

Team Project



Updated: 2021-12-08

New York City (NYC) Citi Bike Sharing

Objective:

Analysis bike trips, stations and system data to identify bike usage characteristics and trends

Data:

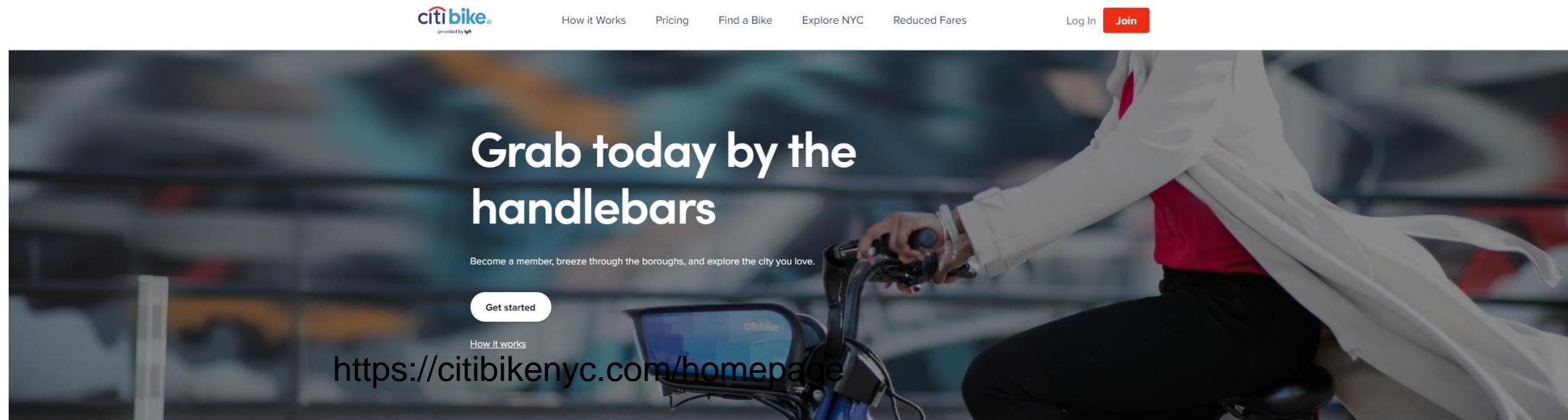
- Big Query Public Datasets
- Data Files
- Additional Data File



New York City (NYC) Citi Bike Sharing

Background

NYC Citi Bike



A fun and affordable way to get around town

Day Passes

\$15/day

Perfect for visitors and tourists.

24 hours of Citi Bike access, includes unlimited 30-minute rides on a classic Citi Bike.

[Learn more](#)

Annual Membership

\$15/month*

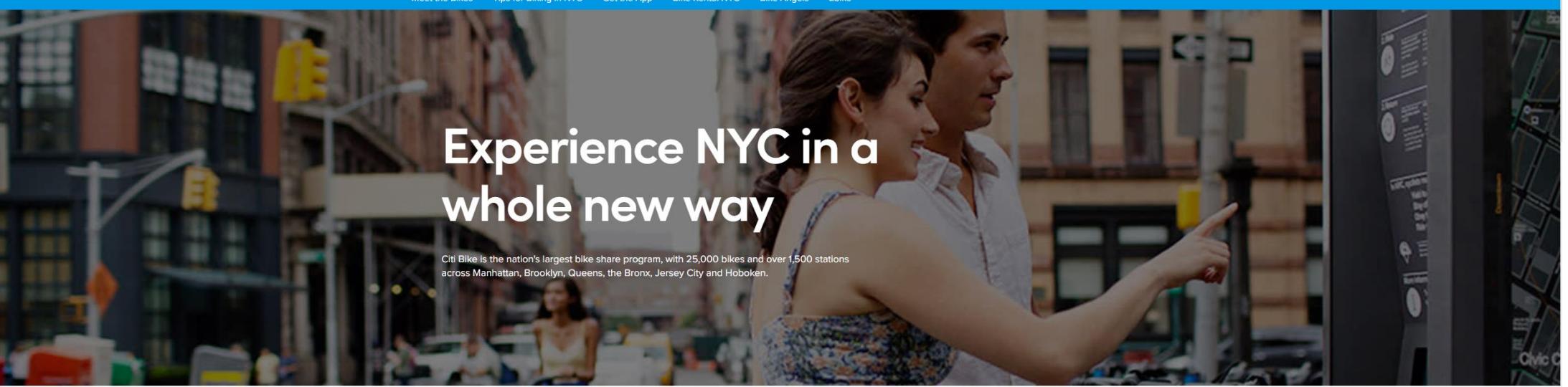
Best deal for New York locals.

Unlimited 45-minute rides on a classic Citi Bike, discounted ebikes, and more!

*\$179 billed upfront annually

[Learn more](#)

NYC Citi Bike



The banner features a man and a woman standing next to a Citi Bike station. The woman is pointing at the station's display screen, which shows a map of the NYC bike share network. In the background, there are city buildings and traffic lights.

Experience NYC in a whole new way

Citi Bike is the nation's largest bike share program, with 25,000 bikes and over 1,500 stations across Manhattan, Brooklyn, Queens, the Bronx, Jersey City and Hoboken.

How it Works Pricing Find a Bike Explore NYC Reduced Fares Log In **Join**

Meet the Bikes Tips for Biking in NYC Get the App Bike Rental NYC Bike Angels Ebike

1  **Join**
Become an Annual Member or buy a short-term pass through the [Citi Bike app](#).
[See prices](#)

2  **Unlock**
Find an available bike nearby, and get a ride code or use your member key to unlock it.
[Find a bike](#)

3  **Ride**
Take as many short rides as you want while your pass or membership is active.
[See popular rides](#)

4  **Return**
Return your bike to any station, and wait for the green light on the dock to make sure it's locked.
[Get started](#)

NYC Citi Bike

citibike®
provided by 

How it Works Pricing Find a Bike Explore NYC Reduced Fares Log In **Join**

Day Passes Single Ride Citi Cardholders Citi Bike for Business NYCHA/SNAP Management Benefits Fund Veterans & Military Discount

 For a limited time, save \$50 on annual membership e-gift certificates. [Shop now.](#) 

