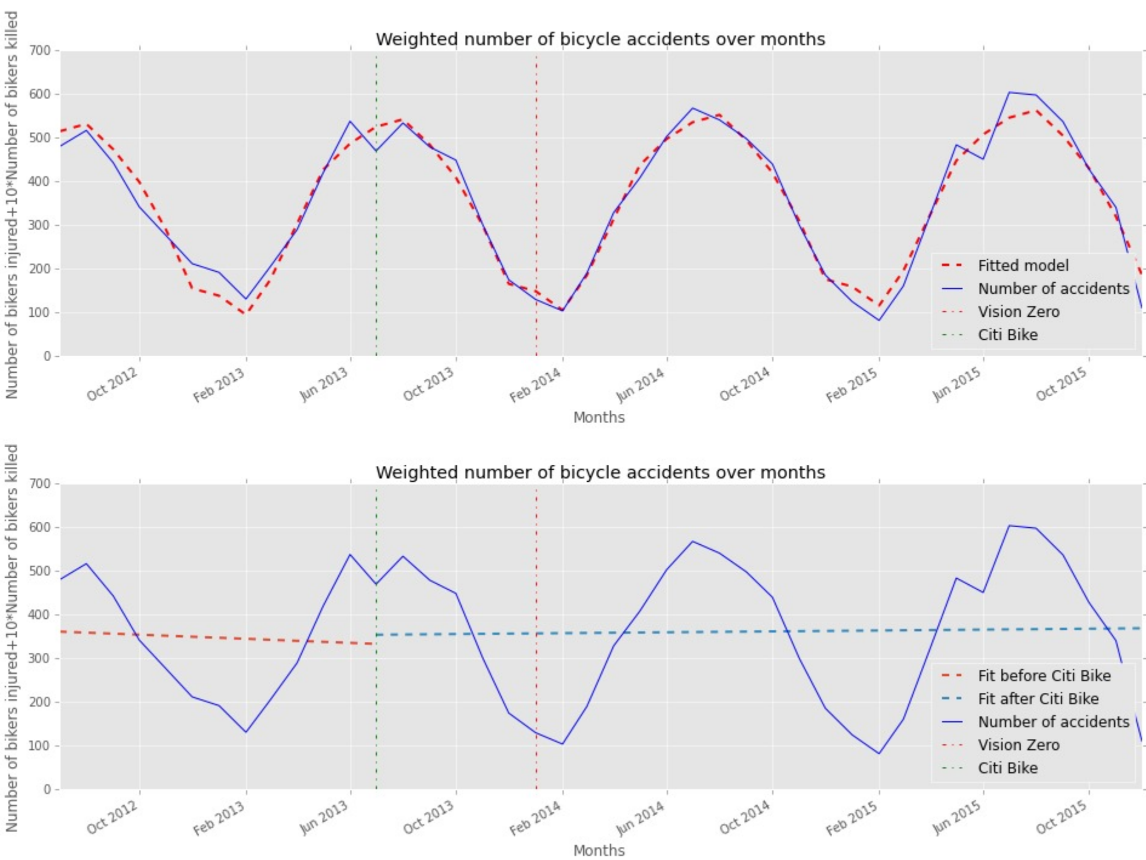

Figure 2 Weighted number of bicycle accidents over months

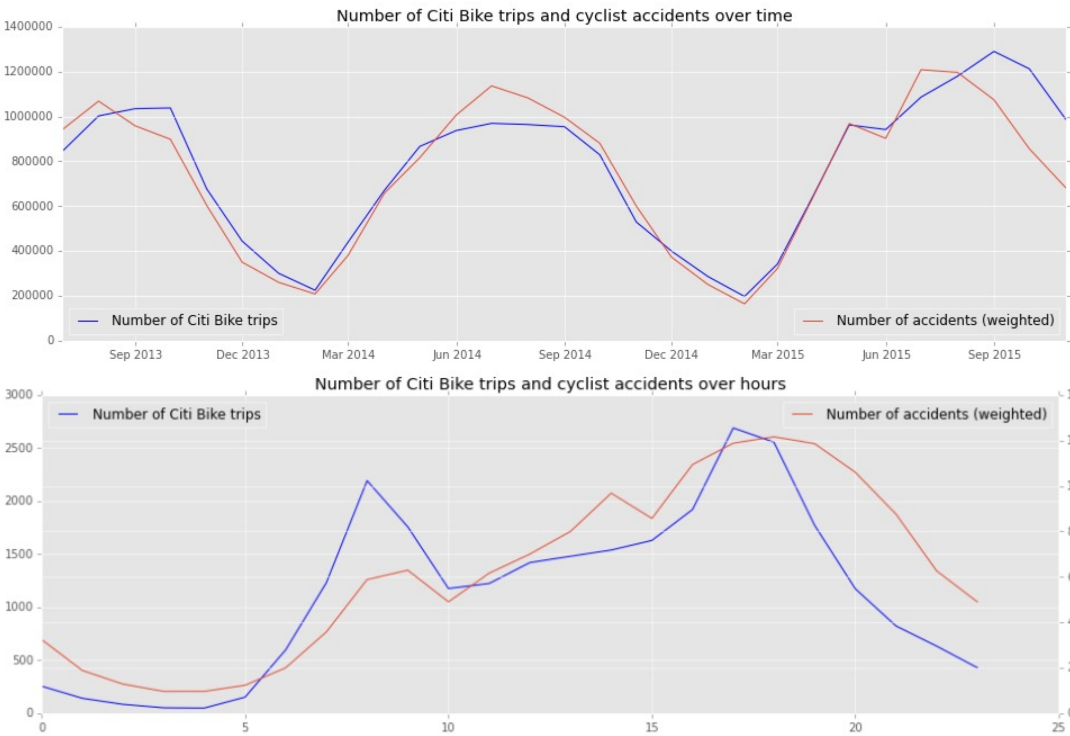

Figure 4 Temporal correlation between cyclist accidents and Citi Bike trips

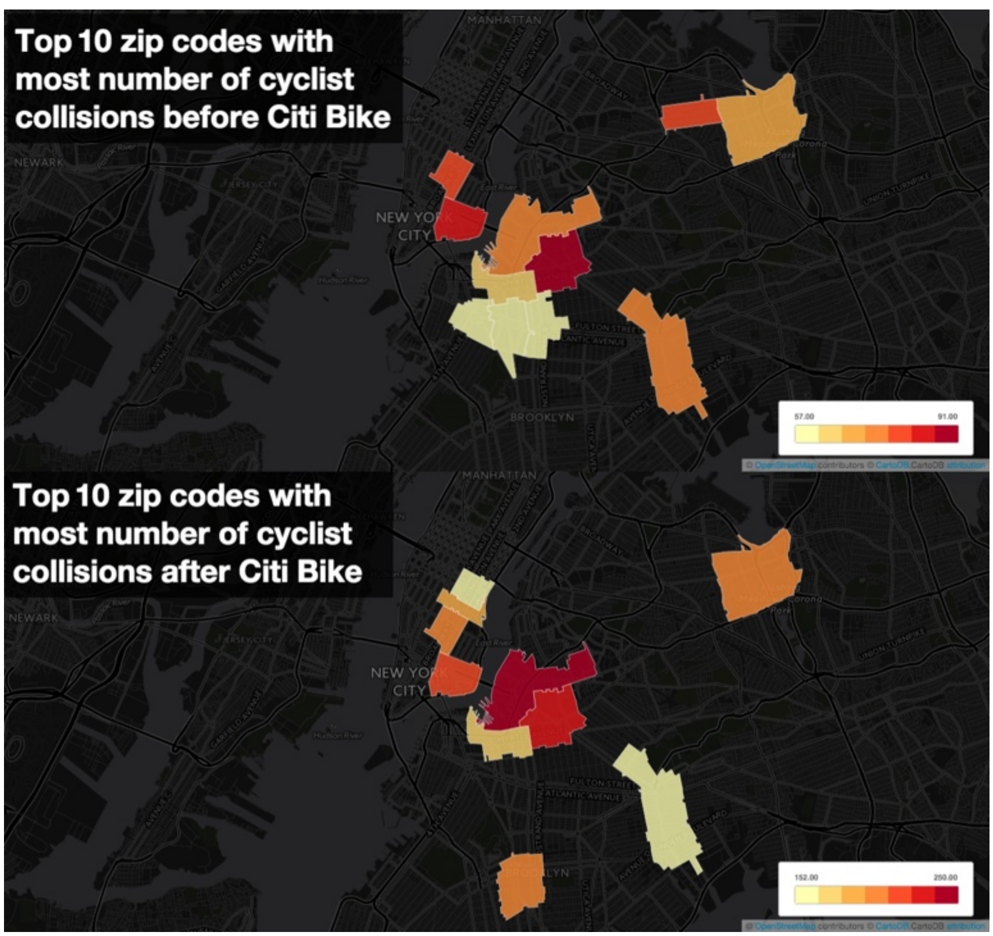

Figure 3 Collision-heavy zones before and after Citi Bike

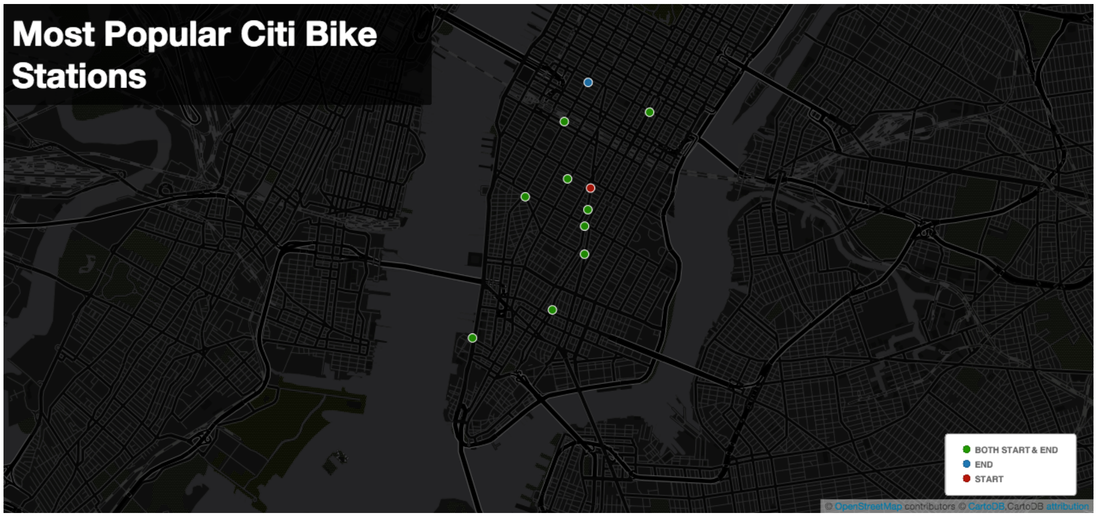

Figure 6 Most popular Citi Bike stations

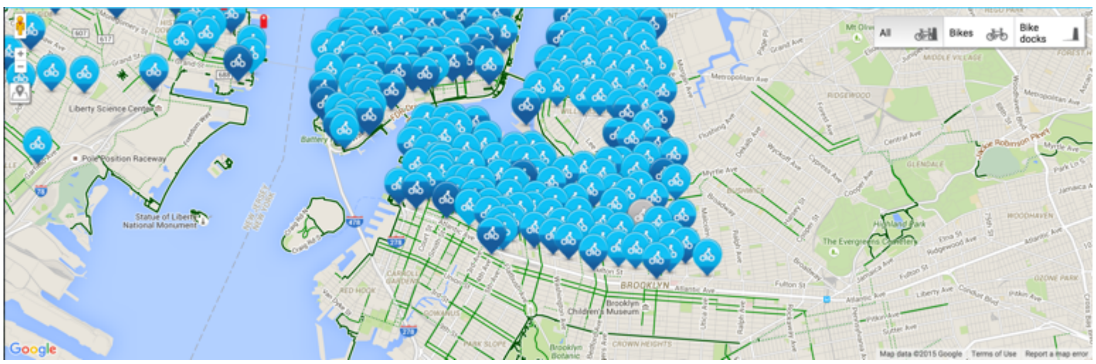

Figure 5 Citi Bike stations [10]

V. Results and Implications

From the temporal analysis, I made three major observations in cyclist collision patterns.

- There are significant higher number of cyclist collisions from July to December than from January to June in each year
- The number of cyclist collisions has been increasing for the second half of each year and descreasing for the first half of each year over the course of the last three years
- After Citi Bike, there has been a growing trend in the number of motor vehicle-bicycle collisions

