

In the Future

How do you think will the collection and analysis and insights evolve in the next 10 years?

I think collection might evolve to include trip paths and conditions and will allow for better urban planning.

I think over time insights will better predict use patterns and needs and operations will become more efficient with prediction accuracy.

I know that citi bike has a bike angels program where users can assist in rebalancing bike flows and hotspots. I think greater analysis of trends will allow for more optimized and dynamic pricing incentives that draw users to redistribute where it is needed most.

In which areas do you see the biggest future potentials or value of the data?

There is huge value in studying the use of human powered transportation which has a lower environmental impact, takes up less space, and allows people to experience their cities. Knowing how people use bike programs gives much more controlled measures of commuting and patterns over time than user owned bikes can provide. It could help justify new bike lanes, send users to the best stations near them, and quantify conditions that people are most likely to bike in. I think the biggest future potential is within sustainable development and urban planning where bikes are equitably available and supported through transportation systems.

Citi Bike



```
{  
    "num_docks_available": 0,  
    "is_renting": 1,  
    "station_id": "4d24ab76-2f4a-4004-9fc5-09c00c1d9aa4",  
    "is_installed": 1,  
    "num_bikes_disabled": 3,  
    "num_bikes_available": 0,  
    "last_reported": 1741011100,  
    "vehicle_types_available": [  
        {  
            "vehicle_type_id": "1",  
            "count": 26  
        },  
        {  
            "vehicle_type_id": "2",  
            "count": 1  
        }  
    ],  
    "num_scooters_unavailable": 0,  
    "num_bikes_available": 27,  
    "num_docks_disabled": 0,  
    "is_returning": 1,  
    "num_ebikes_available": 1  
}.
```

What is the technology or service?

Citi bike, under parent company Lyft is a bike sharing system that allows users to rent bikes from hundreds of stations and then return them to other stations. They are available at all hours and unlocked through an app. The collection of json docs that I looked at for Citi Bike NYC follow the General Bikeshare Feed Specification that many other mobility systems also use. It displays real-time data, although admittedly the last_reported value is not a time value that I know how to interpret so I am unsure of the update frequency.

What behaviors does it capture?

This data feed captures movement of bikes and scooters among citi bike docks in New York city. It shows how available different modes of transport are at different stations, geographic distribution of stations, and if logged the patterns of use. It doesn't provide historical data or personal data.

How does the raw data look like?

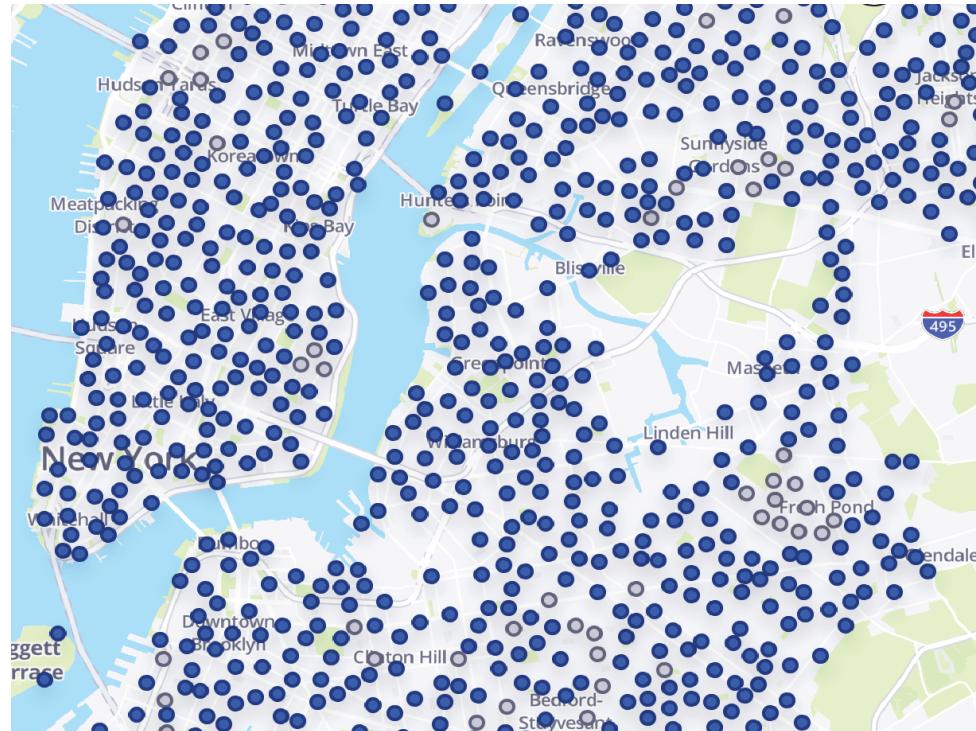
Raw data is split between a few json files. I think I would reference info from station status and station information.

