Yuan Wang

What is Citi Bike?

Citi Bike is a public bicycle sharing system serving New York City, NY and Jersey City, NJ. It is a fun and affordable way to get around the town, faster than walking, cheaper than a taxi, and more fun than the subway. The business started in May 2013 with 332 stations and 6,000 bikes. Annual expansions have brought the totals to 750 stations in 60 neighborhoods and 12,000 bikes as of October 2017, making this service the largest bike sharing program in the United States.

Is Citibike popular?

Citi Bike designed for quick trips with convenience in mind. When become a member, the rider can unlock thousands of bikes around Manhattan, Brooklyn, Queens and Jersey City. The riders took an average of 38,491 rides per day in 2016. The system reached a total of 50 million rides in October 2017. As of November 2017, the total number of annual subscriber breakthrough 250,000.

The daily trip number of Citi Bike changes dramatically over a year. Figure 1 displays a very clear seasonal pattern. Not surprisingly, there are extremely fewer Citi Bike rides during the cold and snowy winter with only 5 thousand trips a day, but the daily trip number can jump up to nearly 75 thousand during the summer holidays.

This paper uses the Citi Bike trip data of September 2017. There are total 1.9 million trips during this month: the most popular station has 34,415 transactions; the most popular bike has been ridden for 419 trips. Unlike taxis, Citi Bikes cannot pick up and drop off at any arbitrary point in the city. Instead, riders can pick up and drop off bike at finite number of stations across the city.

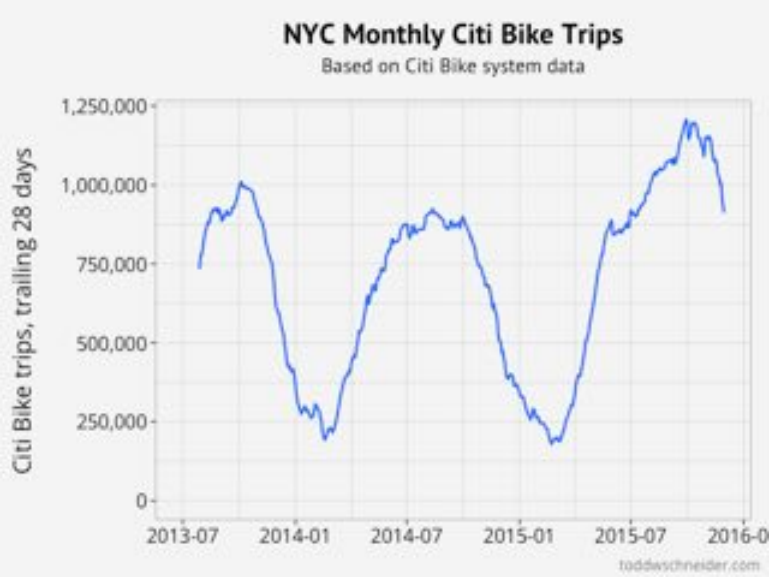

Figure 1. Monthly total trips from the year 2013 to 2015 [1]

Figure 2 shows the popularity of the stations utilization in one month, the bike docks in midtown and downtown Manhattan are frequently used than other locations.

