

# Data Source

---

## Website:

New York City : <https://www.citibikenyc.com/system-data>

some other cities' data : <https://github.com/BetaNYC/Bike-Share-Data-Best-Practices/wiki/Bike-Share-Data-Systems>

# Data Description

---

## Data schema:(from website)

The data includes:

- Trip Duration (seconds)
- Start Time and Date
- Stop Time and Date
- Start Station Name
- End Station Name
- Station ID
- Station Lat/Long
- Bike ID
- User Type (Customer = 24-hour pass or 3-day pass user; Subscriber = Annual Member)
- Gender (Zero=unknown; 1=male; 2=female)
- Year of Birth

This data has been processed to remove trips that are taken by staff as they service and inspect the system, trips that are taken to/from any of our “test” stations (which we were using more in June and July 2013), and any trips that were below 60 seconds in length (potentially false starts or users trying to re-dock a bike to ensure it's secure).

## Data duration

The data contains cycling record from 2013.07 to 2017.09.

Until 2017.09, the data contains 49669470 riding records and 827 bike stations.

# Get the appear time for each station and visualization

---

## Hypothesis

If the bike station is being used, it will appear in the data.

## Method to find the appear time:

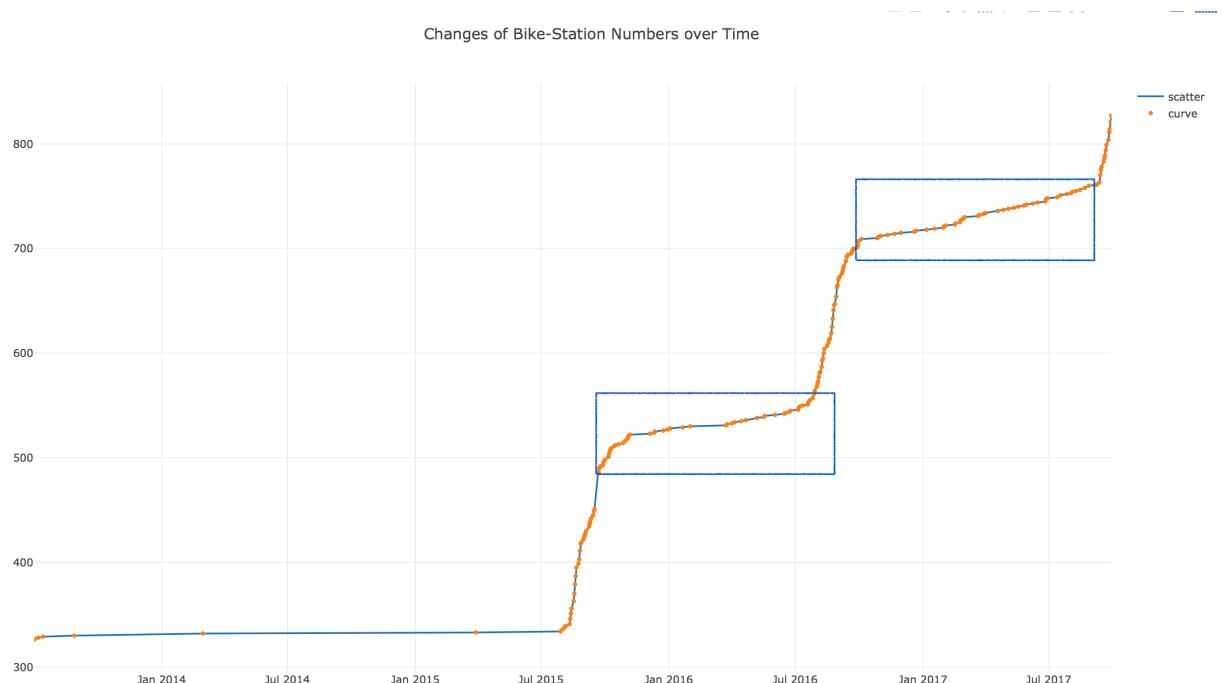
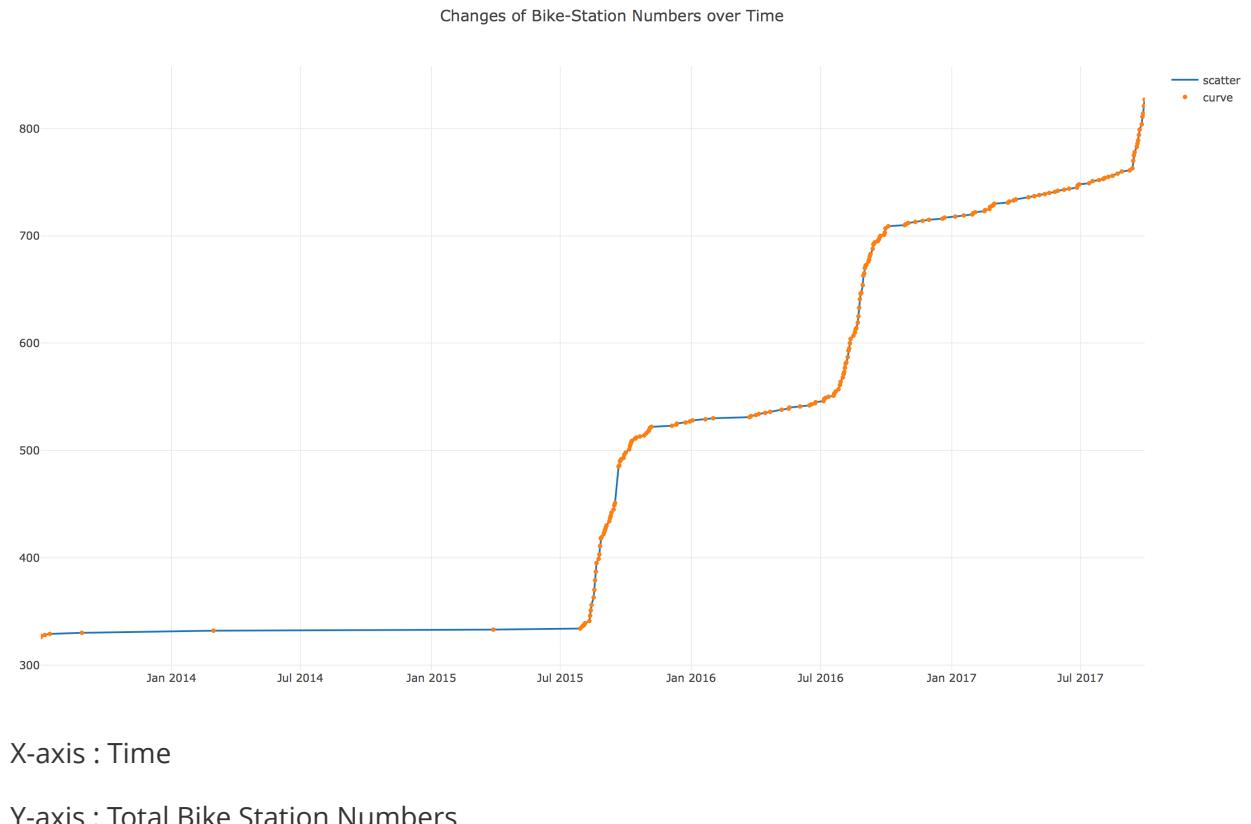
Go through all the records, and get the earliest appear time for each station.

