Citibike Analysis Jan - May 2019

A static map that plots all bike stations with a visual indication of the most popular locations to start and end a journey with zip code data overlaid on top.

start station latitude and start station longitude

end station latitude and end station longitude

Quan SHUANG

Citibike Analysis Jan - May 2019

A dynamic map that shows how each station's popularity changes over time (by month and year) -- with commentary pointing to any interesting events that may be behind these phenomena.

Count of starttime and start station id by start station latitude and start station longitude

starttime

1/2/2019 5/31/2019

NYC OpenData, State of New Jersey, Esri, HERE, Garmin, METI/...

Quan SHUANG

Citibike Analysis Jan - May 2019

A chronic over-achiever: Find at least two unexpected phenomena in the data and provide a visualization and analysis to document their presence.

1. Too many 61 seconds trips...
2. Too many 100 year old plus riders... could be possible

start station name	tripduration	Count of tripduration
1 Ave & E 110 St	61	1
1 Ave & E 16 St	61	16
1 Ave & E 18 St	61	6
1 Ave & E 30 St	61	5
1 Ave & E 44 St	61	1
1 Ave & E 62 St	61	1
1 Ave & E 68 St	61	2
11 Ave & W 27 St	61	1
11 Ave & W 41 St	61	1
11 Ave & W 59 St	61	3
11 St & 35 Ave	61	1
12 Ave & W 40 St	61	1
2 Ave & E 31 St	61	2
2 Ave & E 96 St	61	2
21 St & 43 Ave	61	2
21 St & Queens Plaza North	61	1
24 Ave & 26 St	61	1
24 St & 41 Ave	61	1
3 St & 7 Ave	61	1
31 Ave & 30 St	61	1
Total		**7060686**

age	Count of tripduration
162	6
156	2
134	62
133	125
132	128
131	397
130	103
129	298
126	10
125	18
124	90
123	220
122	17
120	95
119	1611
118	294
115	13
114	1
113	2
112	27
Total	**7060686**

Quan SHUANG

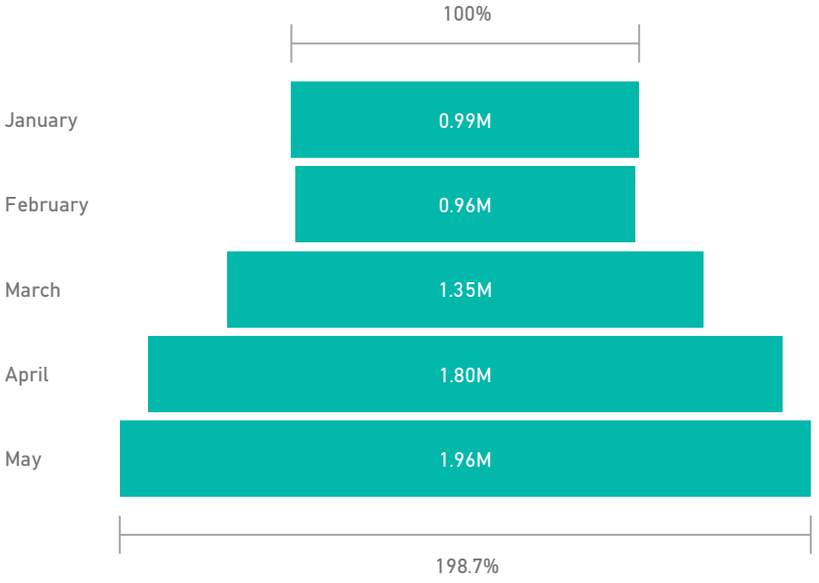
Citibike Analysis Jan - May 2019

How many trips have been recorded total during the chosen period?
7,060,686 Trips Jan-May 2019.