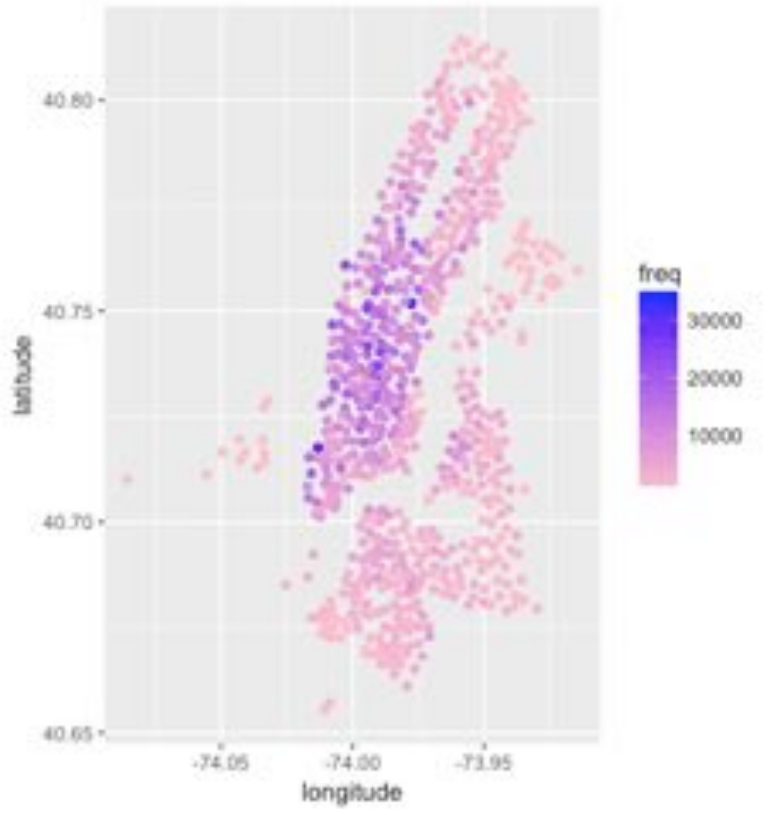

Figure 2. Bike stations utilization in September 2017

Who use the Citibike?

The minimal age to be a Citi Bike subscriber is 16-year old. From Figure 3, 60% of biker are between age 20 to 40, and 30% of whom are between 40 to 60 years old. The mean age of the subscriber is 37 years old.

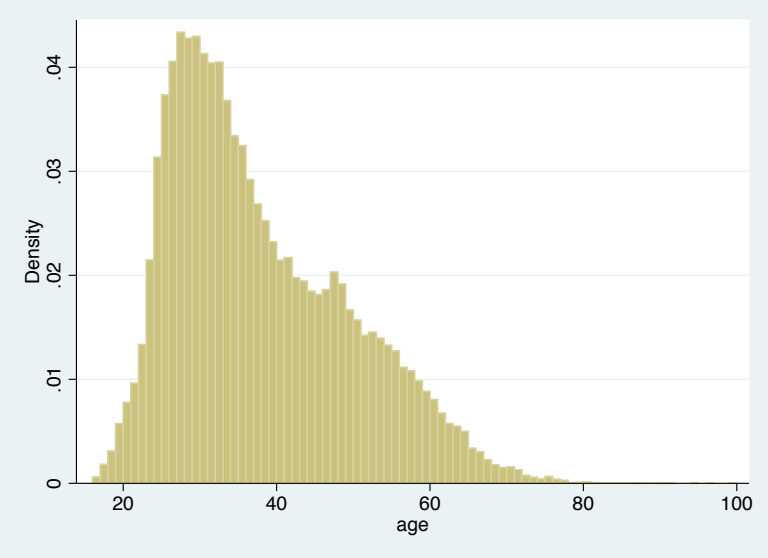

Figure 3. Age distribution of Citi Bike riders

About three quarters of bike riders are male, while female riders are much less than male. The average time duration in a single trip for male rider is 12.6 minutes, and female riders take 2 more minutes than male. This is because, on average, women's riding speed is slower than man.

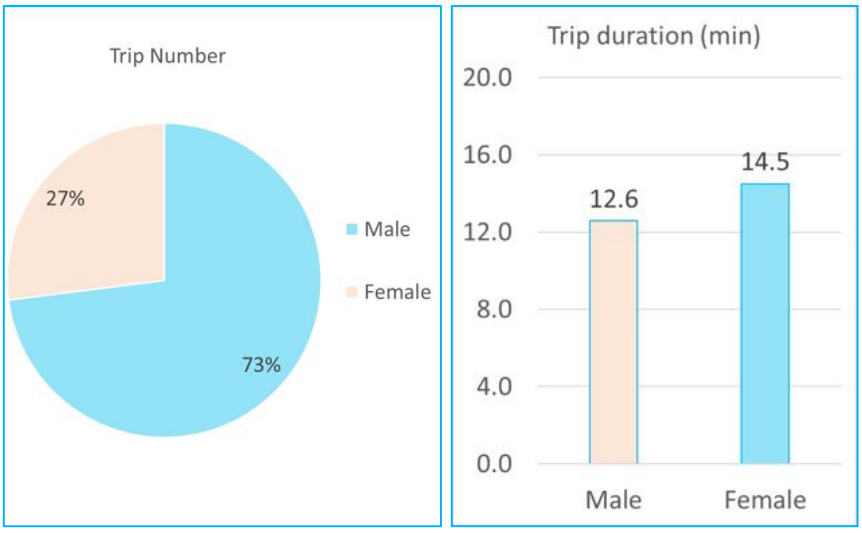

Figure 4. Trip number and duration for male and female

Citi Bike offers 2 plans, day pass and annual membership. The annual membership is the best deal for residents and other frequent riders. Members enjoy the 45 minutes of each ride, and as many rides as they want throughout the year. The short-term passes are perfect for tourists and visitors. It is a fun way to get around NYC on Citibike with a 30 minutes of single ride and take as many rides as you want while your pass is active.

