

Citi Bike NYC Trip Segmentation and Station Network Analysis

Team Citi Slickers

Jessica Rudd

Shashank Hebbar

Soujanya Mandalapu



Citi Bike is the nation's largest bike share program, with 10,000 bikes and 600 stations across Manhattan, Brooklyn, Queens and Jersey City. It was designed for quick trips with convenience in mind, and it's a fun and affordable way to get around town.

Day Pass

\$12

unlimited 30-minute rides in
a 24-hour period

[Get a Pass](#)

3-Day Pass

\$24

unlimited 30-minute rides in
a 72-hour period

[Get a Pass](#)

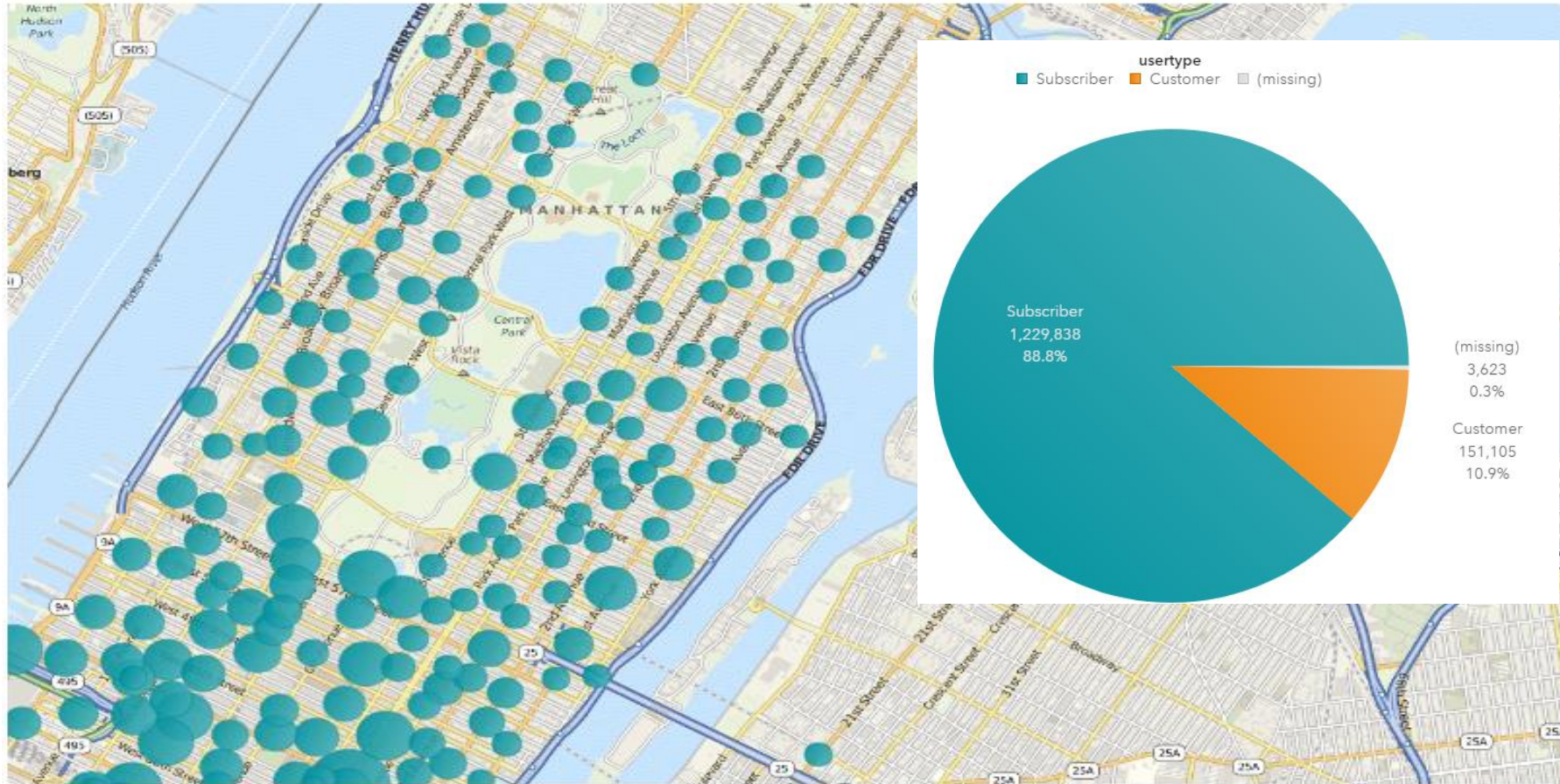
Annual Membership

\$163/year

unlimited 45-minute rides

[Join Now](#)

→ Or \$14.95/mo, annual commitment



- Above visualization is a map that shows in which places these bike share companies located and bubbles that you see are the stations where these bikes available.
- Of all of our total users, 89% of the users are Subscribers and rest 11% were Customers.
- Subscribers are typically the working population that resides in these areas where as customers are likely to be tourists.