Month	Count of starttime
January	986963
February	962309
March	1351566
April	1799150
May	1960698
Total	**7060686**

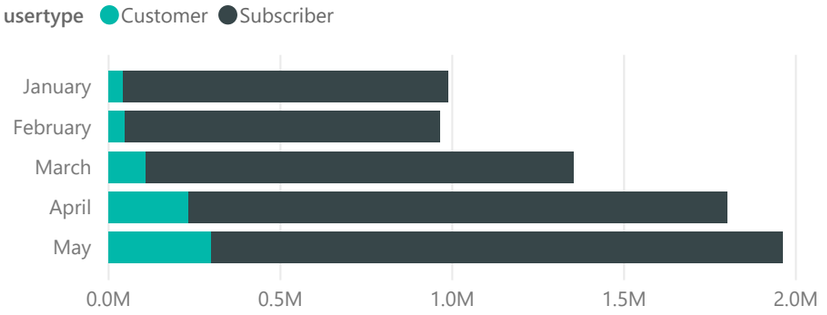
By what percentage has total ridership grown?
198.7%

Count of starttime by Month



How has the proportion of short-term customers and annual subscribers changed?
Customer : 735.1%
Subscriber: 175.7%

Count of usertype and %GT Count of usertype by Month and usertype



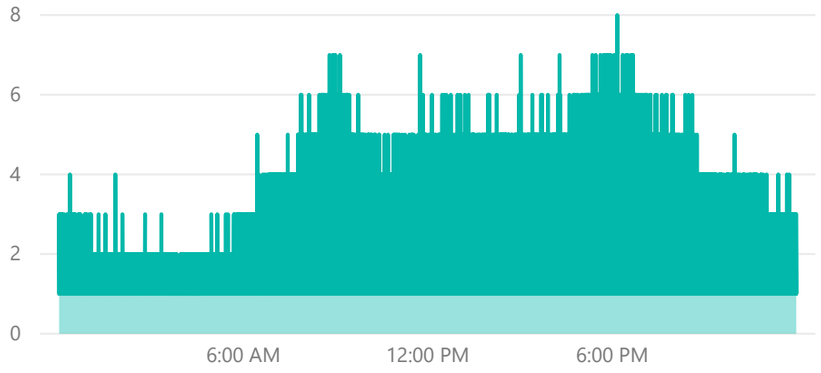
usertype
- Customer
- Subscriber

Quan SHUANG

Citibike Analysis Jan - May 2019

What are the peak hours in which bikes are used during summer months? May: 8AM to 8PM
What are the peak hours in which bikes are used during winter months? Jan: 8-9AM and 5:30 to 7PM

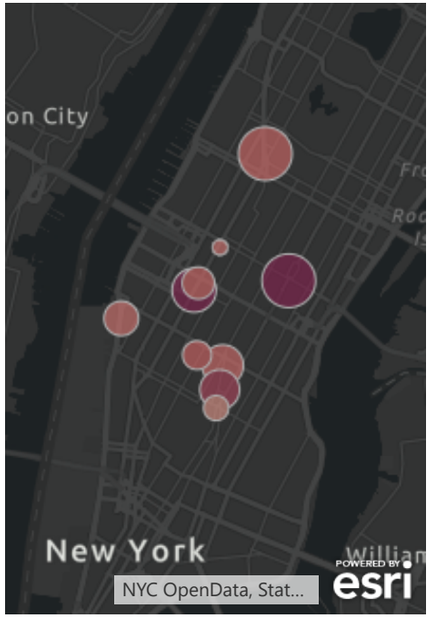
Count of Time and Count of Time by Time



1/1/2019 5/31/2019

Today, what are the top 10 stations in the city for starting a journey? (Based on data, why do you hypothesize these are the top locations?)
As shown on below map, based on starting station count.

tripduration and start station id by start...



tripduration by end station name, e...

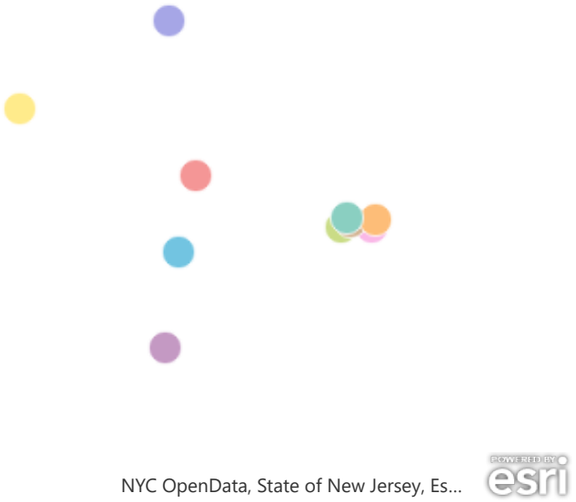
Today, what are the top 10 stations in the city for ending a journey? (Based on data, why?)
As shown on above map, based on starting station count.