Station status: station_id, count of vehicle types available (bike and ebike), num_bikes_disabled, num_docks_available

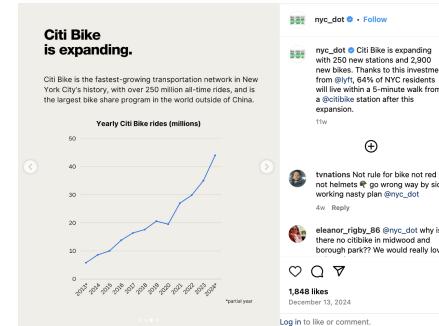
Station information: lat, long, station_id, name, region_id, capacity

Visualizations

Visual Examples of this technology:



Station map for NYC



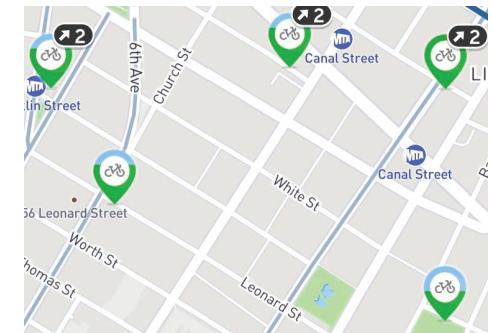
Graph on instagram of citi bike expansion

```
{  
  "lon": -73.8973876833916,  
  "short_name": "6417.11",  
  "station_id": "fa98fb78-51ab-4248-b7e7-f402902385a6",  
  "name": "65 St & Broadway",  
  "rental_uris": {  
    "android": "https://bkn.lft.to/lastmile_qr_scan",  
    "ios": "https://bkn.lft.to/lastmile_qr_scan"  
  },  
  "lat": 40.7493722738996,  
  "region_id": "71",  
  "capacity": 18  
},  
  
{  
  "lon": -73.8558,  
  "short_name": "6255.02",  
  "station_id": "1854538387432581604",  
  "name": "49 Ave & 108 St",  
  "rental_uris": {  
    "android": "https://bkn.lft.to/lastmile_qr_scan",  
    "ios": "https://bkn.lft.to/lastmile_qr_scan"  
  },  
  "lat": 40.74535,  
  "region_id": "71",  
  "capacity": 22  
},
```

Station information json snippet



citibike marketing image



Bike angels point map

Insights

What are the main types of insights that can be derived from the data?

Station Distribution

Geocoordinates of station locations provide insights to where they are distributed and therefore likely the biking population.

Vehicle type

Electric assist and human powered bikes are both present at many stations. We could learn where electric bikes are most available within the city.

Maintence

Disabled bikes and disabled docks are shown in the dataset so we could see if the most borrowed from station also have the most broken parts.

Total use

The available documents show snapshots of station states not paths of trips, but aggregated, the rise and fall of available bikes could show how many are in use.

Usage

How are the insights currently being used?
By who?

Company

The operations of citi bike depend heavily on display of available vehicles through their app and I'm sure internally as well to track maintenance and needs.

Urban Analysts

Transportation is a huge sector in cities like NYC and knowing real-time usage of different modes helps navigate congestion as well as planning

Sustainability Researchers

Policy proposals and new initiatives for lower environmental impact depend on understanding daily use of current systems and opportunities for growth.

Users

Citi bike users especially those who participate in the bike angels program look at what stations need more or less bikes in order to be compensated most for their redistribution.