There are totally 1.66 million Citi Bike trips in September 2017, 97% of which is ridden by annual members, and their average trip duration is only 13 minutes. The annual member are mostly NYC residents who use the Citi Bike to commute to work or school, run errands, get to appointments, and connect to the nearby subway or bus stations. For the less than 3% of trips ridden by day pass user, the average trip length is longer than 20 minutes. They use the Citibike to get to local attractions, or take a bike out for a ride around Central Park or the Hudson River Greenway. Thus the tourists would like to utilize most of the time allowance to visit as many as the place of interests.

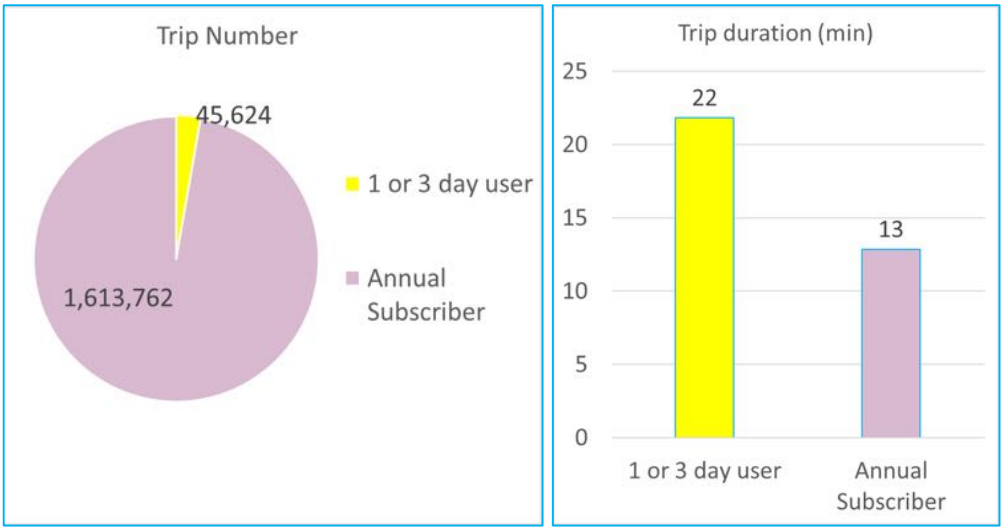

Figure 5. Trip number and duration for day pass user and annual member

We also observe the important pattern for the Citi Bike trips during one day and over the week. Figure 6 shows clearly that the Citi Bike system gets highly used on weekdays than on weekends, and if we look at trips by hour of the day, weekday riders primarily use Citi Bike to commute to and from work, with peak hours from 8–9 AM and 5–7 PM, while weekend riders, on the other hand, prefer a more leisurely schedule, with most weekend rides occurring in the midday hours from 10am-8pm.