Quan SHUANG

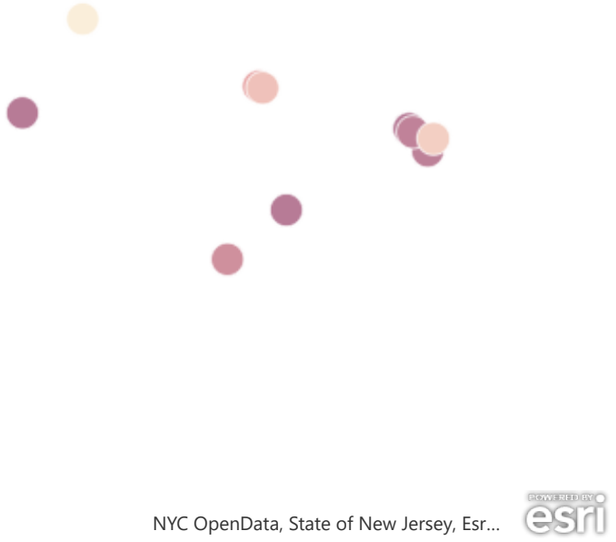
Citibike Analysis Jan - May 2019

Today, what are the bottom 10 stations in the city for starting a journey? (Based on data, why?)
Today, what are the bottom 10 stations in the city for ending a journey (Based on data, why?)
see below map

end station id, end station latitude, end station longitude and end station ...



start station id by start station name, start station latitude and start station l...



Quan SHUANG

Citibike Analysis Jan - May 2019

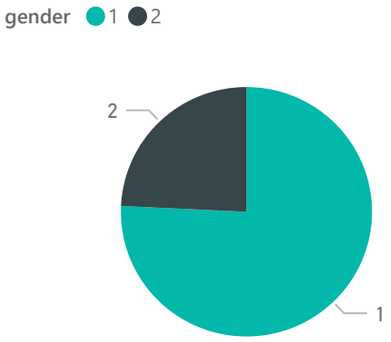
Today, what is the gender breakdown of active participants (Male v. Female)?

Male: 1, Female: 2

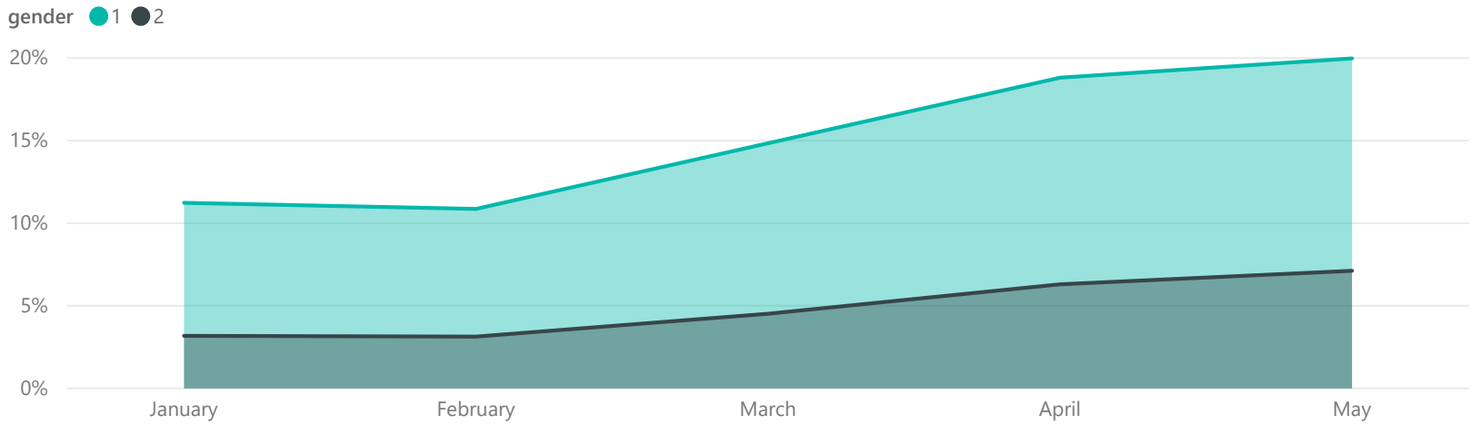
How effective has gender outreach been in increasing female ridership over the time span?

Female riders has increased from 22.07% to 26.28% between two genders.