So

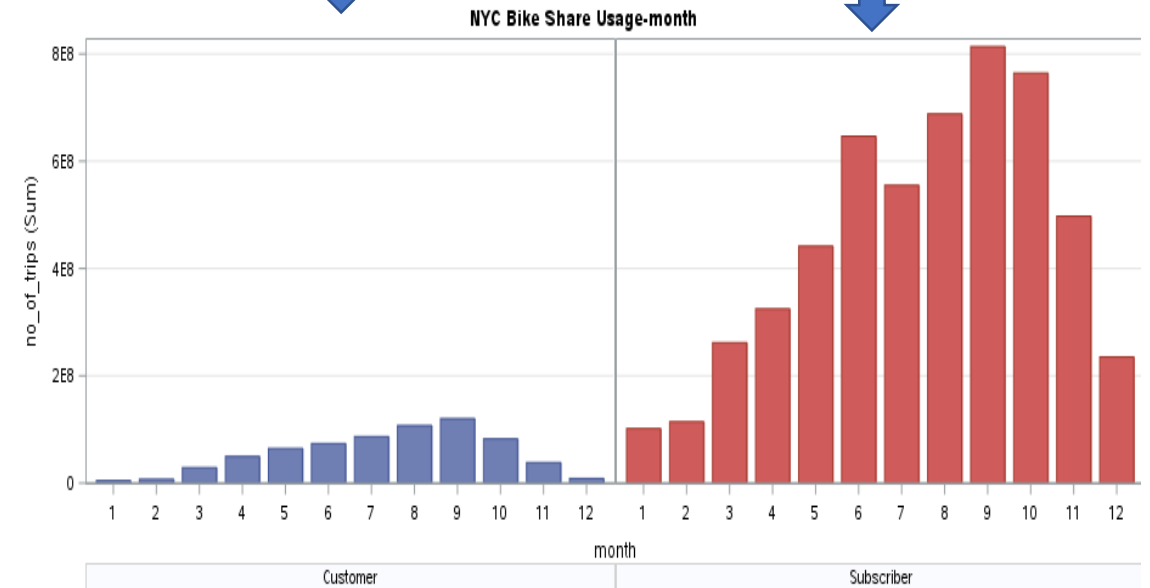
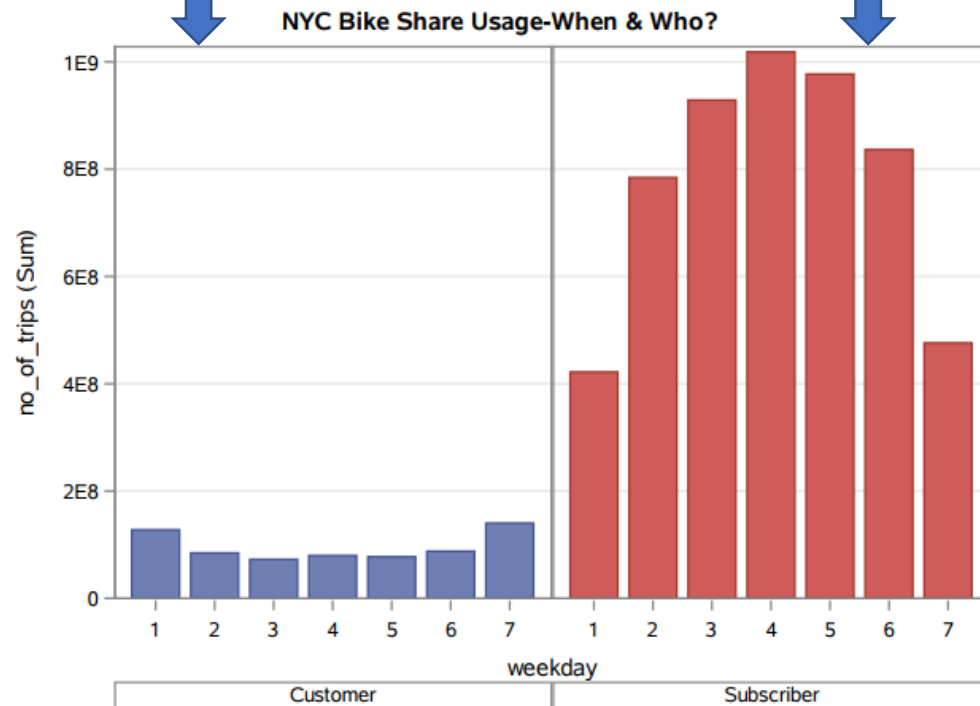
- Where do city Bikers ride?
- When do they ride?
- How far do they go?
- Which Stations are most popular?
- What days of the week are most riders take on?

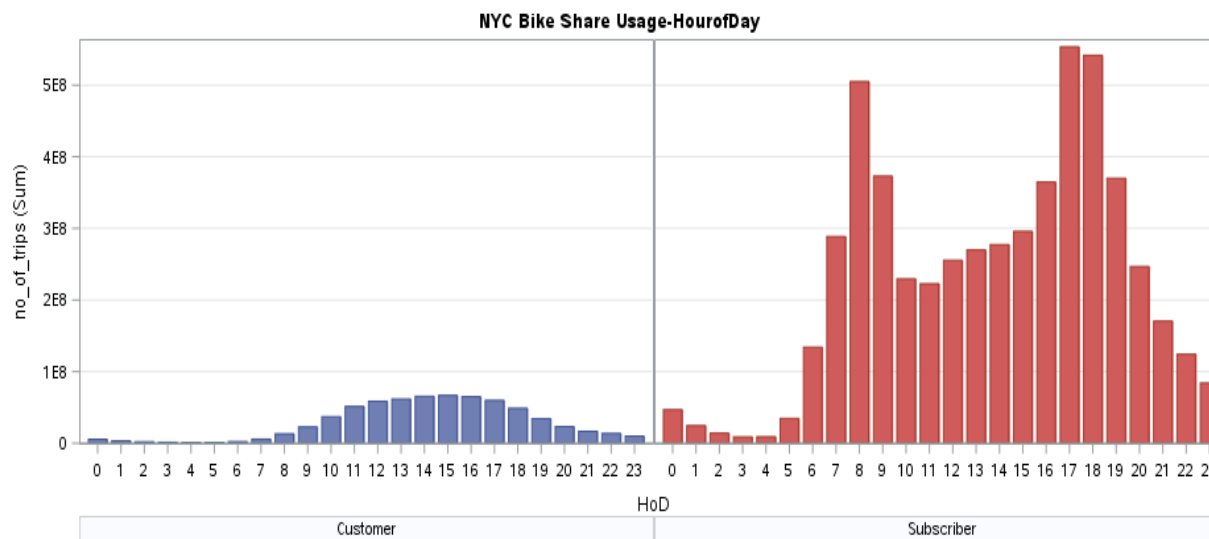
Customers seem riding more on weekends when compared to weekdays