These observations are able to shed some light on how Citi Bike has generally affected the bicycle safety situation in New York City and which months we may want to prioritize in terms of approaching the issue of biking safety. From the correlation plot, I was able to draw conclusions that the number of Citi Bike trips is highly correlated to the number of motor vehicle-bicycle collisions in New York City, and such correlation is stronger by months than by hours. The reason for why the correlation is less in the hours case may be that Citi Bike users tend to bike mostly for commute purposes, according to what the peaks of the blue line suggests in Figure 3, whereas there can be a great number of recreational rides which are equally susceptible to cyclist collisions. The difference in the usage pattern between Citi Bike and the whole biking set creates such a less correlation between the number of total collisions and the number of Citi Bike trips during each hour within a day. In addition, Figure 3 suggests that lots of cyclist accidents happen at night, so it is crucial for the city to install certain lighting equipments on the street to prevent accidents from happening. On the other hand, it is also crucial for Citi Bike to install effective lighting on their bikes as well to mitigate the risk of collision. Nevertheless, the high correlation by months implies that the increased volume of biking trips in the summer may result in a greater number of cyclist accidents. Therefore, if certain agencies want to start implementing safety countermeasures to cope with the issue of cyclist collisions, it can be potentially helpful to take seasonal features into the design of such countermeasures. For example, if bike lanes are to be added to a biking-heavy area in Manhattan, it will be useful to install additional bike lanes particularly to be implemented for the summer. During the summer when biking density is extremely high, the additional bike lanes will be used; during the winter when biking density is relatively low, the additional lanes will cease operations to facilitate the overall traffic flow. Similar policy can be applied to hours within a day as well. It may be helpful to prioritize bike lanes during rush hours to offset the higher risk of cyclist collisions due to the increased biking density. On the other side of the operation, Citi Bike may inform their users on the road safety conditioins based on real-time biking traffic. In addition, they may require Citi Bike users to wear helmets under certain scenarios, such as during rush hour in the summer or on a certain day that has horrible weather.

From the spatial analysis, I also made three major observations in cyclist collision patterns.

- All the top ten most popular Citi Bike stations are located in Manhattan
- Manhattan has become more subject to cyclist collisions since the initiation of Citi Bike
- The cyclist collision patterns in Brooklyn have shifted in accordance to the use of Citi Bike

