

NYC Citibike Trips Data Analysis in 2018

ALI ROHMAN
• • • •

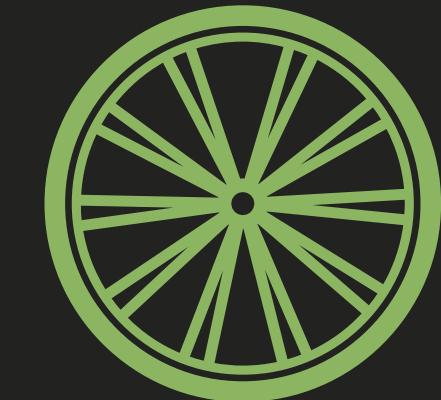
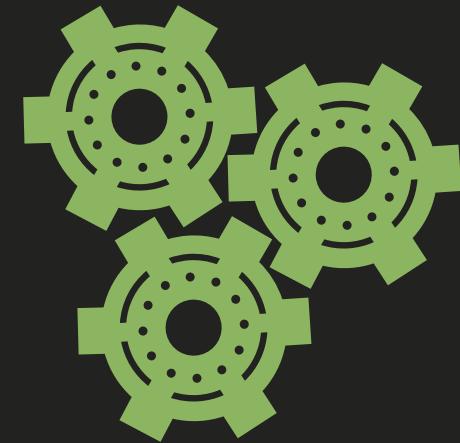


NYC Citibike Trips

Citi Bike is New York City's bike-sharing system with thousands of bikes at hundreds of stations and is available 24/7 every day of the year. Citi Bike is a convenient solution for quick trips around the City. Users could unlock a bike at any station and ride wherever they want.

Dataset Description

The dataset that is used comes from bigquery public data with table ID "bigquery-public-data:new_york_citibike.citibike_trips". The dataset contains information about NYC Citibike trips from July 1, 2013, to May 31, 2018. The dataset consists of 58,937,715 rows and 16 columns.



Problem

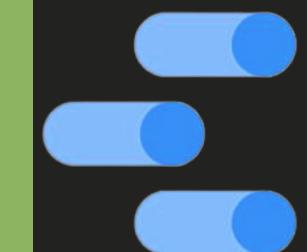
1. Who was the largest customer group in 2018?
2. What was the total, min, max and average trip duration (in minutes) in 2018?
3. What is the average duration of using Citibike per month in 2018?
4. What stations are the most popular in 2018?
5. What is the overall trend of NYC Citibike travel data per year from 2013 to 2018?

Tools

1. Google BigQuery (for dataset and SQL Query)
2. Google Data Studio



Google
Big Query



Data Studio

Who is our largest group of customers in 2018?

.....