Choose your plan



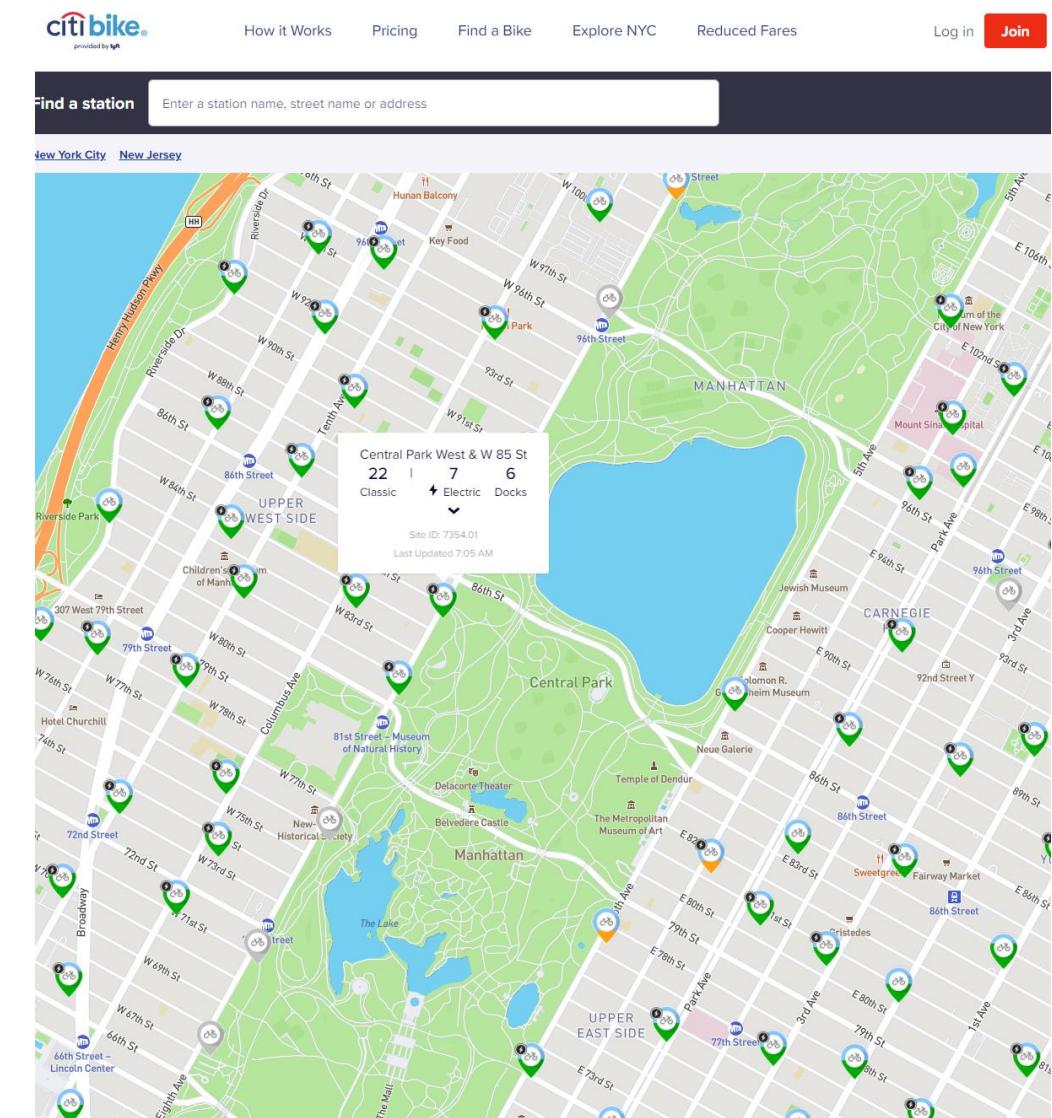
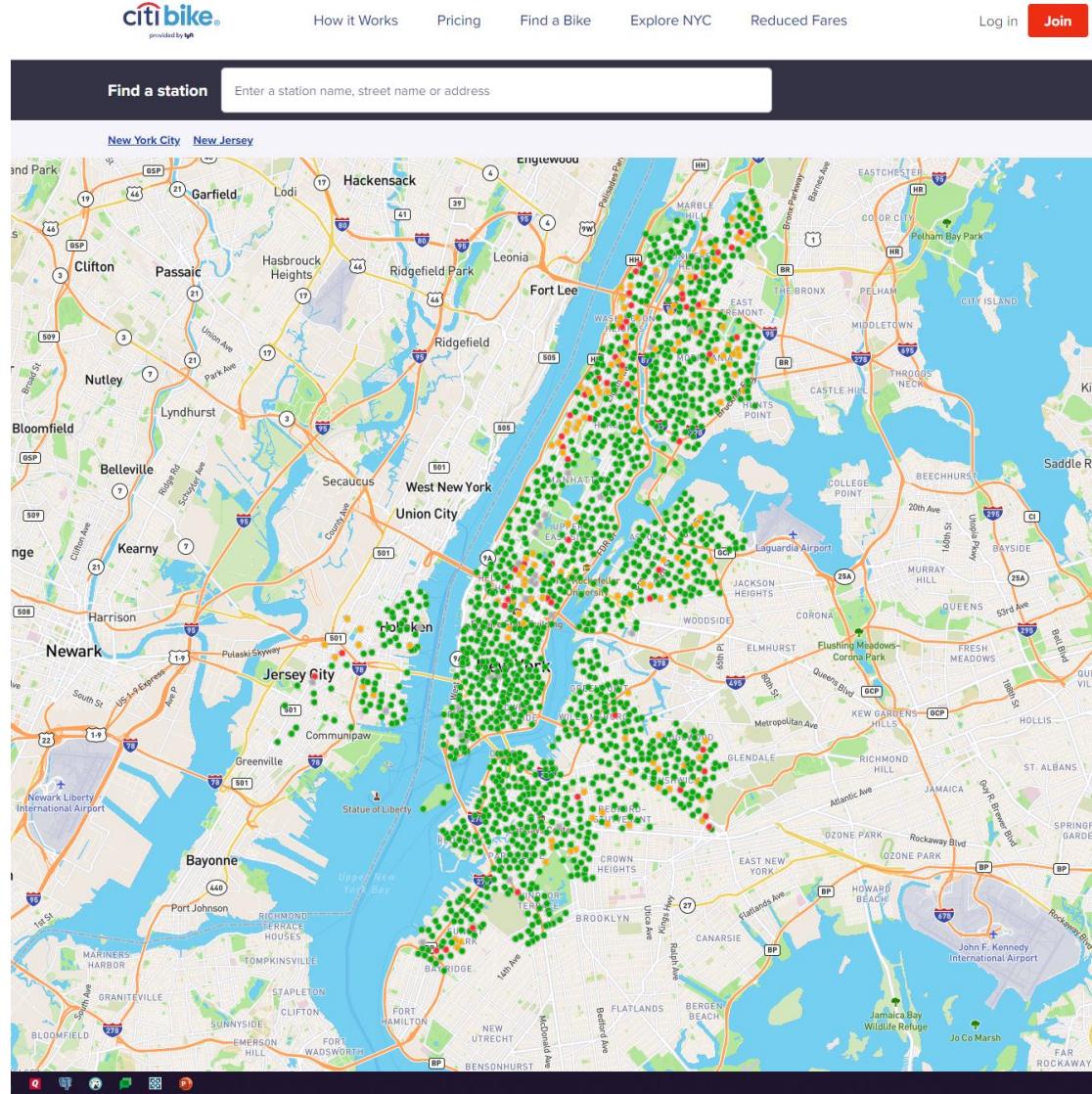
Annual Membership \$15*/month <small>unlimited 45-minute rides on a Classic bike.</small> Join Now <small>*\$179 billed upfront annually</small>	Single Ride \$3.50/trip <small>one ride up to 30 minutes on a Classic bike.</small> Get a pass	Day Pass \$15/day <small>unlimited 30-minute rides in a 24-hour period on a Classic bike.</small> Get a pass
--	---	---

More options

[Bike for Business](#) [Management Benefits](#) [NYCHA and SNAP](#) [Credit Union Members](#)

[Citi® Cardholders](#) [Fund](#) [JCHA and JC SNAP](#)

NYC Citi Bike



NYC Citi Bike

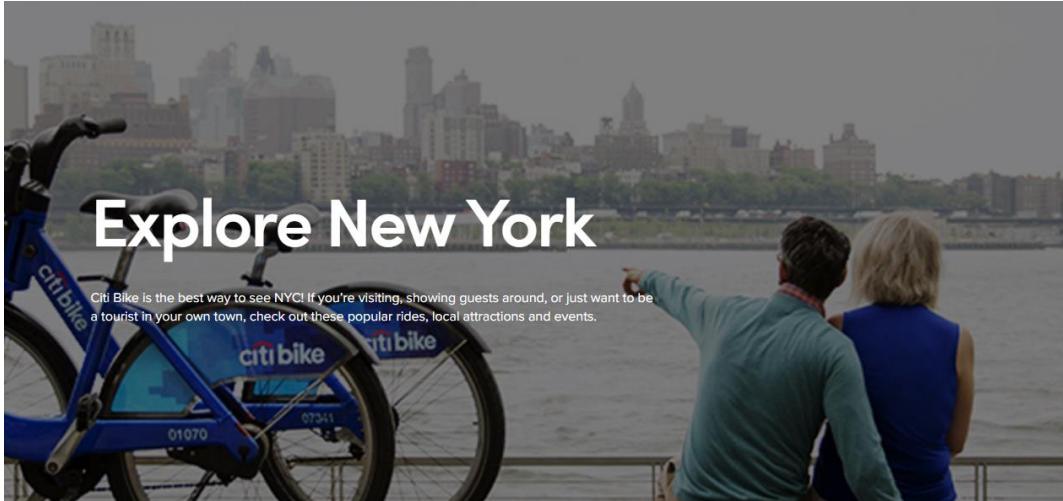


How it Works Pricing Find a Bike Explore NYC Reduced Fares

Log In

Join

For a limited time, save \$50 on annual membership e-gift certificates. [Shop now.](#)



Explore New York

Citi Bike is the best way to see NYC! If you're visiting, showing guests around, or just want to be a tourist in your own town, check out these popular rides, local attractions and events.