Count of gender by gender



%GT Count of gender by Month and gender



gender	%GT Count of gender
1	75.74%
2	24.26%
Total	**100.00%**

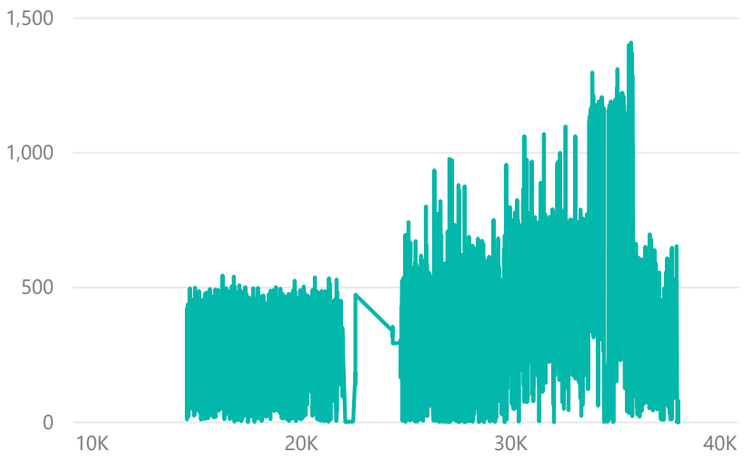
Month	1	2	**Total**
January	77.93%	22.07%	**100.00%**
February	77.60%	22.40%	**100.00%**
March	76.66%	23.34%	**100.00%**
April	74.90%	25.10%	**100.00%**
May	73.72%	26.28%	**100.00%**
Total	**75.74%**	**24.26%**	**100.00%**

Quan SHUANG

Citibike Analysis Jan - May 2019

What is the average distance in miles that a bike is ridden? 420 miles

Count of distance average per bikeid by bikeid



Calculation formula for distance using lat and log.

dlon = lon2 - lon1

dlat = lat2 - lat1

a = (sin(dlat/2))^2 + cos(lat1) * cos(lat2) * (sin(dlon/2))^2

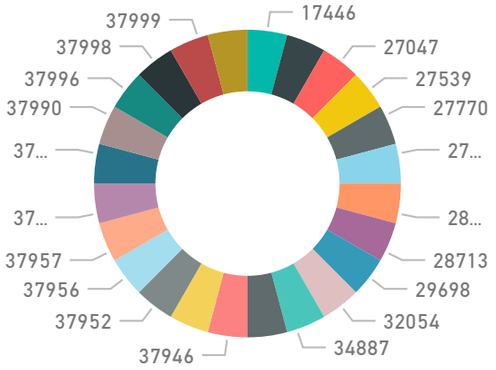
c = 2 * atan2(sqrt(a), sqrt(1-a))

d = R * c (where R is the radius of the Earth)

Which bikes (by ID) are most likely due for repair or inspection in the time span?

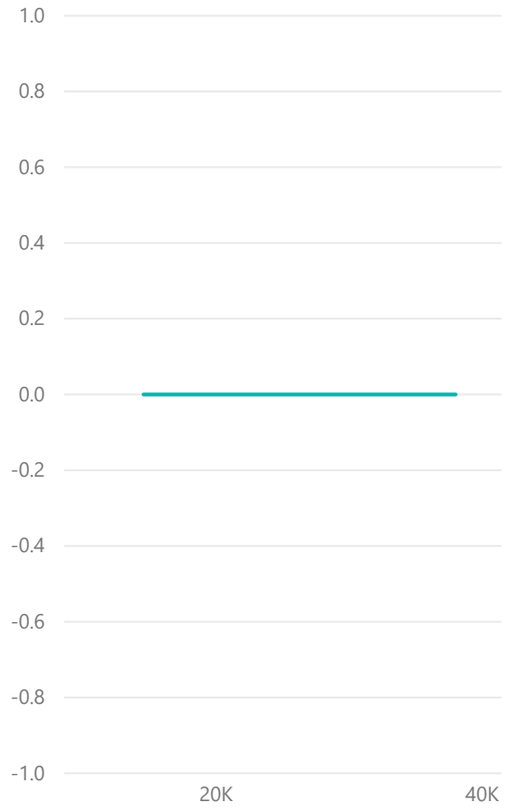
Below listed bike ids which had very low usage counts. So most likely they are in repair or inspection.

Count of distance by bikeid



How variable is the utilization by bike ID?

Count of distance variance per bikeid 2 by bike...



Quan SHUANG