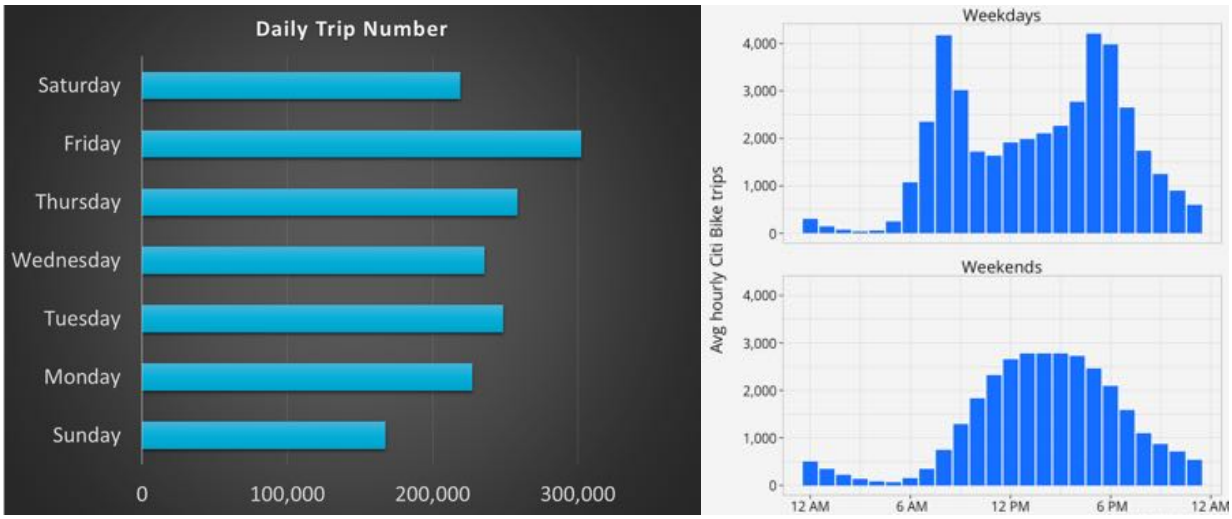

Figure 6. Trip number during one day and over the week

How to estimate the cycling time?

A previous study of bike riders shows the average speed across all Citi Bike trips is 8.3 miles per hour, and its graph makes clearly that younger riders tend to travel faster than older riders, men tend to travel faster than women, and trips covering longer distances have higher average speeds than shorter distance trips.

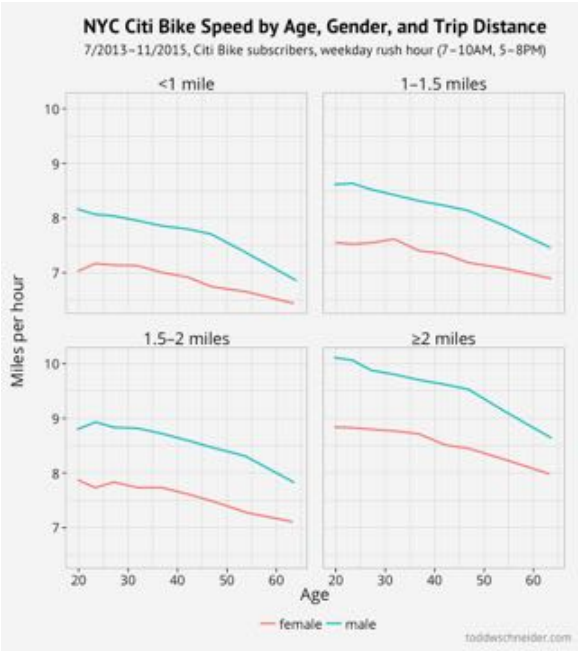

Figure 7. Bike speed by age, gender, and trip distance [1]

For each individual Citi Bike trip, our data includes demographic information about its riders, gender, birth year (age), annual membership subscriber status, and the trip date and time. These demographic data can address a host of interesting questions, including:

- How fast do Citi Bike riders tend to travel?
- How do age and gender impact biking speed?
- Are residents and visitors riding bikes in the same way?
- Are people riding the bike for same time for work and leisure over the days of week?

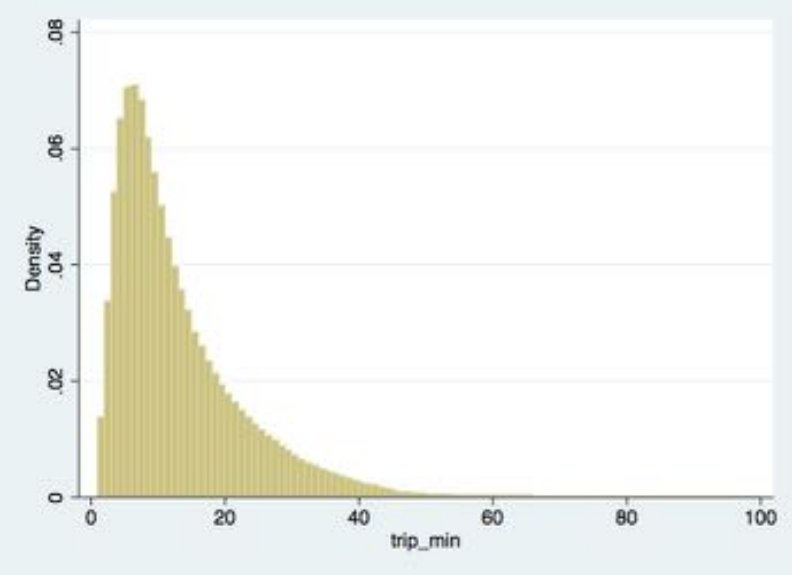

Figure 8. Distribution of trip time duration

To answer the questions above, we can conduct a qualitatively and quantitatively analysis of bike trip duration with respect to the riders' information. The regression model is

$$Tripduration = \beta_0 + \beta_1 \cdot Age + \beta_2 \cdot Gender + \beta_3 \cdot Membership + \beta_4 \cdot Weekday$$