[citibike](#)

01070

07341

Popular Rides



Central Park Southern Loop

The Southern Loop of Central Park is one of the most popular bike rides in NYC. It's scenic, easy – and completely car-free!

[Learn more](#)

Hudson River Greenway

The greenway along the Hudson River is one of the most popular places to ride. It offers beautiful water views and a dedicated bike path that is closed to car traffic.

[Learn more](#)



Brooklyn Bridge Park

The waterfront along DUMBO and Brooklyn Heights has been revitalized into a park with a bike path offering breathtaking views of the Manhattan skyline.

[Learn more](#)

Local Attractions



The High Line

Take a stroll on Manhattan's gorgeous elevated park.

[Learn more](#)

NYC Citi Bike

citibike
provided by lyft

How it Works Pricing Find a Bike Explore NYC Help Log In **Join**

POPULAR RIDES LOCAL ATTRACTIONS EVENTS & CLASSES CENTRAL PARK BIKE RENTAL PROSPECT PARK HUDSON GREENWAY

EXPLORE ON A CITI BIKE

Central Park Southern Loop

The Southern Loop of Central Park is one of the most popular bike rides in NYC. It's scenic, easy – and completely car-free!



Start Here

Head to the Citi Bike station at [Central Park South & 6th Ave](#) and use your bike key or pre-purchased ride pass to begin. If you need to get a pass, use the [Citi Bike app](#) or kiosk at the station and follow the on-screen prompts to complete your purchase.

Remember, the first 30 minutes of each ride on a classic bike are included in the price of a [single ride](#) and [day pass](#). If you're an [annual member](#), the first 45 minutes of each ride on a classic bike are included. For more information about Citi Bike pricing and pass options, please head [here](#).

Once you have your bike, ride north into the park and take Center Drive a short distance to connect with Park Drive, the main road ("loop") that runs around Central Park.

Try this ride

Pick up a single ride or day pass through the mobile app and be on a bike within minutes!

[Buy a Pass](#)

No smartphone?

You can also purchase a day pass at any Citi Bike station kiosk.

[Learn more](#)

Nearby Stations

- Central Park S & 6 Ave
- Grand Army Plaza & Central Park S
- 5 Ave & E 63 St

[Find a bike](#)

Ride Here

Follow the few steps below and biking the Central Park Southern Loop will be a breeze!

- Follow Park Drive north up the East Side of the park
- Turn left onto the Terrace Drive transverse that crosses the park (this will be around 72nd St)
- When you reach the end of the transverse, turn left to pick up Park Drive again and travel south, down the West Side of the park
- Loop back around to Center Drive, where you can return your bike

The total distance is about 1.7 miles, and takes most people 15-30 minutes on a classic Citi Bike.

See the interactive map below and download [Central Park Bike Map](#) for more information.



See this

Located along this biking route are many great Central Park attractions worth checking out. These include:

- Bethesda Fountain
- The Carousel
- Strawberry Fields

NYC Citi Bike

News

NYC cycling surge: 530K daily bike trips represents 26% increase in recent years

Updated: Oct. 11, 2021, 11:13 a.m. | Published: Oct. 09, 2021, 4:56 p.m.



Bike ridership in New York City has been on a steady rise in recent years, with daily trips increasing 26% since 2014. (Reena Rose Sibayan | The Jersey Journal)



By Erik Bascome | tbascome@siadvance.com

STATEN ISLAND, N.Y. -- Cycling is on the rise in New York City as more and more residents ditch their cars in favor of a two-wheeled commute.

On Thursday, the Department of Transportation (DOT) released its annual [Cycling in the City](#) report, detailing the steady increase in daily cycling in New York City in recent years.

"Our fellow New Yorkers are seeing the light and embracing transportation on two wheels – helping us fight congestion and climate change at the same time," said DOT Commissioner Hank Gutman.

NEWS · WE THE COMMUTERS

Bike Boom Shows No Signs Of Slowing As Citi Bike Sets New Ridership Records

BY JAKE OFFENHARTZ

MAY 18, 9:16 PM • 149 COMMENTS



► Mayor de Blasio riding a Citi Bike last week NYC MAYOR FLICKR

Gorgeous spring weather and the growing pace of reopening has been a major boon for New York City's bike-share program, according to new data released by Citi Bike.

Cyclists took 631,314 rides on Citi Bike last week, smashing an all time record for the eight-year-old bike share program. Both Friday and Saturday saw consecutive single day highs, with 103,159 and 118,590 rides, respectively.

NYC Citi Bike

Citi Bike set to become world's biggest bike-share outside of China

By David Meyer

April 7, 2021 | 2:15pm | Updated



A Citi Bike is seen at a docking station
Getty Images

MORE ON: **CITIBIKE**

[Homeless pitch camp under Manhattan Bridge colonnade](#)

[Rerouting Fifth Ave traffic before the holidays is madness](#)

[Gang member on probation busted for allegedly shooting man on Citi Bike](#)

[Watch: Citi Bike explodes on subway tracks after being hit by oncoming train](#)

Citi Bike will expand into Upper Manhattan this month — making the bike-share system the largest in the world outside of those in China, the company said Wednesday.

The incursion into the Washington Heights and Inwood area will bring the total number of bikes in the system to 20,000, Citi Bike said — triple the number when bike-share started in 2013, and more than the sizable systems in Seoul and Paris.

By 2024, Citi Bike expects to have doubled in size again, surpassing systems in China to become the world's biggest municipal bike-share.

According to data compiled by Citi Bike-owner Lyft, the largest network in the world is currently in Suzhou, China, which has a 38,679-bike fleet that also operates in nearby Wujiang.

The two largest networks outside of China are currently in Seoul and Paris, Lyft said.

NBC NEWS

POLITICS COVID-19 U.S. NEWS WORLD OPINION BUSINESS

WATCH NOW

DATA GRAPHICS

Covid fueled the rise of the e-bike. See where ridership grew in the U.S.

The share of e-bike usage more than tripled from May 2020 to May 2021.



A bicyclist rides a Citi Bike bike-share e-bike through New York City's Midtown neighborhood. The system's e-bike usage has increased nine-fold from May 2020 to May 2021. Noam Galai / Getty Images file

New York City (NYC) Citi Bike Sharing

Data

NYC Citi Bike - System Data



How it Works Pricing Find a Bike Explore NYC Help

System Data

Where do Citi Bikers ride? When do they ride? How far do they go? Which stations are most popular? What days of the week are most rides taken on? We've heard all of these questions and more from you, and we're happy to provide the data to help you discover the answers to these questions and more. We invite developers, engineers, statisticians, artists, academics and other interested members of the public to use the data we provide for analysis, development, visualization and whatever else moves you.

This data is provided according to the [NYCBI Data Use Policy](#).

Citi Bike Trip Histories

We publish [downloadable files of Citi Bike trip data](#). The data includes:

- Ride ID
- Rideable type
- Started at
- Ended at
- Start station name
- Start station ID
- End station name
- End station ID
- Start latitude
- Start longitude
- End latitude
- End Longitude
- Member or casual ride

Data format previously:

- 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).

Please be aware of your software program's row limitations as you are viewing the data.

Many of the CSV files contain more than 1 million rows. After downloading, you will need to use a large data tool / visualizer (like Tableau, Alteryx, R, or others) to view and analyze the full data sets.