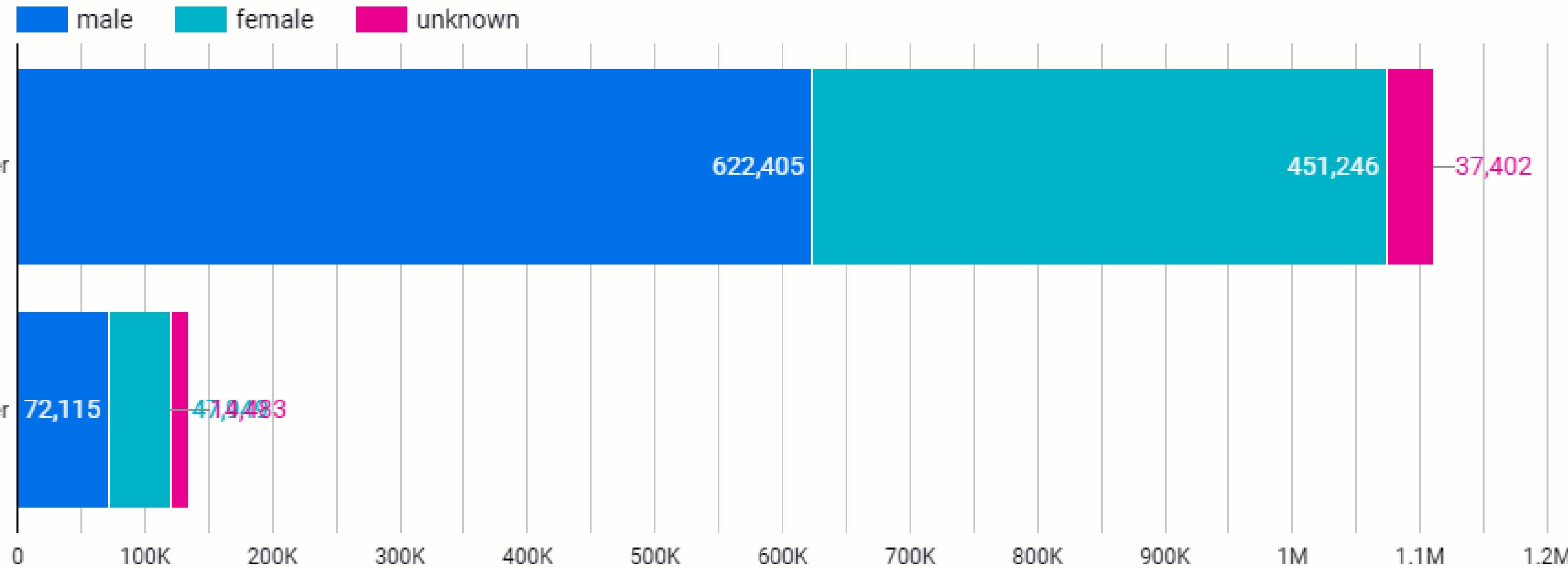
SQL QUERY

```
SELECT
    DISTINCT (usertype),
    birth_year,
    gender,
    COUNT (DISTINCT (bikeid)) AS unique_users,
    COUNT (bikeid) AS total_trip
FROM bigquery-public-data.new_york_citibike.citibike_trips`
WHERE
    DATE(starttime) >= '2018-01-01'
    AND DATE(stoptime) <= '2018-12-31'
GROUP BY 1, 2, 3
ORDER BY unique_users ASC
```



VISUALIZATION

| | usertype | gender | unique_users | total_trip |
|----|------------|---------|--------------|------------|
| 1. | Subscriber | male | 622.4K | 3.9M |
| 2. | Subscriber | female | 451.2K | 1.2M |
| 3. | Customer | male | 72.1K | 91.2K |
| 4. | Customer | female | 47.9K | 59.6K |
| 5. | Subscriber | unknown | 37.4K | 98.7K |
| 6. | Customer | unknown | 14.5K | 355.6K |



Insight

- From the table of query results, we can see that the largest NYC Citibike users are Users with the Subscriber type (Annual Member), with a total of 622.5K users with a total trip of 3.9M trips.
- Users with Subscriber type have a much larger percentage than Users with Customer type, which means that of all Citibike users in New York in 2018, the majority have become customers (Annual Members).
- It can also be seen that there are more male users than female, whether for subscribers of the Subscriber type or not.

Total, Min, Max, and Average Trips Duration (in minutes) in 2018



SQL QUERY

```
SELECT  
    ROUND(SUM(tripduration) / 60,3) as total_duration_minutes,  
    ROUND(MIN(tripduration) / 60,3) as min_duration_minutes,  
    ROUND(MAX(tripduration) / 60,3) as max_duration_minutes,  
    ROUND(AVG(tripduration) / 60,3) as avg_duration_minutes  
  
FROM  
    bigquery-public-data.new_york_citibike.citibike_trips`  
  
WHERE  
    EXTRACT(year FROM starttime) = 2018
```



VISUALIZATION

| total_duration_minutes | max_duration_minutes | min_duration_minutes | avg_duration_minutes |
|------------------------|----------------------|----------------------|----------------------|
| 1. 91871798.617 | 325167.483 | 1.017 | 16.2 |

Insight

- The table of query results shows the total duration time, largest duration, smallest duration, and the average duration of NYC Citibike trips during the year 2018 presented in minutes.
- It shows that the total duration of NYC Citibike trips is 91871798.617 minutes, the largest (longest) usage time is 325167.483 minutes, the smallest usage time is 1.017 minutes, and the average usage time of NYC Citibike usage is 16.2 minutes.