Subscribers seem riding more on weekdays when compared to weekends

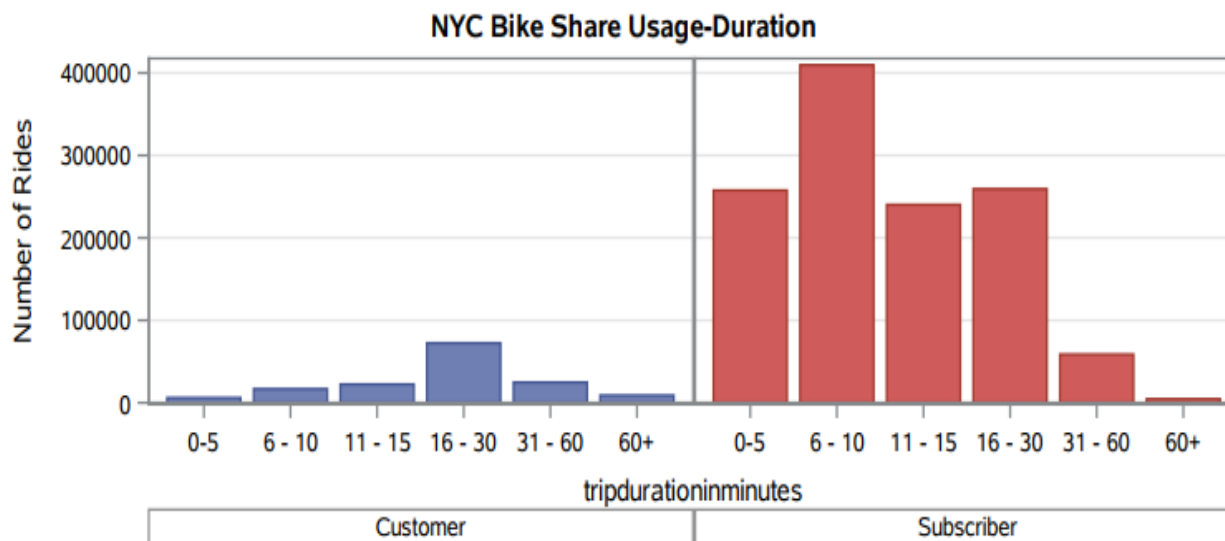
Customers seem riding more on summer and fall than winter

Subscribers have similar pattern as of customers but notice there was a dip in rides in the month of July