→ [Download Citi Bike trip history data](#)

NYC Citi Bike - System Data

Index of bucket "tripdata"

Name	Date Modified	Size	Type
201810-citibike-tripdata.csv.zip	Nov 7th 2018, 02:43:33 pm	79.98 MB	ZIP file
201811-citibike-tripdata.csv.zip	Dec 4th 2018, 11:27:25 am	53.73 MB	ZIP file
201812-citibike-tripdata.csv.zip	Jan 8th 2019, 03:51:08 pm	43.61 MB	ZIP file
201901-citibike-tripdata.csv.zip	Feb 11th 2019, 09:43:30 am	41.47 MB	ZIP file
201902-citibike-tripdata.csv.zip	Mar 4th 2019, 11:21:38 am	40.47 MB	ZIP file
201903-citibike-tripdata.csv.zip	Apr 15th 2019, 02:34:41 pm	57.20 MB	ZIP file
201904-citibike-tripdata.csv.zip	May 6th 2019, 03:07:32 pm	75.70 MB	ZIP file
201905-citibike-tripdata.csv.zip	Jun 11th 2019, 03:37:54 pm	82.18 MB	ZIP file
201906-citibike-tripdata.csv.zip	Jul 15th 2019, 01:42:08 pm	91.08 MB	ZIP file
201907-citibike-tripdata.csv.zip	Aug 5th 2019, 05:34:01 pm	93.36 MB	ZIP file
201908-citibike-tripdata.csv.zip	Sep 18th 2019, 04:33:14 pm	100.02 MB	ZIP file
201909-citibike-tripdata.csv.zip	Oct 11th 2019, 03:42:04 pm	104.19 MB	ZIP file
201910-citibike-tripdata.csv.zip	Nov 5th 2019, 05:10:55 pm	89.20 MB	ZIP file
201911-citibike-tripdata.csv.zip	Dec 20th 2019, 01:44:54 pm	63.20 MB	ZIP file
201912-citibike-tripdata.csv.zip	Jan 21st 2020, 04:23:32 pm	41.08 MB	ZIP file
202001-citibike-tripdata.csv.zip	Feb 4th 2020, 02:05:46 pm	53.47 MB	ZIP file
202002-citibike-tripdata.csv.zip	Mar 24th 2020, 04:07:44 pm	49.51 MB	ZIP file
202003-citibike-tripdata.csv.zip	Apr 17th 2020, 07:58:01 pm	46.66 MB	ZIP file
202004-citibike-tripdata.csv.zip	May 26th 2020, 08:18:14 pm	30.12 MB	ZIP file
202005-citibike-tripdata.csv.zip	Jun 5th 2020, 12:09:37 pm	65.38 MB	ZIP file
202006-citibike-tripdata.csv.zip	Jul 5th 2020, 08:37:43 pm	83.40 MB	ZIP file
202007-citibike-tripdata.csv.zip	Aug 11th 2020, 10:13:25 pm	94.09 MB	ZIP file
202008-citibike-tripdata.csv.zip	Sep 4th 2020, 03:37:28 pm	104.26 MB	ZIP file
202009-citibike-tripdata.csv.zip	Oct 13th 2020, 03:55:55 pm	111.93 MB	ZIP file
202010-citibike-tripdata.csv.zip	Nov 4th 2020, 09:51:10 am	101.84 MB	ZIP file
202011-citibike-tripdata.csv.zip	Dec 4th 2020, 06:26:03 pm	78.69 MB	ZIP file
202012-citibike-tripdata.csv.zip	Jan 5th 2021, 09:25:44 am	49.55 MB	ZIP file
202101-citibike-tripdata.csv.zip	Feb 8th 2021, 10:11:26 am	50.12 MB	ZIP file
202102-citibike-tripdata.csv.zip	Jun 11th 2021, 01:23:12 pm	21.18 MB	ZIP file

Downloaded to class OneDrive

OneDrive - Northeastern University 2 > damg6210_2021_03_fall > Data - CitiBike			
Name	Date modified	Type	Size
201511-citibike-tripdata.csv	12/2/2015 9:01 AM	Microsoft Excel Comma Separated Values File	189,054 KB
201810-citibike-tripdata.csv	11/7/2018 2:02 PM	Microsoft Excel Comma Separated Values File	349,235 KB
201912-citibike-tripdata.csv	1/21/2020 3:49 PM	Microsoft Excel Comma Separated Values File	176,609 KB
202012-citibike-tripdata.csv	1/5/2021 9:17 AM	Microsoft Excel Comma Separated Values File	201,917 KB
202101-citibike-tripdata.csv	2/8/2021 9:39 AM	Microsoft Excel Comma Separated Values File	203,119 KB
202102-citibike-tripdata.csv	6/11/2021 10:09 AM	Microsoft Excel Comma Separated Values File	113,762 KB
202106-citibike-tripdata.csv	7/15/2021 6:58 PM	Microsoft Excel Comma Separated Values File	568,910 KB
202110-citibike-tripdata.csv	11/4/2021 3:09 PM	Microsoft Excel Comma Separated Values File	561,847 KB
201511-citibike-tripdata.zip	12/1/2021 8:37 AM	Zip-File	33,550 KB
201810-citibike-tripdata.csv.zip	11/30/2021 10:25 AM	Zip-File	78,109 KB
201811-citibike-tripdata.csv.zip	11/30/2021 10:25 AM	Zip-File	52,473 KB
201812-citibike-tripdata.csv.zip	11/30/2021 10:25 AM	Zip-File	42,584 KB
201901-citibike-tripdata.csv.zip	11/30/2021 10:25 AM	Zip-File	40,496 KB
201912-citibike-tripdata.csv.zip	11/30/2021 10:08 AM	Zip-File	40,121 KB
202002-citibike-tripdata.csv.zip	11/30/2021 10:23 AM	Zip-File	48,354 KB
202003-citibike-tripdata.csv.zip	11/30/2021 10:23 AM	Zip-File	45,565 KB
202004-citibike-tripdata.csv.zip	11/30/2021 10:23 AM	Zip-File	29,415 KB
202005-citibike-tripdata.csv.zip	11/30/2021 10:23 AM	Zip-File	63,847 KB
202006-citibike-tripdata.csv.zip	11/30/2021 10:23 AM	Zip-File	81,445 KB
202007-citibike-tripdata.csv.zip	11/30/2021 10:23 AM	Zip-File	91,886 KB
202008-citibike-tripdata.csv.zip	11/30/2021 10:24 AM	Zip-File	101,814 KB
202009-citibike-tripdata.csv.zip	11/30/2021 10:24 AM	Zip-File	109,309 KB
202010-citibike-tripdata.csv.zip	11/30/2021 10:24 AM	Zip-File	99,457 KB
202011-citibike-tripdata.csv.zip	11/30/2021 10:24 AM	Zip-File	76,848 KB
202012-citibike-tripdata.csv.zip	11/30/2021 10:08 AM	Zip-File	48,387 KB
202101-citibike-tripdata.csv.zip	11/30/2021 10:13 AM	Zip-File	48,947 KB
202102-citibike-tripdata.csv.zip	11/30/2021 10:24 AM	Zip-File	20,687 KB
202103-citibike-tripdata.csv.zip	11/30/2021 10:24 AM	Zip-File	49,359 KB
202104-citibike-tripdata.csv.zip	11/30/2021 10:24 AM	Zip-File	63,261 KB
202105-citibike-tripdata.csv.zip	11/30/2021 10:24 AM	Zip-File	77,525 KB
202106-citibike-tripdata.csv.zip	11/30/2021 10:25 AM	Zip-File	111,846 KB
202107-citibike-tripdata.csv.zip	11/30/2021 10:25 AM	Zip-File	95,064 KB
202108-citibike-tripdata.csv.zip	11/30/2021 10:25 AM	Zip-File	103,900 KB
202109-citibike-tripdata.csv.zip	11/30/2021 10:25 AM	Zip-File	109,424 KB
202110-citibike-tripdata.csv.zip	11/30/2021 9:47 AM	Zip-File	100,319 KB

bigquery-public-data:new_york_citibike.citibike_trips

Type to search

Viewing pinned projects.