The first two points have already been emphasized in the Methodology section, whereas the last point requires further explanation. From Figure 4, we know that south Williamsburg used the to be the most cyclist collision-heavy neighborhood before Citi Bike. After Citi Bike, however, the most cyclist collision-heavy neighborhood has become its neighbor, which is the waterfront part of Williamsburg, closer to Manhattan. The potential implication we can draw from this is that after the public bike sharing system came into being, people living in the waterfront area tend to use bikes more often, since it creates significant convenience for them to commute to Manhattan. Citi Bike seems to be a more efficient and economical means of transportation. The increased number of trips in the region due to the installation of Citi Bike stations makes North Williamsburg more susceptible to cyclist collisions than before. In addition, there seems to be a decrease in the number of bicycle accidents in lower Downtown Brooklyn, as the three adjacent zip codes marked in yellow in the "before" map disappeared all together in the "after" map. It may be tempted to think that there is a relative decrease in the volume of biking trips within the area, which leads to a decrease in number of crashes. A screenshot of the distributed Citi Bike stations, Figure 6, shows that there is no stations provided below the Atlantic Avenue, reassuring us that the area does not have as much of an influx of biking trips as its northern neighbors. The distribution pattern of Citi Bike stations may have somehow impacted the biking densities around Downtown Brooklyn. The increased number of biking trips in the surrounding neighborhoods of lower Downtown Brooklyn may have made the number of cyclist collisions within the area seem to be decreasing. In fact, it may just be that the surrounding areas with Citi Bike stations have more biking trips nowadays which result in more accidents than before. Comparing Figure 4 and 6, we do see that areas with plenty of Citi Bike stations, such as lower Manhattan and Williamsburg, tend to have a certain level of increase in the number of cyclist collisions. From an operational standpoint, the city should consider installing subsequent safety countermeasures to accommodate the influx of biking trips in those areas equipped with lots of Citi Bike stations.

Additionally, there is one area on the map that deserves attention, which is the polygon at the center of the bottom on the "after" map. This area does not seem to have any Citi Bike stations. Therefore, the plausible speculation for the increased number of cyclist collisions is that it is one of those areas where "people are more likely to bike to work but their environments do not provide enough safety for them" [8]. This confirms my previous assumption that some of those bicycle accident-heavy areas may be far away from the CBDs. Different from those areas that have a complex land use pattern, the main reason for such high crash rate in these regions is the lack of biking infrastructure rather than the excess of biking density. Therefore, from an operatioanl standpoint, it is imperative to install necessary facilities like bike lanes for such area, because its residents rely heavily on them.

One major limitation in my approach is that only the Citi Bike trips are taken into account of the calculation, whereas the total biking trips may exhibit a different temporal and spatial pattern in terms of quantity and density. Therefore, my analysis and conclusions are based on the assumption that Citi Bike trips have similar behavior, temporally and spatially, to all the biking trips in New York City. Taking the North Williamsburg as an example, I previously concluded that the increased number of cyclist accidents in the area is due to an influx of biking trips generated by Citi Bike. On the other hand, however, the lack of biking trip data collected for other types of bicycles makes it extremely difficult to get a handle on the big picture of biking density in New York City. Also, the lack of such data makes it extremey difficult to quantify the impact that Citi Bike actually has on the overall biking situation in the city.

Another limitation particularly in the case of spatial analysis is that the most popular routes may be a potentially better indicator for the biking density aross different neighborhoods than the most popular stations. However, we do not have data on the specific routes of each Citi Bike trip. Knowing the start and end locations of each trip, we can at most simulate or estimate the routes that each biker takes, but handling such process will require painstaking efforts. Collecting relevant information on the specific routes of each trip will result in a richer database, which can give us a better handle on the real-time scenario of biking traffic flow in the city.

VI. Conclusions and Recommendations

To sum up the findings from the analysis, Citi Bike does have a certain degree of impact, temporally and spatially, on the cyclist collision patterns in New York City. Temporally, the rate of collisions has increased since the implementation of the public bike sharing system. Besides, there is a high correlation between the number of monthly cyclist collisions and the monthly usage pattern of Citi Bike, and summer has the highest cyclist collision rate. From the city's operational standpoint, New York City may take on certain safety countermeasures based on the temporal patterns, such as creating additional bike lanes for the summer and reinforcing road lighting at night. From Citi Bike's operational standpoint, they may require bikers to wear helmets during rush hours and install effective lighting equipment on their bikes.