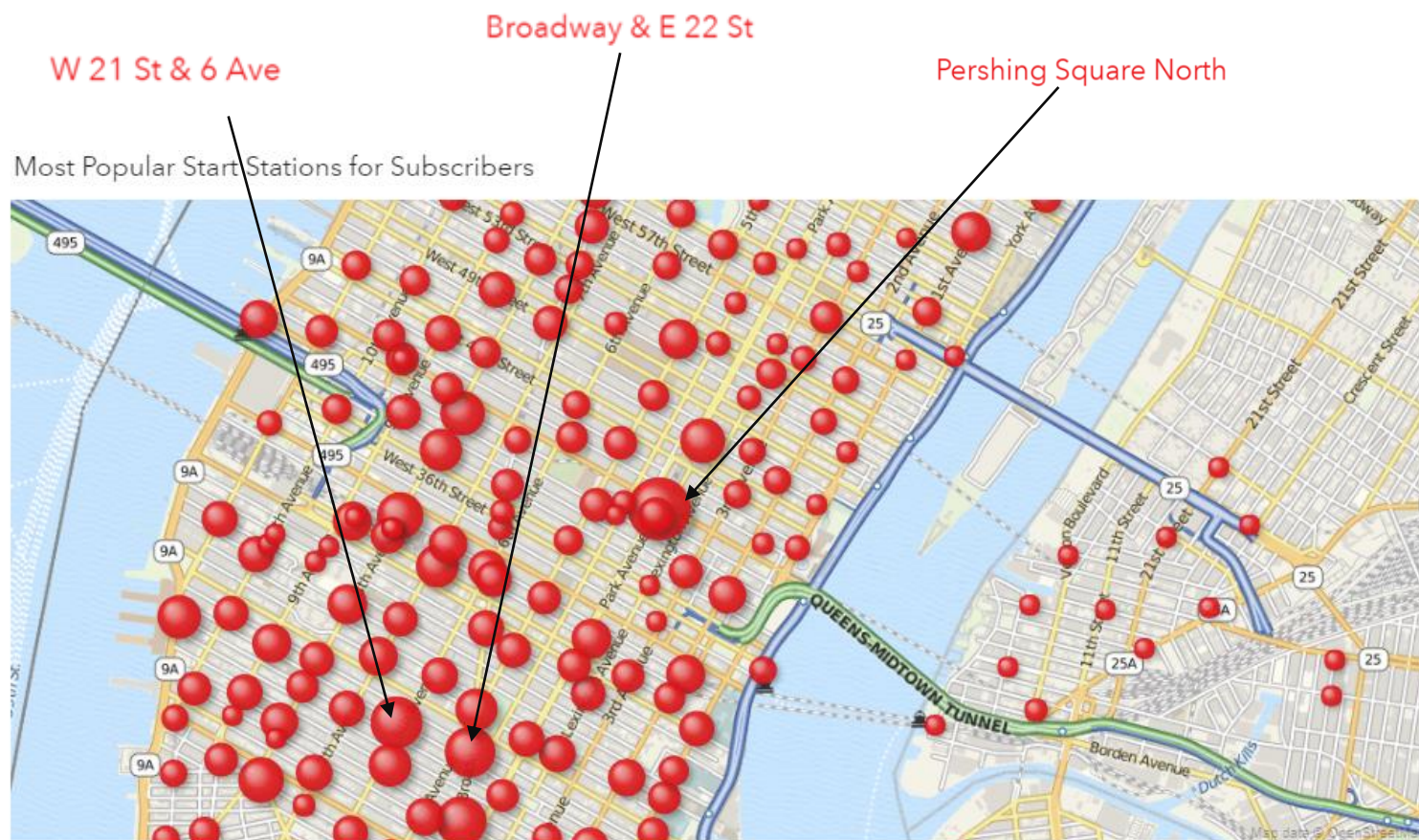
Subscribers take more rides during morning and evening rush hours but customers take more rides as the day progresses which peaks in the afternoon



More number of customers rent bike for longer duration (16-30 mins) where as subscribers rent bike typically for shorter duration (6-10 mins)

Geospatial representation of popular stations using SAS Visual Analytics

Most Popular Start Stations for Subscribers



Note that most popular stations for Subscribers are around Grand Central Terminal

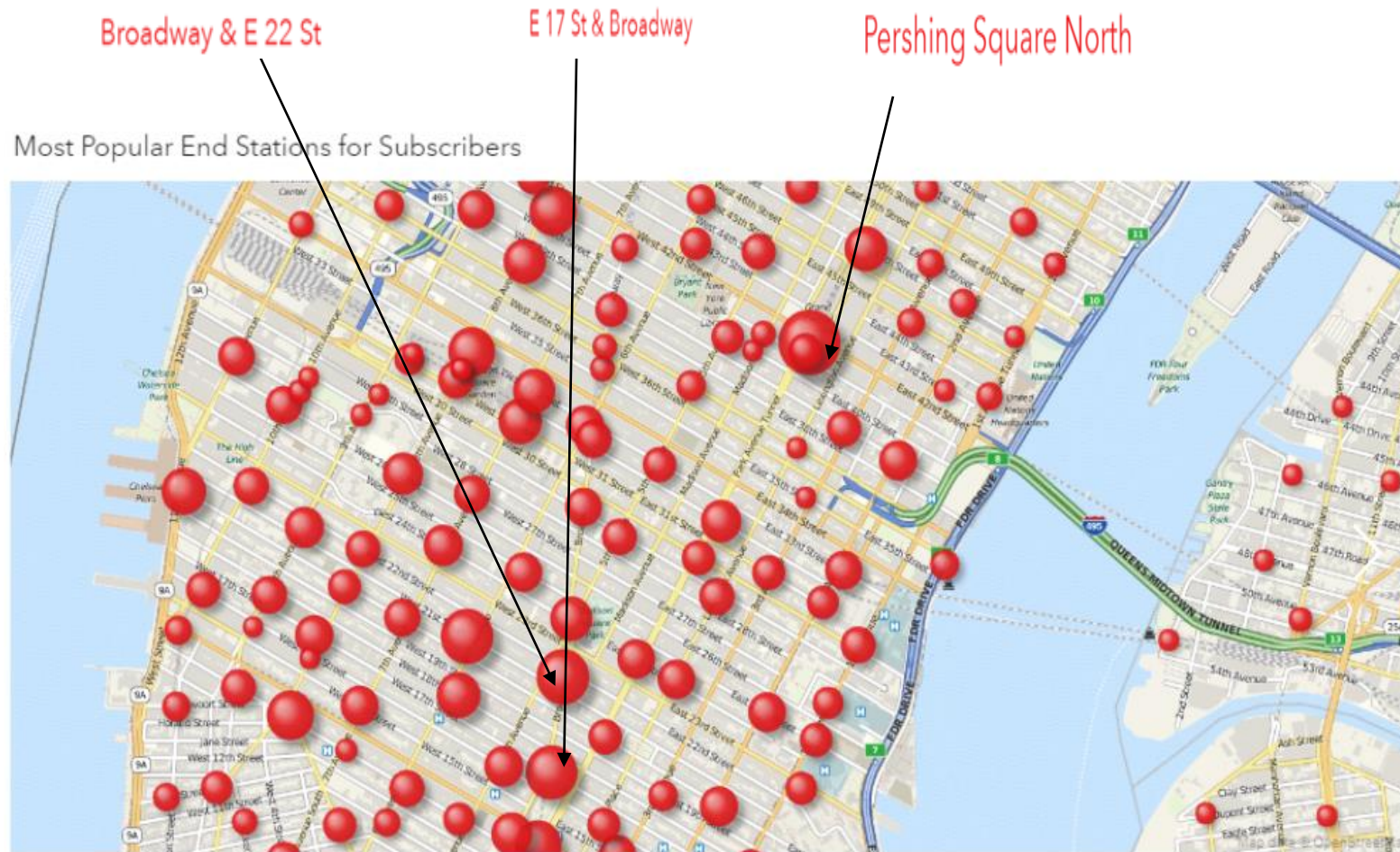
Most Popular Start Stations for Subscribers