- new_york_citibike
 - citibike_stations
 - citibike_stations-2021-11-01
 - citibike_trips**
 - citibike_trips-2021-11-01
- new_york_mv_collisions
- new_york_subway

citibike_trips

SCHEMA DETAILS PREVIEW

Table info

Table ID	bigquery-public-data:new_york_citibike.citibike_trips
Table size	7.47 GB
Long-term storage size	7.47 GB
Number of rows	58,937,715
Created	Apr 12, 2017, 3:40:35 PM UTC-4
Last modified	Sep 10, 2018, 9:29:21 PM UTC-4
Table expiration	NEVER
Data location	US
Description	

Data from 2013-06-01 until 2018-05-31
And not being updated

WARNING

Updated: 2021-12-07

citibike_trips

SCHEMA DETAILS PREVIEW

Table schema

Filter Enter property name or value

Field name	Type	Mode	Policy Tags	Description
tripduration	INTEGER	NULLABLE		Trip Duration (in seconds)
starttime	DATETIME	NULLABLE		Start Time, in NYC local time.
stoptime	DATETIME	NULLABLE		Stop Time, in NYC local time.
start_station_id	INTEGER	NULLABLE		Start Station ID
start_station_name	STRING	NULLABLE		Start Station Name
start_station_latitude	FLOAT	NULLABLE		Start Station Latitude
start_station_longitude	FLOAT	NULLABLE		Start Station Longitude
end_station_id	INTEGER	NULLABLE		End Station ID
end_station_name	STRING	NULLABLE		End Station Name
end_station_latitude	FLOAT	NULLABLE		End Station Latitude
end_station_longitude	FLOAT	NULLABLE		End Station Longitude
bikeid	INTEGER	NULLABLE		Bike ID
usertype	STRING	NULLABLE		User Type (Customer = 24-hour pass or 7-day pass user, Subscriber = Annual Member)
birth_year	INTEGER	NULLABLE		Year of Birth
gender	STRING	NULLABLE		Gender (unknown, male, female)
customer_plan	STRING	NULLABLE		The name of the plan that determines the rate charged for the trip

EDIT SCHEMA VIEW ROW ACCESS POLICIES

bigquery-public-data:new_york_citibike.citibike_stations

Type to search

Viewing pinned projects.

- new_york_citibike
 - citibike_stations
 - citibike_stations-2021-...
 - citibike_trips
 - citibike_trips-2021-11-...
- new_york_mv_collisions
- new_york_subway
- new_york_taxi_trips
- new_york_trips

citibike_stations

SCHEMA DETAILS PREVIEW

Table info

Table ID	bigquery-public-data:new_york_citibike.citibike_stations
Table size	205.37 KB
Long-term storage size	0 B
Number of rows	1,584
Created	Apr 12, 2017, 3:35:34 PM UTC-4
Last modified	Nov 9, 2021, 3:01:06 AM UTC-5
Table expiration	NEVER
Data location	US
Description	

Data from 2013-06-01, updated monthly and is being updated

citibike_stations

SCHEMA DETAILS PREVIEW

Table schema

Filter Enter property name or value

Field name	Type	Mode	Policy Tags	Description
station_id	INTEGER	REQUIRED		Unique identifier of a station.
name	STRING	NULLABLE		Public name of the station.
short_name	STRING	NULLABLE		Short name or other type of identifier, as used by the data publisher.
latitude	FLOAT	NULLABLE		The latitude of station. The field value must be a valid WGS 84 latitude in decimal degrees format.
longitude	FLOAT	NULLABLE		The longitude of station. The field value must be a valid WGS 84 longitude in decimal degrees format.
region_id	INTEGER	NULLABLE		ID of the region where station is located.
rental_methods	STRING	NULLABLE		Array of enumerables containing the payment methods accepted at this station.
capacity	INTEGER	NULLABLE		A number of total docking points installed at this station, both available and unavailable.
eighthd_has_key_dispenser	BOOLEAN	NULLABLE		
num_bikes_available	INTEGER	NULLABLE		Number of bikes available for rental.
num_bikes_disabled	INTEGER	NULLABLE		Number of disabled bikes at the station.
num_docks_available	INTEGER	NULLABLE		Number of docks accepting bike returns.
num_docks_disabled	INTEGER	NULLABLE		Number of empty but disabled dock points at the station.
is_installed	BOOLEAN	NULLABLE		Is the station currently on the street?
is_renting	BOOLEAN	NULLABLE		Is the station currently renting bikes?
is_returning	BOOLEAN	NULLABLE		Is the station accepting bike returns?
eighthd_has_available_keys	BOOLEAN	NULLABLE		
last_reported	DATETIME	NULLABLE		Timestamp indicating the last time this station reported its status to the backend, in NYC local time.

EDIT SCHEMA VIEW ROW ACCESS POLICIES

Sample Summary Data



How it Works Pricing Find a Bike Explore NYC Help

Citi Bike Monthly Operating Reports

Lyft, the operator of the Citi Bike program, provides these monthly reports to the NYC Department of Transportation. We've included them here with the intention of being open and transparent with our customers and the residents of the City of New York.

2013 Operating Reports

- [June 2013](#)
- [July 2013](#)
- [August 2013](#)
- [September 2013](#)
- [October 2013](#)
- [November 2013](#)
- [December 2013](#)

2014 Operating Reports

- [January 2014](#)
- [February 2014](#)

A screenshot of a file explorer window titled 'damg6210_2021_03_fall > Data - CitiBike > Summary_Files'. It shows two files: 'Citi_bike_daily_trips.csv' and 'citi_bikes_stations_EndOfMonth.csv'. An orange arrow points from the '2013 Operating Reports' section on the left towards the 'citi_bikes_stations_EndOfMonth.csv' file.

Name	Date modified	Type	Size
Citi_bike_daily_trips.csv	12/1/2021 8:56 AM	Microsoft Excel C...	109 KB
citi_bikes_stations_EndOfMonth.csv	12/1/2021 8:54 AM	Microsoft Excel C...	3 KB

YearMonth	Stations	Avg_active_bikes	EndOfMonth_fleet	Membership
2013-06	322	5130	6000	52130
2013-07	328	5531	6000	66315
2013-08	332	5681	6000	77138
2013-09	332	5513	6000	85241
2013-10	332	5623	6000	93124
2013-11	332	5900	6000	94955
2013-12	332	6499	6000	96125
2014-01	332	5441	6000	97138
2014-02	329	4200	6000	97864
2014-03	327	5028	6000	100600
2014-04	324	5731	5808	105367
2014-05	327	5149	5276	105359
2014-06	325	5362	5276	96318
2014-07	323	5411	5367	93501
2014-08	325	5079	5066	93184
2014-09	325	5043	5351	90879
2014-10	328	5279	5461	89286
2014-11	325	4995	4592	88495
2014-12	325	4077	3716	88405

Sample Summary Data

Name	Date modified	Type	Size
Citi_bike_daily_trips.csv	12/1/2021 8:56 AM	Microsoft Excel C...	109 KB
citi_bikes_stations_EndOfMonth.csv	12/1/2021 8:54 AM	Microsoft Excel C...	3 KB