Average Usage Duration per Month in 2018



SQL QUERY

```
SELECT
    EXTRACT(month FROM starttime) AS month,
    ROUND(AVG(tripduration) / 60,3) as avg_duration_minutes
FROM
    bigquery-public-data.new_york_citibike.citibike_trips`
WHERE
    EXTRACT(year FROM starttime) = 2018
GROUP BY
    month
HAVING
    month is NOT NULL
ORDER BY
    month
```



VISUALIZATION

| month | avg_duration_minutes |
|-------------|----------------------|
| 1. January | 13.78 |
| 2. February | 14.62 |
| 3. March | 14.7 |
| 4. April | 16.18 |
| 5. May | 18.7 |



Insight

- During 2018, NYC Citibike usage data was recorded from January to May 2018, starting in January having an average trip duration of 13.78 minutes, February was 14.62 minutes, March was 14.7 minutes, April was 16.18 minutes, and May was 18.7 minutes which is the largest average trip duration for NYC Citibike usage recorded in 2018.
- From the graph it can be seen that there is an increase in average trip duration, this could be because people are less interested in cycling or fewer people are traveling in winter and when entering spring the average trip duration has increased significantly,

The Most Popular Stations in 2018

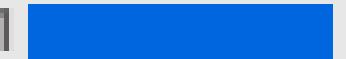


QUERY

```
SELECT  
    start_station_name,  
    start_station_latitude,  
    start_station_longitude,  
    COUNT(*) as total_num_trips  
FROM  
    bigquery-public-data.new_york.citibike_trips`  
GROUP BY  
    1,2,3  
ORDER BY  
    total_num_trips DESC  
LIMIT 10
```



VISUALIZATION

| | start_station_name | start_station_longitude | start_station_latitude | total_num_trips |
|-----|-----------------------|-------------------------|------------------------|---|
| 1. | E 17 St & Broadway | -73.99009296 | 40.73704984 | 291,615  |
| 2. | Lafayette St & E 8 St | -73.9907647 | 40.73028666 | 277,060  |
| 3. | W 21 St & 6 Ave | -73.99415556 | 40.74173969 | 275,348  |
| 4. | West St & Chambers St | -74.01322069 | 40.71754834 | 260,911  |
| 5. | Pershing Square North | -73.977706 | 40.751873 | 246,181  |
| 6. | Broadway & E 14 St | -73.99074142 | 40.73454567 | 244,420  |
| 7. | Broadway & E 22 St | -73.98955109 | 40.7403432 | 237,394  |
| 8. | 8 Ave & W 33 St | -73.993934 | 40.751551 | 223,970  |
| 9. | W 20 St & 11 Ave | -74.007756 | 40.746745 | 211,840  |
| 10. | Broadway & W 24 St | -73.98915076 | 40.7423543 | 211,713  |

Insight

From the 10 most popular stations visited by NYC Citibike users, the most popular for starting trips used by users is E 17 St & Broadway station.

Overall NYC Citibike trips usage data from 2013 - 2018

.....

QUERY

SELECT

```
EXTRACT(year FROM starttime) as year,  
COUNT(*) as num_trips,  
ROUND(SUM(tripduration) / 60,3) as total_duration_minutes
```