- Bigger the size of text more the number of trips starting or ending at these stations
- Most famous start stations for subscribers are Pershing square in Manhattan in front of Grand Central Terminal, Broadway & E 22 St , W21st & 6 Ave etc.

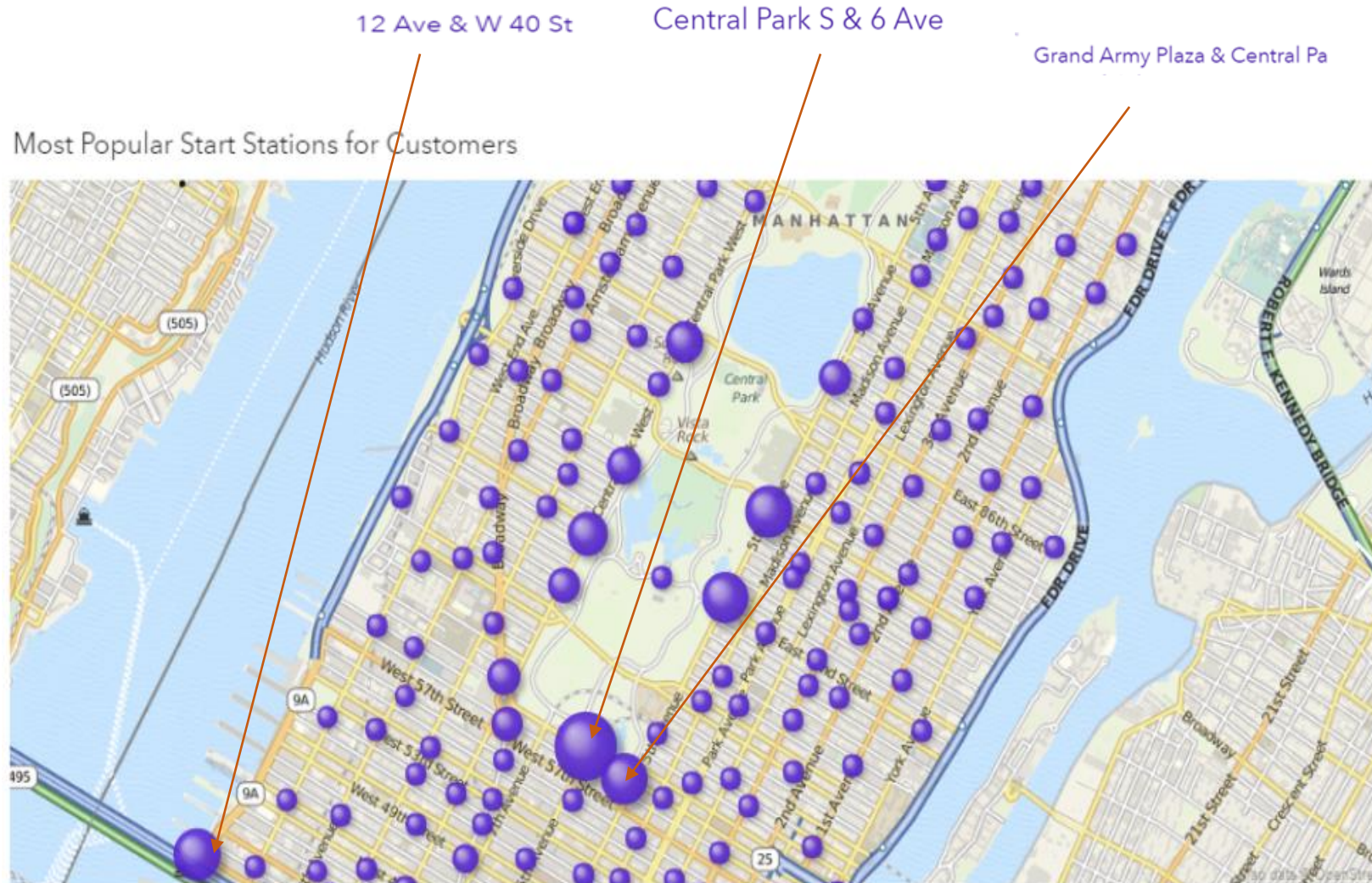
Geospatial representation of popular stations using SAS Visual Analytics