Spatially, the rate of collisions have increased in those areas that have plenty of Citi Bike stations, and those areas that are not equipped with Citi Bike stations exhibit a different, if not opposite, behavior. Specifically, as a result of the biking trip influx generated by Citi Bike, Manhattan has more cyclist collision-heavy zones than before. Also, the distributing patterns of Citi Bike have caused the collision patterns to shift a certain degree in Brooklyn, being that areas not equipped with Citi Bike stations have a lower cyclist collision rate than those that are. From an operational standpoint, it is important for the city to install subsequent safety countermeasures to offset the collision risks followed by the influx of biking trips due to Citi Bike stations. Nevertheless, remote areas with inefficient biking infrastructure should never be ignored, and necessary facilities such as bike lanes should be installed for such remote areas that rely heavily on bicycles.

In the future, it will be helpful to collect data regarding the routes of each Citi Bike trip, in order to get a better handle on the real-time biking density situation in New York City. For city agencies, it will be helpful to collect biking trip data on more sets of bikes in addition to Citi Bike, in order to get a sense of the bigger picture. Such data should be collected on a neighborhood scale with the purpose understanding how the biking traffic flow shifts over time for each particular neighborhood. With such data available, we will have a better way to quantify, instead of estimating, the actual impact that Citi Bike has had on each local community across the city.

Acknowledgment

The author would like to acknowledge NYU Libararies; Center for Urban Science and Progress; NYPD; NYC Open Data; Citi Bike for providing special access to peer-reviewed research articles and public access to open datasets.

References

[1] H. Lucy Wang, “U.S. Bicyclist Deaths in Motor Vehicle Accidents Increased Dramatically,” in American Journal of Nursing (AJN), Vol. 115, No. 2, p. 17, Feburuary 2015.

[2] W. Allan, “Bicyclist Safety—Prepared for Governors Highway Safety Association,” in Spotlight on Highway Safet, pp. 1–19, 2014.

[3] B. H. Corey, Z. A. Emily, R. Sonali, E. Danna, and B. E. Charles, “Helmet Use Among Cyclists in New York City,” in Journal of Community Health, Springer Science+Business Media B.V., pp. 956–958, 16 February 2014.

[4] C. Li, C. Cynthia, S. Raghavan, M. E. Claire, W. Reid, and R. Matthew. “Evaluating the Safety Effects of Bicycle Lanes in New York City,” in American Journal of Public Health, Vol. 102, No. 6, p. 1120–1127, June 2012.

[5] C. Li, C. Cynthia, E. Reid, M. E. Claire, S. Raghavan, and R. Matthew. “Safety countermeasures and crash reduction in New York City—Experience and lessons learned,” in Sciverse Science Direct, Accident Analysis and Prevention, pp. 312–322, 2 May 2012.

[6] T. Ivana, and P. J. Richard, “Modeling spatial relationships between multimodal transportation infrastructure and traffic safety outcomes in urban environments,” in ScienceDirect, Safety Science, pp. 325–337, 27 October 2015.

[7] C. Peng, and S. Qing. “Built environment effects on cyclist injury severity in automobile-involved bicycle crashes,” in ScienceDirect, Accident Analysis and Prevention, pp. 239–246, 21 November 2015.

[8] Y. Chia-Yuan. “Environmental supports for walking/biking and traffic safety: Income and ethnicity disparities,” in ScienceDirect, Preventive Medicine, pp. 12–16, 28 June 2014.

[9] C. Shailesh. “Safety-based path finding in urban areas for older drivers and bicyclists,” in ScienceDirect, Transportation Research Part C, pp. 143–157, 18 September 2014.

[10] Citi Bike Station Map. URL: <https://member.citibikenyc.com/map/