## Result

The result is stored in 'stationAppearTime.json', the schema is like:

```
{'stationID': [AppearTime, lat, lng, StationName], ...}
```

## Visualization (1)

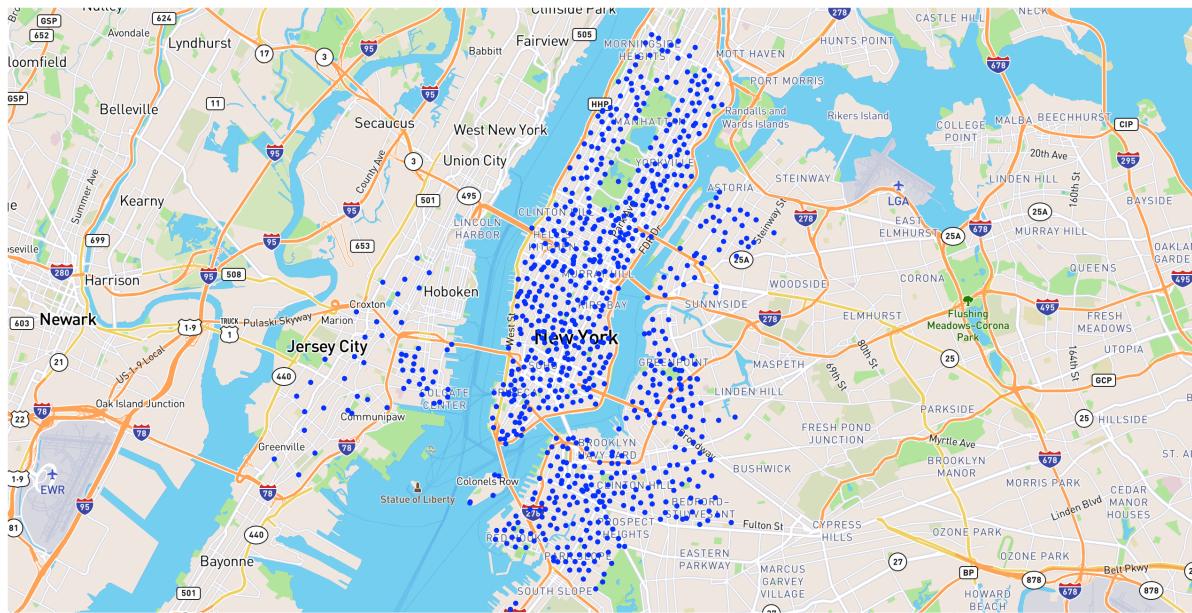


The curves in the two blue boxes are relatively flat and they both appear after a dramatic increase, so I think we can analyse the absorption/stimulation phenomenon over this two time blocks in next step.

## Visualization (2)

A OverView of Bike Stations' location (the blue points are Bike Stations)

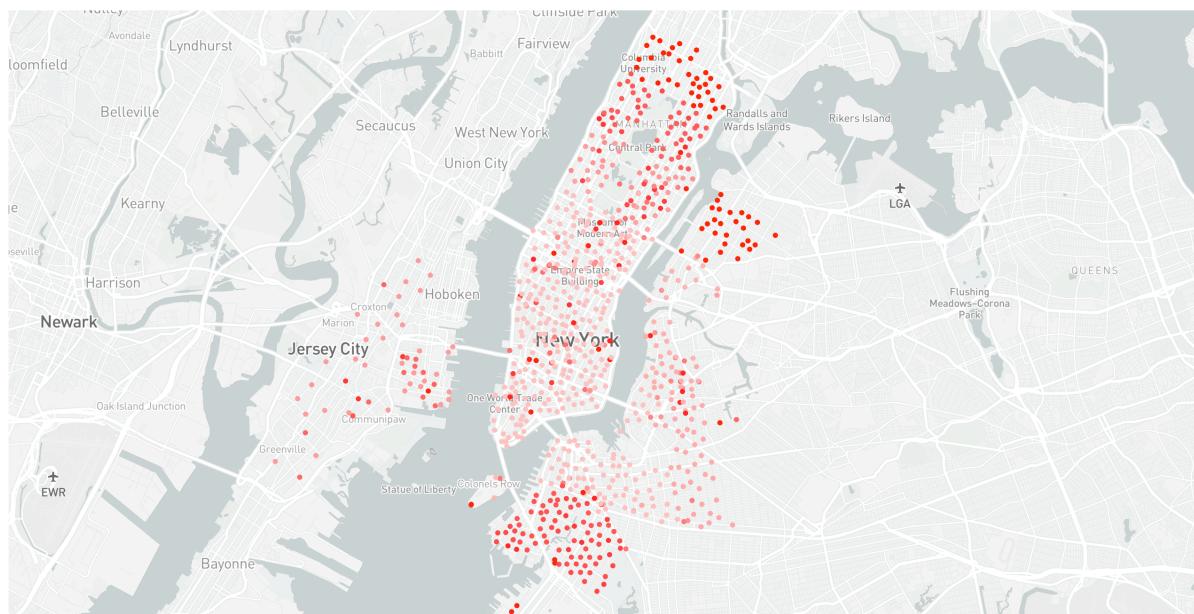
Bike Station Location



## Visualization (3)

Bike Stations' location with appear time(the red points are Bike Stations)

Bike Station Location & The latest built stations with deeper color



The color is deeper if the station is built recently.

## Source Code and Visualization File

The source code and visualization html can be found here:

<https://github.com/Di-Chai/BikeSharingResearch>