SK	date	count	Year	Month	Week	Day	Day of Week	Day of Year
0	6/1/2013	8722	2013	6	22	1	5	152
1	6/2/2013	15971	2013	6	22	2	6	153
2	6/3/2013	7598	2013	6	23	3	0	154
3	6/4/2013	15782	2013	6	23	4	1	155
4	6/5/2013	15690	2013	6	23	5	2	156
5	6/6/2013	12420	2013	6	23	6	3	157
6	6/7/2013	1226	2013	6	23	7	4	158
7	6/8/2013	18007	2013	6	23	8	5	159
8	6/9/2013	21034	2013	6	23	9	6	160
9	6/10/2013	3897	2013	6	24	10	0	161
10	6/11/2013	12596	2013	6	24	11	1	162
11	6/12/2013	16795	2013	6	24	12	2	163
12	6/13/2013	8016	2013	6	24	13	3	164
13	6/14/2013	14942	2013	6	24	14	4	165
14	6/15/2013	25178	2013	6	24	15	5	166
15	6/16/2013	21966	2013	6	24	16	6	167
16	6/17/2013	19659	2013	6	25	17	0	168
17	6/18/2013	13724	2013	6	25	18	1	169
18	6/19/2013	24142	2013	6	25	19	2	170

New York City (NYC) Citi Bike Sharing Assignment

Data Sources for Analysis

1. Use Big Query Public datasets
 - citibike_stations
 - citibike_trips - trip data 6/1/2013-5/31/2018
2. Load Big Query tables in a dataset in your project folder
 - citibike_stations
 - citibike_trips - trip data 1/1/2014-12/31/2020 (NOT 2021)
3. Load Data into Oracle ADW, Azure SQL or MySQL (**your choice**)
 - citibike_stations
 - citibike_trips - trip data 1/1/2019-12/31/2020

WARNING

Updated: 2021-12-07

Options for Loading Data for Analysis

- Load Big Query tables in a dataset in your project folder
 - 2014-01-01 to 2018-05-31
 - Query BigQuery and save results into your project table
 - Data from 2018-06-01 to 2020-12-31, use one of the following
 - Use Tableau Data Prep (**this is the preferred method**)
 - Upload data into Google Storage and then into tables
 - DBeaver import
- Load Data into Oracle ADW, Azure SQL or MySQL for 2019 and 2020.
 - **Use one of the following:**
 - Tableau Data Prep
 - DBeaver import
 - Import csv files (there may be limitations)

*Hint: use **Timestamp** versus **Datetime** for starttime and stoptime*



Updated: 2021-12-07

Answering questions and performing analysis

- Perform analysis and answer questions for each data set (1, 2, 3)
 - 1: 25% of grade
 - 2: 50% of grade
 - 3: 25% of grade
- Upload
 - Table row counts
 - SQL & query result sets
 - Questions asked
 - Additional analysis
 - Any data visualizations with Google Data Studio are for bonus and are NOT required

New York City (NYC) Citi Bike Sharing Assignment

NYC Citi Bike Share: Questions to answer

- The number of bike-sharing trips is a typical example of a time series. It has the following components:
 - Trend: overall year-over-year growth
 - Seasonality: users make more trips in the summer months and less in the winter.
 - Cycles: weekly cycle where more trips happen on weekdays and less on weekends
- NOTE:
 - All the following questions need to be based on the data in the query set
 1. 2013-06-01 to 2018-05-31 (fyi these are fiscal years) - Big Query Public Datasets
 2. 2014-01-01 to 2020-12-31 - Big Query Your Project Datasets
 3. 2019-01-01 to 2020-12-31 – Oracle ADW, Azur SQI or MySQL (your choice)



Updated: 2021-12-08

NYC Citi Bike Share: Questions to answer

- How many bike trips and what is the trend over time
 - calendar year (2017)
 - calendar month & year (2017-09)
 - calendar week (2017-36)
 - calendar day (2017-09-01)
- Average number of bike trips and may also want to compare YOY (year over year) changes
 - per day, i.e., Sunday, Monday,...Saturday
 - weekday versus weekend
 - month
 - calendar week

NYC Citi Bike Share: Questions to answer

Examining when trips occur during the day

- Breakdown of subscriber (or member) vs customer (non-subscriber) rides
- How many bike trips occur during what periods of the day:
 - By hour (0900, 1300)
 - During periods during the day i.e., morning 6am-11:59am, afternoon noon-5:59pm, evening 6pm-10pm
- How many round trips (start and end station are the same) versus non-round trips (different start and end station)
 - Further breakdown subscriber vs customer
 - Average length of trip

NYC Citi Bike Share: Questions to answer

Examining length of trips

- Length of trips (in time)
 - What is the average length (in minutes) of bike trip?
 - Further breakdown by subscriber versus customer
 - How many trips are less than 15 minutes in duration?
 - How many trips are greater than 15 minutes but less than 1 hour in duration?
 - How many trips are greater than 1 hour in duration?

Note:

May help identify commuters, occasional users (getting around town) and tourists

NYC Citi Bike Share: Questions to answer

Most popular stations

- What are the top 10 bike stations based on bike trips?
 - For starting a trip
 - As a trip destination
 - Overall, as either a start and/or destination
- Determine the most popular routes (starting to end location)
 - Total number of trips
 - Total number of trips by subscriber and customer
 - Further breakdowns (examples): time of day, weekday vs weekend, month

NYC Citi Bike Share: Questions to answer

- **Gender analysis**
 - How many trips - female vs male
 - Average length of trip - female vs male
 - Subscriber versus customer trips - female vs male
 - Any trends from 2014 to 2020 regarding above
- **Age analysis** (note: age at the time of the ride i.e., age is year of bike trip vs birth year)
 - What are the number of trips per age
 - Average length of trip per age
 - Subscriber versus customer trips – per age
 - Any trends from 2014 to 2020 regarding above

NYC Citi Bike Share: Questions to answer

Bike Analysis

- What are the top 5 bikes (bike id) for latest (full) year by
 - Total trips
 - Total time
 - Average trip time
- For top bike (by total trips) create a history for the latest (full) year
 - Bike id
 - Route (state and end location)
 - Start time
 - Start location (id, longitude and latitude)
 - Duration



Updated: 2021-12-08

New York City (NYC) Citi Bike Sharing Appendix

Moving data

- Cloud to Cloud
 - Same Cloud Service (GCP, Azure, AWS, Snowflake, Oracle) and Same “region”
 - Same Cloud Service but different “region” – speed & ability to create view
 - Different Cloud Services & Different Cloud Applications i.e., Salesforce, Marketo
- Files to/from Cloud Service, database, application
- On-premise (your company network or your local database) to/from Cloud
- Note:
 - Cloud databases, storage and applications are in physical data centers
 - An enterprise's databases and applications
 - An enterprise's physical data centers in their network
 - On one or more Cloud Services

Considerations

- Data Types
 - datetime, timestamp
 - Big Query string, MySQL text
- Database Functions such as datetime, timestamp, cast
 - Vary between databases
- Trying to do it all at once

Considerations: Data Types

ISO SQL Data Types

Data Type	Declarations							
boolean	BOOLEAN							
character	CHAR	VARCHAR						
bit [†]	BIT	BIT VARYING						
exact numeric	NUMERIC	DECIMAL	INTEGER	SMALLINT	BIGINT			
approximate numeric	FLOAT	REAL	DOUBLE PRECISION					
datetime	DATE	TIME	TIMESTAMP					
interval	INTERVAL							
large objects	CHARACTER LARGE OBJECT	BINARY LARGE OBJECT						

- [MySQL Data Types](#)
- [Oracle Data Types](#)
- [Big Query Data Types](#)

Considerations: Database Functions

Google Cloud

Why Google Solutions Products Pricing Getting Started



BigQuery Overview Guides Reference Samples Support Resources

https://cloud.google.com/bigquery/docs/reference/standard-sql/date_functions

BigQuery command-line tool
bq command-line tool reference

BigQuery SQL

Quickstart

Standard SQL reference

Concepts

Functions and operators

Function call rules

Conversion rules

AEAD encryption functions

Aggregate analytic functions

Aggregate functions

Approximate aggregate functions

Array functions

Bit functions

Conversion functions

Date functions

Datetime functions

Debugging functions

Federated query functions

Geography functions

Hash functions

HyperLogLog++ functions

Interval functions

JSON functions

Mathematical functions

Navigation functions

Net functions

Numbering functions

Security functions

Statistical aggregate functions

String functions

Time functions

Timestamp functions

UUID functions

Operators

Conditional expressions

Expression subqueries

One-page function reference

Statements

Legacy SQL reference

BigQuery audit logging

BigQuery audit logging reference

BigQuery > Documentation > Reference

Was this helpful?