, where ***Tripduration*** and ***Age*** are continuous variables, and ***Gender*** (female =1, male = 0), ***Membership*** (annual subscriber = 1, day pass user = 0), and ***Weekday*** (weekday = 1, weekend = 0) are three dummy variables.

The regression results listed in Table 1 shows that

(1) ***Age*** and ***Gender*** have positive coefficients, which means the trip duration is longer when age is increasing and for female riders. In other words, younger riders tend to travel faster than older riders, men tend to travel faster than women.

(2) ***Weekday*** and ***Membership*** have negative coefficients, indicating that a short bike trip happens in weekdays and for annual subscribers. And the coefficient for membership has the largest scale of effect. Thus, visitors with daily pass have a much longer and leisure trip during the weekend.

Table 1. Regression results for trip duration with rider's information

	Trip duration	
Age	0.03	***
	(0.00)	
Gender	1.81	***
	(0.02)	
Weekday	-0.70	***
	(0.02)	
Membership	-8.71	***
	(0.05)	
Constant	20.54	***
	(0.05)	
R^2	0.03	
N	1659386	

Where are the Citi Bike stations located?

There are more than 600 stations located across Manhattan, Brooklyn, Queens and Jersey City. The user can easily find an available bike nearby on the websites or their mobile app. Most Citi Bike trips start and end in Manhattan. And sure enough, in the mornings there are more rides from outer boroughs (Brooklyn) to Manhattan than vice versa, while in the evenings there are more people riding from Manhattan to outer boroughs (Brooklyn).

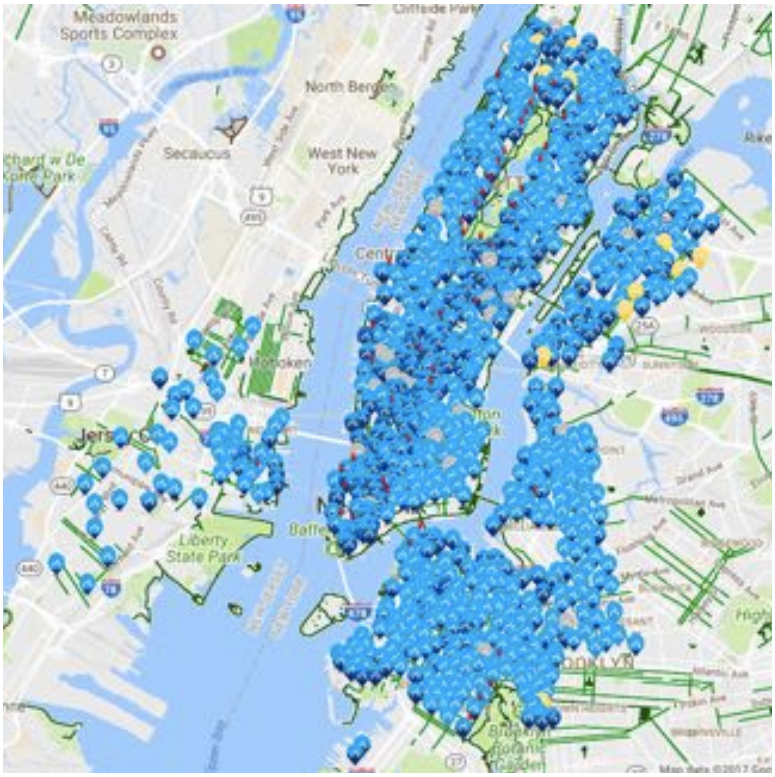

Figure 9. Citi Bike stations on map

It is true that, 88% of trips start and end in Manhattan, 8% of trips start and end in an outer borough, and only 4% of trips travel between Manhattan and an outer borough in our data set. But it might be thoughtless to claim that the bike docks are concentrated in Manhattan and sparsely distributed around three outer boroughs. In this section, I would explore the stations location more carefully by clustering analysis with K-means algorithm.