Most Popular End Stations for Subscribers



Geospatial representation of popular stations using SAS Visual Analytics

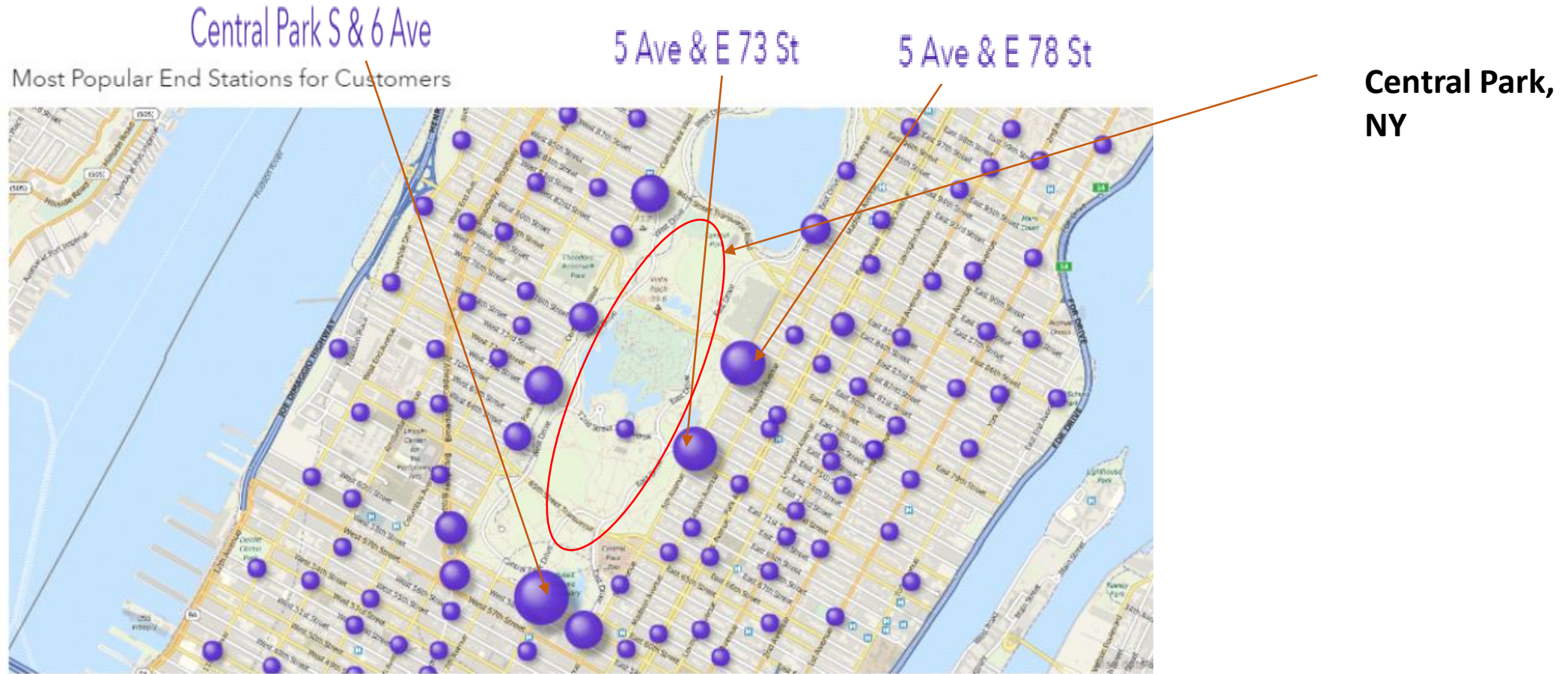
Most Popular Start Stations for Customers