FROM

```
bigquery-public-data.new_york_citibike.citibike_trips`
```

GROUP BY year

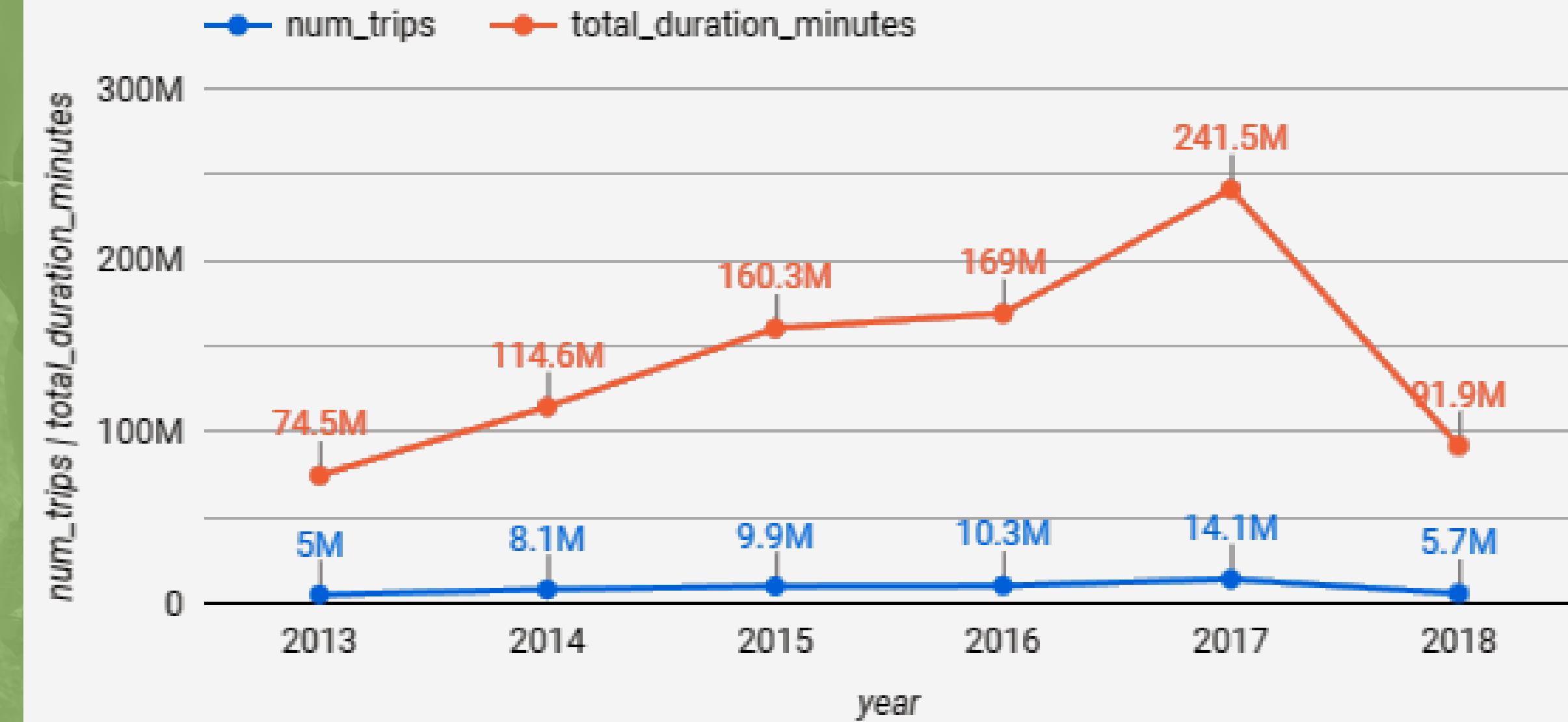
ORDER BY year



| | year | num_trips | total_duration_minutes |
|----|------|------------|------------------------|
| 1. | null | 5,828,994 | null |
| 2. | 2013 | 5,037,185 | 74,519,981.52 |
| 3. | 2014 | 8,081,216 | 114,642,178.33 |
| 4. | 2015 | 9,937,969 | 160,347,084.95 |
| 5. | 2016 | 10,262,649 | 169,017,191.72 |
| 6. | 2017 | 14,118,669 | 241,546,185.13 |
| 7. | 2018 | 5,671,033 | 91,871,798.62 |



VISUALIZATION



Insight

- Overall NYC Citibike trips usage data from July 1, 2013, to May 31, 2018 berdasarkan jumlah trips dan jumlah duration (per menit) penggunaan NYC Citibike, terlihat dari grafik bahwa jumlah trips dan jumlah duration selalu mengalami kenaikan tiap tahunnya, dengan nilai tertinggi yang tercatat yaitu pada Tahun 2017, dengan jumlah trips selama tahun tersebut adalah 14118669 trips dan jumlah duration (per menit) adalah 241546185.133 menit.
- Pada tahun 2018, terlihat bahwa terjadi penurunan pemakaian NYC Citibike kembali dibanding tahun 2017, hal ini dapat dikarenakan data yang tercatat pada tahun 2018 hanya tercatat dari bulan January - May, hanya 5 bulan yang tercatat, jadi hal ini sangat wajar bila .