[Send feedback](#)

Date functions

BigQuery supports the following DATE functions.

CURRENT_DATE

[CURRENT_DATE\(\[time_zone\]\)](#)

On this page

- [CURRENT_DATE](#)
- [EXTRACT](#)
- [DATE](#)
- [DATE_ADD](#)
- [DATE_SUB](#)
- [DATE_DIFF](#)
- [DATE_TRUNC](#)
- [DATE_FROM_UNIX_DATE](#)
- [FORMAT_DATE](#)
- [LAST_DAY](#)
- [PARSE_DATE](#)
- [UNIX_DATE](#)
- [Supported format elements for DATE](#)

Description

Returns the current date as of the specified or default timezone. Parentheses are optional when called with no arguments.

This function supports an optional time_zone parameter. This parameter is a string representing the timezone to use. If no timezone is specified, the default timezone, UTC, is used. See [Timezone definitions](#) for information on how to specify a time zone.

If the time_zone parameter evaluates to NULL, this function returns NULL.

Return Data Type

DATE

Example

[SELECT CURRENT_DATE\(\) AS the_date;](#)

```
+-----+  
| the_date |  
+-----+  
| 2016-12-25 |  
+-----+
```

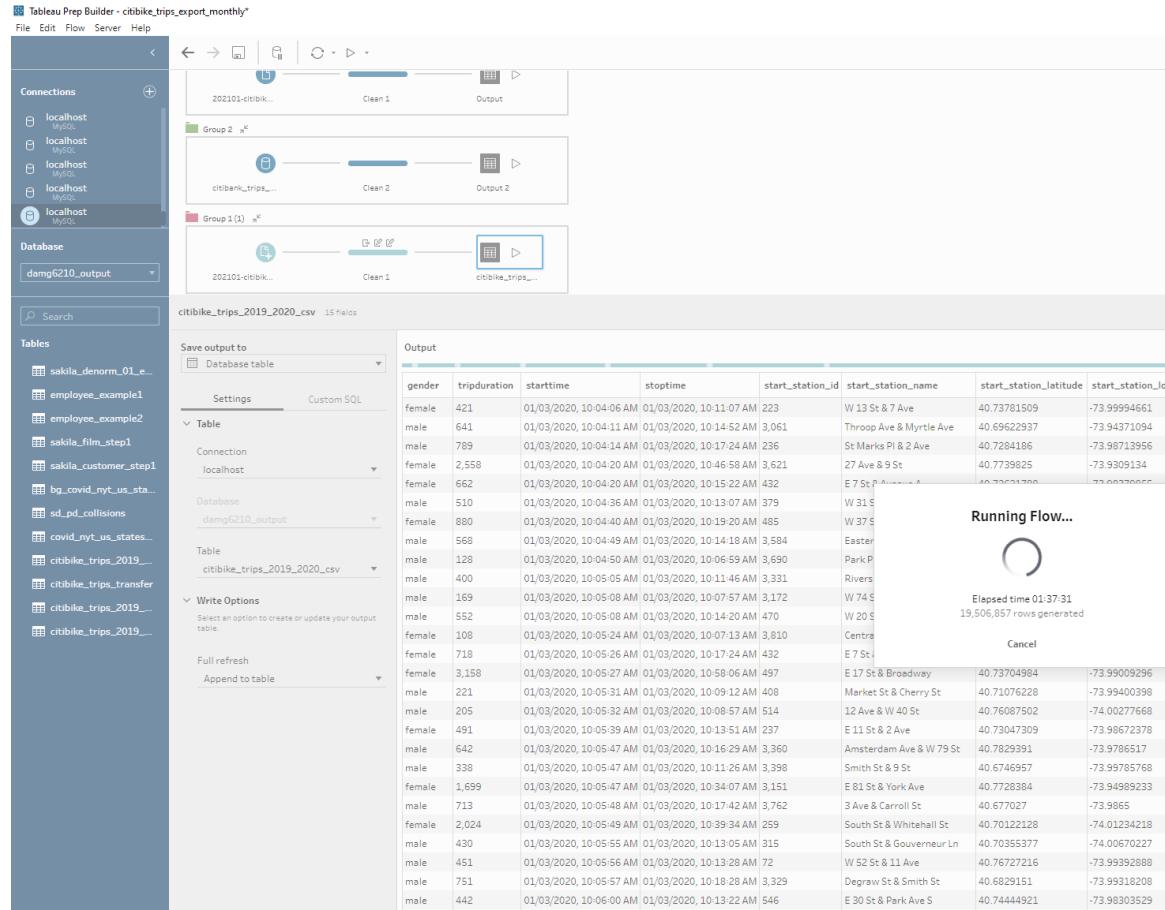
When a column named current_date is present, the column name and the function call without parentheses are ambiguous. To ensure the function call, add parentheses; to ensure the column name, qualify it with its range variable. For example, the following query will select the function in the the_date column and the table column in the current_date column.

[WITH t AS \(SELECT 'column value' AS 'current_date'\)
SELECT current_date\(\) AS the_date, t.current_date FROM t;](#)

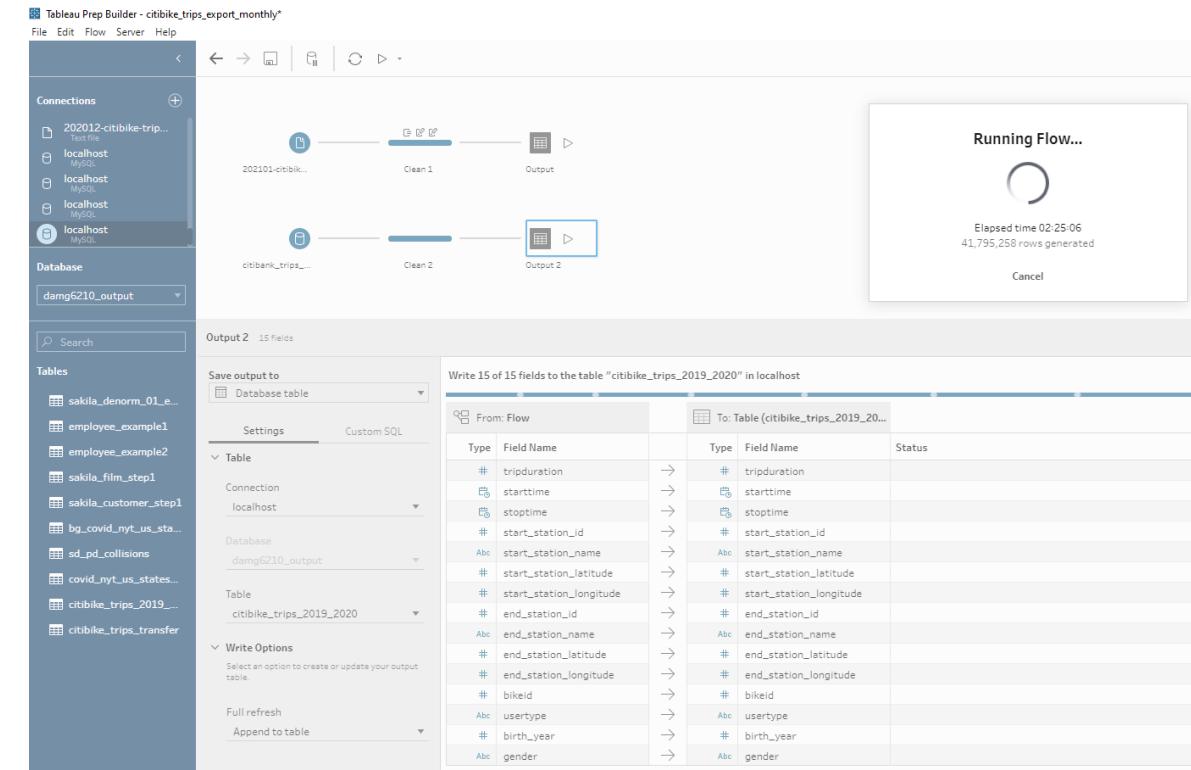
```
+-----+-----+  
| the_date | current_date |  
+-----+-----+
```