Note that most popular stations for customers are around Central Park

Geospatial representation of popular stations using SAS Visual Analytics

Most Popular End Stations for Customers



Popular Routes for Customers and Subscribers

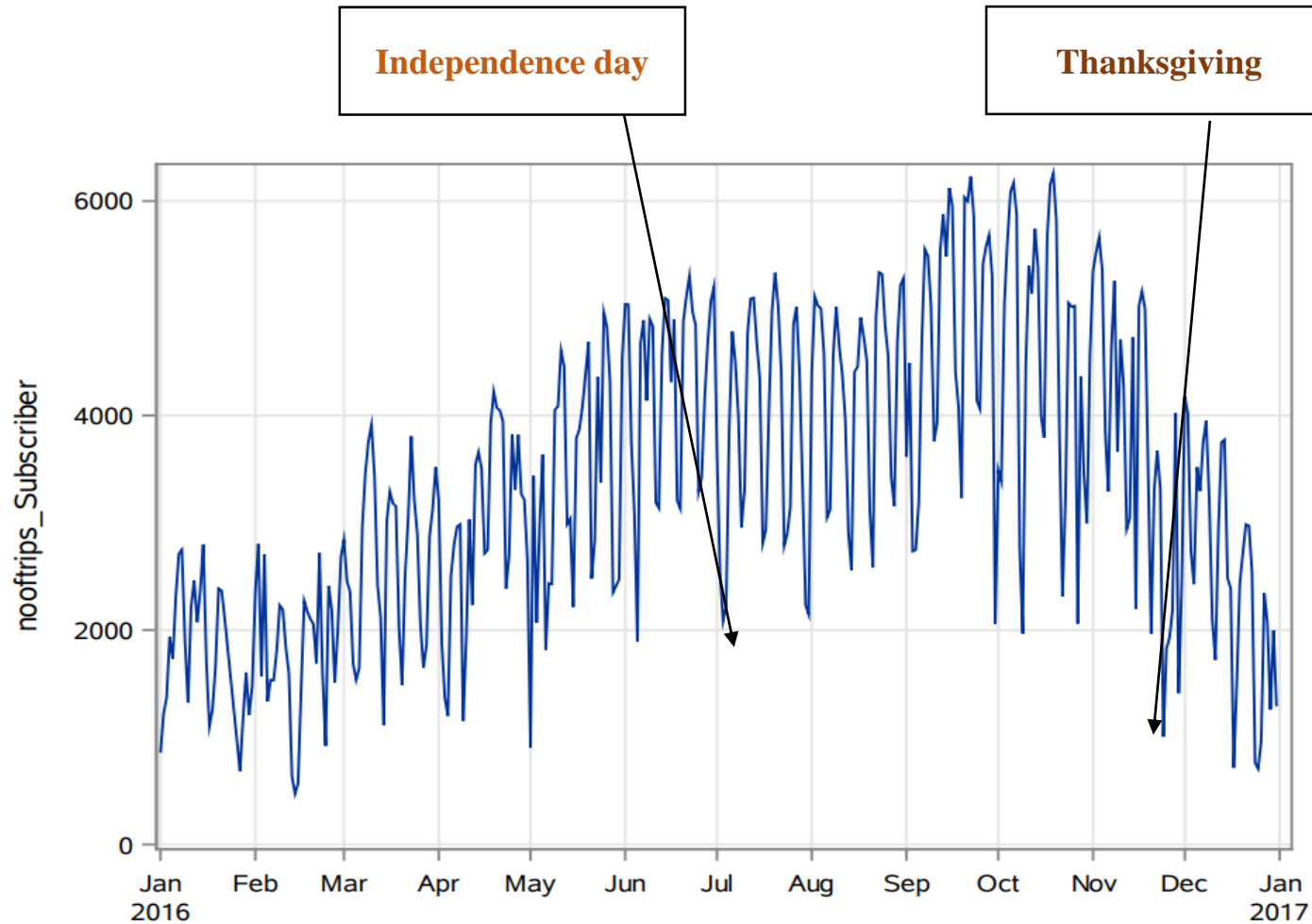
Top 20 busiest Routes for Subscribers

Obs	Route	CountOfRoute
1	Pershing Square North - W 33 St & 7 Ave	463
2	W 21 St & 6 Ave - 9 Ave & W 22 St	458
3	E 85 St & York Ave - E 85 St & 3 Ave	441
4	N 6 St & Bedford Ave - Wythe Ave & Metropolitan Ave	423
5	Pershing Square North - Broadway & W 32 St	418
6	Pershing Square North - E 24 St & Park Ave S	407
7	Pershing Square North - W 41 St & 8 Ave	403
8	Vernon Blvd & 50 Ave - 46 Ave & 5 St	379
9	Wythe Ave & Metropolitan Ave - N 6 St & Bedford Ave	349
10	1 Ave & E 68 St - Lexington Ave & E 63 St	348
11	N 6 St & Bedford Ave - S 4 St & Wythe Ave	348
12	W 17 St & 8 Ave - W 20 St & 11 Ave	335
13	1 Ave & E 62 St - 1 Ave & E 68 St	328
14	E 30 St & Park Ave S - Pershing Square North	321
15	Lexington Ave & E 63 St - 1 Ave & E 68 St	321
16	W 21 St & 6 Ave - W 22 St & 10 Ave	317
17	Bedford Ave & Nassau Ave - N 8 St & Driggs Ave	313
18	S 4 St & Wythe Ave - N 6 St & Bedford Ave	312
19	11 Ave & W 41 St - 8 Ave & W 33 St	311
20	E 47 St & Park Ave - W 41 St & 8 Ave	310