The Citibike trip data includes the start- and end-station name, ID, latitude and longitude, thus the stations can be mapped into a Lat-Long grid. Then I set the cluster parameter K equals 1, 2, 3, 4, and repeat each calculation with random initial partitions by 20 times. As shown in Figure

10, the bike stations are first grouped into two with a cutting line across the midtown and the dividing line between Brooklyn and Queens. With the number of clusters increasing from 2 to 3, the Brooklyn area becomes a new cluster separating from lower Manhattan area. When K equals 4, the upper and lower Brooklyn become two clusters, and several stations in Manhattan belongs to the lower Brooklyn group due to the connection by Brooklyn Bridge. The Queens borough is still bound with upper town, due to a lower density of stations distribution around Central Park in Manhattan and Long Island City in Queens. As the number of clusters increases to 5 or 6, the results are not as significant as before.

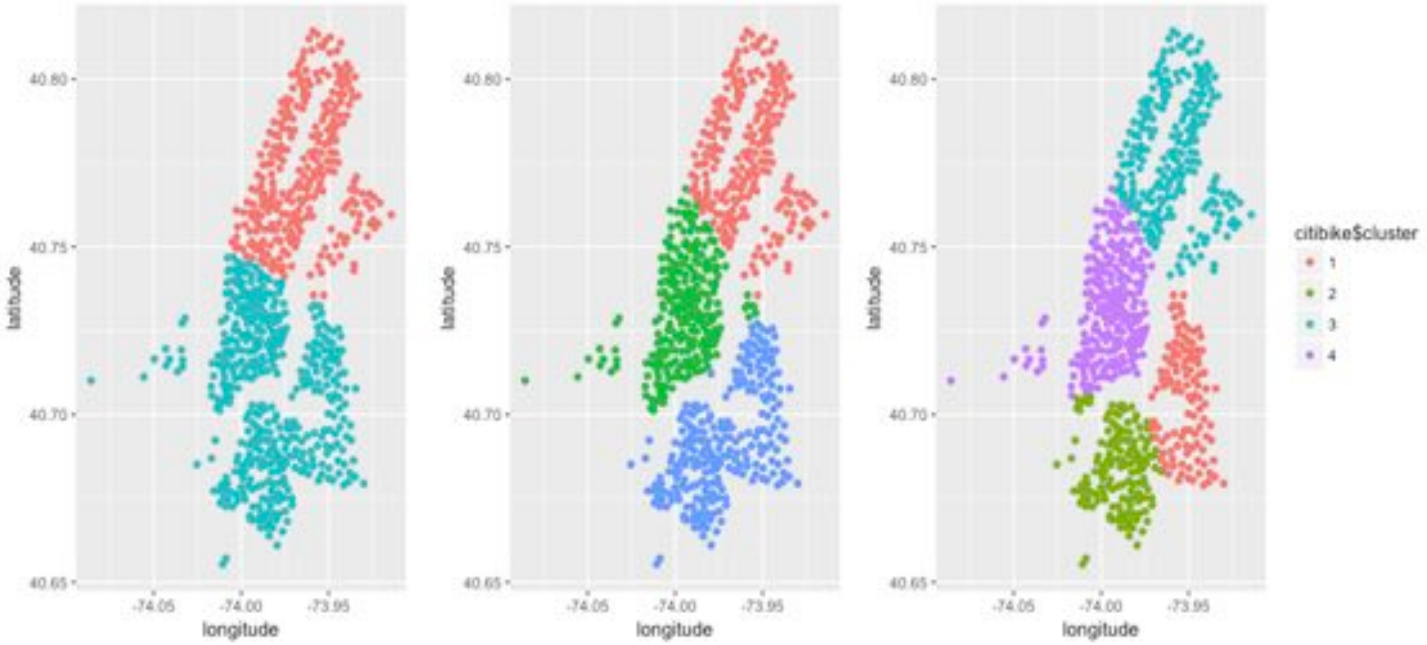

Figure 10. Clustering analysis of Citi Bike stations

At the end…

Citi Bike provides a fun and affordable way to get around the city, and a great alternative to the subway, taxis, buses, and walking. Enjoy bike riding, enjoy the NYC!

---

[1] *A Tale of Twenty-Two Million Citi Bike Rides: Analyzing the NYC Bike Share System*, by Todd W. Schneider
[2] Citi Bike trip data and station map are available from www.citibikenyc.com.