Considerations: Trying to do it all at once

Trying to load 2020 data from multiple files to MySQL



Trying to load 2019-2020 data from Big Query to MySQL



Considerations: Trying to do it all at once

Loading one file (month) at a time (202012) to MySQL

Tableau Prep Builder - citibike_trips_export_monthly*

File Edit Flow Server Help

Connections

- localhost MySQL

Database

- damg6210_output

Search

Tables

- sakila_denorm_01_e...
- employee_exampl...
- employee_exampl...
- sakila_film_step1
- sakila_customer_step1
- bg_covid_nyt_us_st...
- sd_pd_collisions
- covid_nyt_us_st...
- citibike_trips_2019...
- citibike_trips_2019...
- citibike_trips_2019...
- citibike_trips_2019...
- citibike_trips_2019...

citibike_trips_2019_2020_csv 15 Fields

Save output to Database table

Output

gender	tripduration	starttime	stoptime	start_station_id	start_station_name	start_station_latitude	start_station_longitude
unknown	1,298	12/02/2020, 10:41:32 AM	12/02/2020, 11:03:11 AM	3,671	E 81 St & 2 Ave	40.77477944895728	-73.95427465438843
male	147	12/02/2020, 10:41:33 AM	12/02/2020, 10:44:00 AM	239	Willoughby St & Fleet St	40.69196566	-73.9813018
female	139	12/02/2020, 10:41:34 AM	12/02/2020, 10:43:53 AM	432	E 7 St & Avenue A	40.72621788	-73.98379855
male	543	12/02/2020, 10:41:35 AM	12/02/2020, 10:50:39 AM	509	9 Ave & W 22 St	40.7454973	-74.00197139
male	219	12/02/2020, 10:41:37 AM	12/02/2020, 10:45:16 AM	3,259	9 Ave & W 20 St	40.740320345000003	-73.980032041000006
female	1,091	12/02/2020, 10:41:40 AM	12/02/2020, 10:59:51 AM	462	W 22 St		
male	711	12/02/2020, 10:41:41 AM	12/02/2020, 10:53:33 AM	3,949	Brooklyn		
female	442	12/02/2020, 10:41:43 AM	12/02/2020, 10:49:05 AM	3,641	Broadway		
female	737	12/02/2020, 10:41:44 AM	12/02/2020, 10:54:01 AM	3,410	Dean St		
male	562	12/02/2020, 10:41:44 AM	12/02/2020, 10:51:06 AM	497	E 17 St		
male	195	12/02/2020, 10:41:44 AM	12/02/2020, 10:45:00 AM	505	6 Ave		
male	131	12/02/2020, 10:41:47 AM	12/02/2020, 10:43:58 AM	301	E 2 St		
male	94	12/02/2020, 10:41:47 AM	12/02/2020, 10:43:21 AM	3,315	Henry Street		
female	1,164	12/02/2020, 10:41:48 AM	12/02/2020, 11:01:12 AM	3,175	W 70 St		
female	583	12/02/2020, 10:41:48 AM	12/02/2020, 10:51:32 AM	3,490	E 116 St & 2 Ave	40.796879	-73.937261
male	278	12/02/2020, 10:41:49 AM	12/02/2020, 10:46:27 AM	3,521	Lenox Ave & W 111 St	40.798789	-73.9523
male	1,886	12/02/2020, 10:41:52 AM	12/02/2020, 11:13:19 AM	487	E 20 St & FDR Drive	40.73314259	-73.97573881
male	1,114	12/02/2020, 10:41:54 AM	12/02/2020, 11:00:28 AM	3,998	W 135 St & Lenox Terrace Pl	40.813619	-73.939336
male	774	12/02/2020, 10:41:54 AM	12/02/2020, 10:54:49 AM	3,137	S 6 Ave & 72 St	40.77282817	-73.9685276
male	248	12/02/2020, 10:41:57 AM	12/02/2020, 10:46:06 AM	3,102	Driggs Ave & Lorimer St	40.72179134	-73.9504154
male	275	12/02/2020, 10:41:58 AM	12/02/2020, 10:46:33 AM	3,498	Pleasant Ave & E 120 St	40.7974772	-73.9311847
female	643	12/02/2020, 10:41:58 AM	12/02/2020, 10:52:41 AM	143	Clinton St & Joralemon St	40.69239502	-73.99337909
male	179	12/02/2020, 10:42:00 AM	12/02/2020, 10:45:00 AM	228	E 48 St & 3 Ave	40.75460110265	-73.971878855
male	189	12/02/2020, 10:42:03 AM	12/02/2020, 10:45:12 AM	349	Rivington St & Ridge St	40.71850211	-73.98329859
male	1,089	12/02/2020, 10:42:04 AM	12/02/2020, 11:00:14 AM	418	Front St & Gold St	40.70224	-73.982578
male	134	12/02/2020, 10:42:04 AM	12/02/2020, 10:44:19 AM	223	W 13 St & 7 Ave	40.73781509	-73.99994661
female	535	12/02/2020, 10:42:07 AM	12/02/2020, 10:51:02 AM	3,926	5 Ave & W 131 St	40.81014	-73.93973
male	280	12/02/2020, 10:42:09 AM	12/02/2020, 10:46:50 AM	4,110	Avenue D & E 10 St	40.72506551015475	-73.97515833777838
female	2,092	12/02/2020, 10:42:10 AM	12/02/2020, 11:17:03 AM	3,734	E 58 St & 1 Ave (NW Corner)	40.759125	-73.962658
male	1,259	12/02/2020, 10:42:11 AM	12/02/2020, 11:03:11 AM	2,021	W 45 St & 8 Ave	40.75929124	-73.98859651
male	471	12/02/2020, 10:42:11 AM	12/02/2020, 10:50:03 AM	151	Cleveland Pl & Spring St	40.72210378668603	-73.99724900722504
female	899	12/02/2020, 10:42:14 AM	12/02/2020, 10:57:14 AM	3,690	Park Pl & Church St	40.71334184275907	-74.00935515761375
female	1,082	12/02/2020, 10:42:16 AM	12/02/2020, 11:00:18 AM	3,497	Pleasant Ave & E 116 St	40.7949879	-73.9333349
male	703	12/02/2020, 10:42:16 AM	12/02/2020, 10:53:59 AM	3,282	5 Ave & 88 St	40.78307	-73.95939
male	941	12/02/2020, 10:42:16 AM	12/02/2020, 10:57:58 AM	3,918	Avenue D & E 8 St	40.72387	-73.975767
male	1,070	12/02/2020, 10:42:17 AM	12/02/2020, 11:00:07 AM	3,486	Schermerhorn St & Bond St	40.68841742754083	-73.98451656103134
male	444	12/02/2020, 10:42:20 AM	12/02/2020, 10:49:45 AM	3,707	Lexington Ave & E 26 St	40.741459	-73.983293

Run Flow

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

- adventureworks2017
- adventureworks2019
- adventureworksdw2017
- chinook
- cif
- crm
- customer
- damg6210_input
- damg6210_output
- Tables
- bg_covid_nyt_us_states
- citibike_trips_2019_2020
- citibike_trips_2019_2020_int_csv
- citibike_trips_2019_2020_int_cv
- citibike_trips_transfer
- covid_nyt_us_states_damg6210
- employee_exampl...
- employee_exampl...
- sakila_customer_step1
- sakila_film_step1
- sd_pd_collisions

1 • SELECT count(*), min(starttime) min_trip, max(starttime) max_trip FROM damg6210_output.citibike_trips_2019_2020_cv

Result Grid

count(*)	min_trip	max_trip
1088929	2020-12-01 00:00:04	2020-12-31 23:59:57