Top 20 busiest Routes for Customers

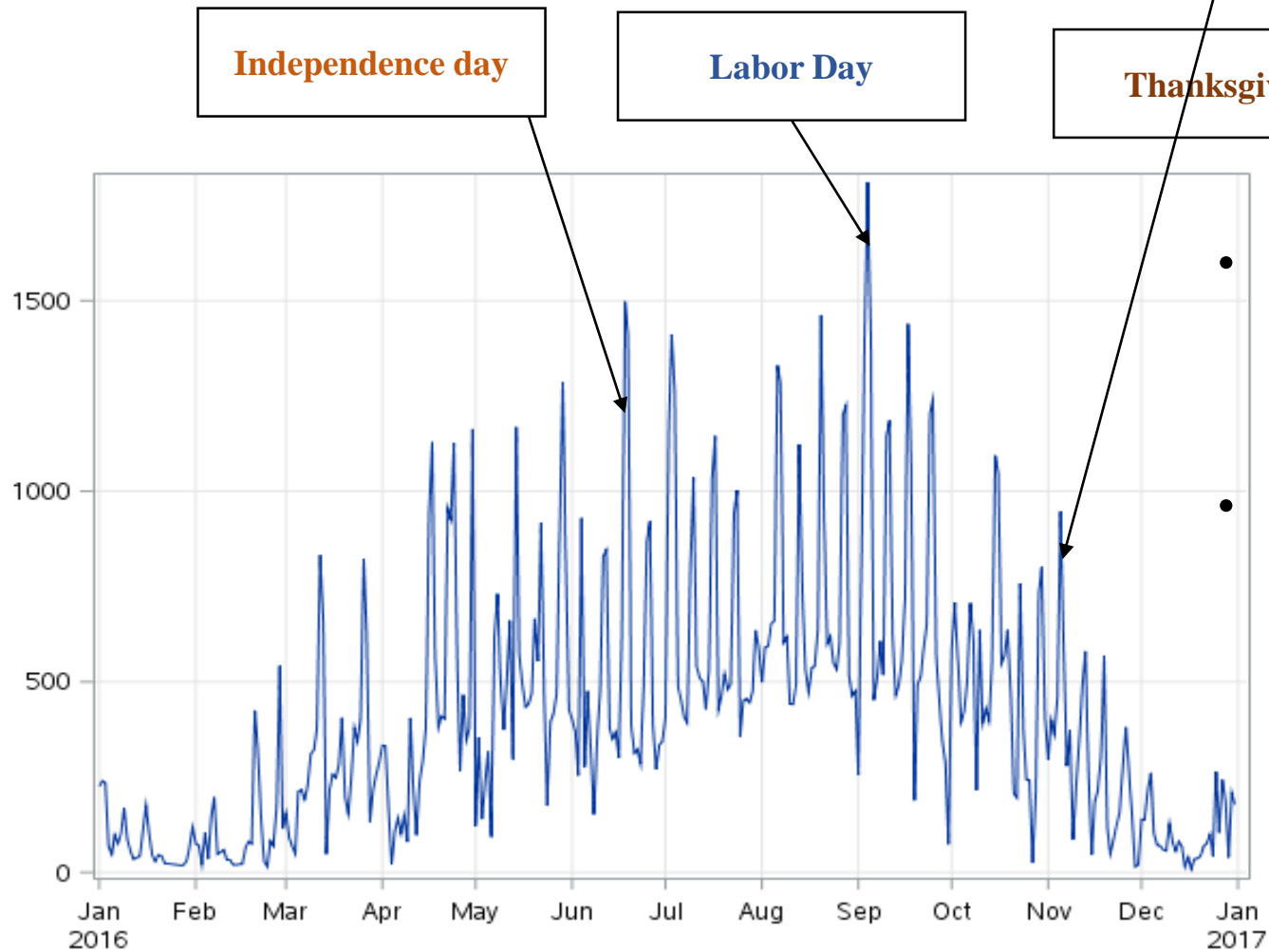
Obs	Route	CountOfRoute
1	Central Park S & 6 Ave - Central Park S & 6 Ave	823
2	Central Park S & 6 Ave - 5 Ave & E 78 St	527
3	Centre St & Chambers St - Centre St & Chambers St	381
4	Central Park S & 6 Ave - 5 Ave & E 73 St	334
5	Grand Army Plaza & Central Pa - Grand Army Plaza & Central Pa	295
6	Centre St & Chambers St - Cadman Plaza E & Tillary St	294
7	5 Ave & E 73 St - 5 Ave & E 73 St	291
8	Central Park S & 6 Ave - 5 Ave & E 88 St	290
9	Central Park S & 6 Ave - Central Park West & W 72 St	283
10	Old Fulton St - Centre St & Chambers St	278
11	12 Ave & W 40 St - West St & Chambers St	264
12	Grand Army Plaza & Central Pa - 5 Ave & E 78 St	242
13	5 Ave & E 78 St - Central Park West & W 85 St	238
14	Central Park West & W 72 St - Central Park West & W 72 St	238
15	Central Park West & W 85 St - Central Park S & 6 Ave	238
16	5 Ave & E 78 St - 5 Ave & E 78 St	235
17	Cadman Plaza E & Red Cross Pl - Centre St & Chambers St	235
18	Central Park S & 6 Ave - Central Park West & W 85 St	213
19	5 Ave & E 78 St - Central Park S & 6 Ave	207
20	West St & Chambers St - 12 Ave & W 40 St	205

Time Series of Rides over the last year (Jan.2016 - Dec.2016) for Subscribers



- Reason for Rise and Fall in the rides is due to usage of bikes on weekdays and weekends.
- Subscribers tend to ride bikes low on the weekends and holidays but rent bikes more during weekdays to commute to their workplace.

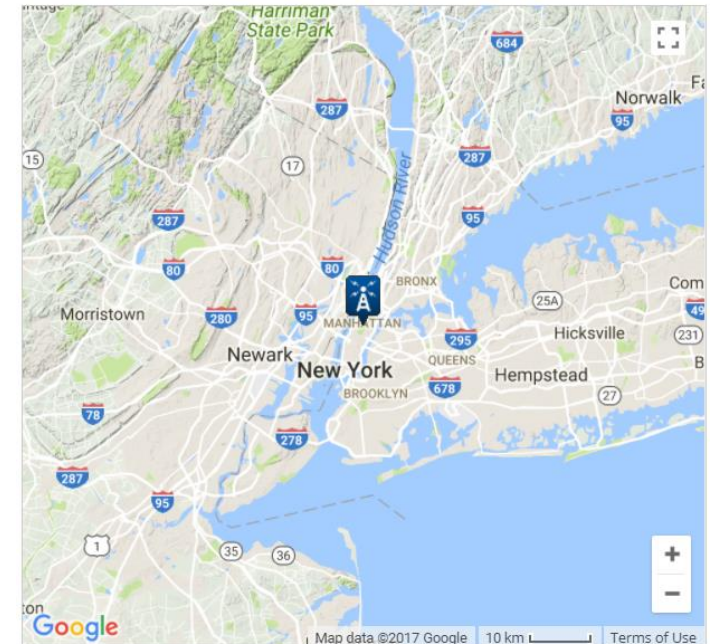
Time Series of Rides over the last year (Jan.2016 - Dec.2016) for Customers



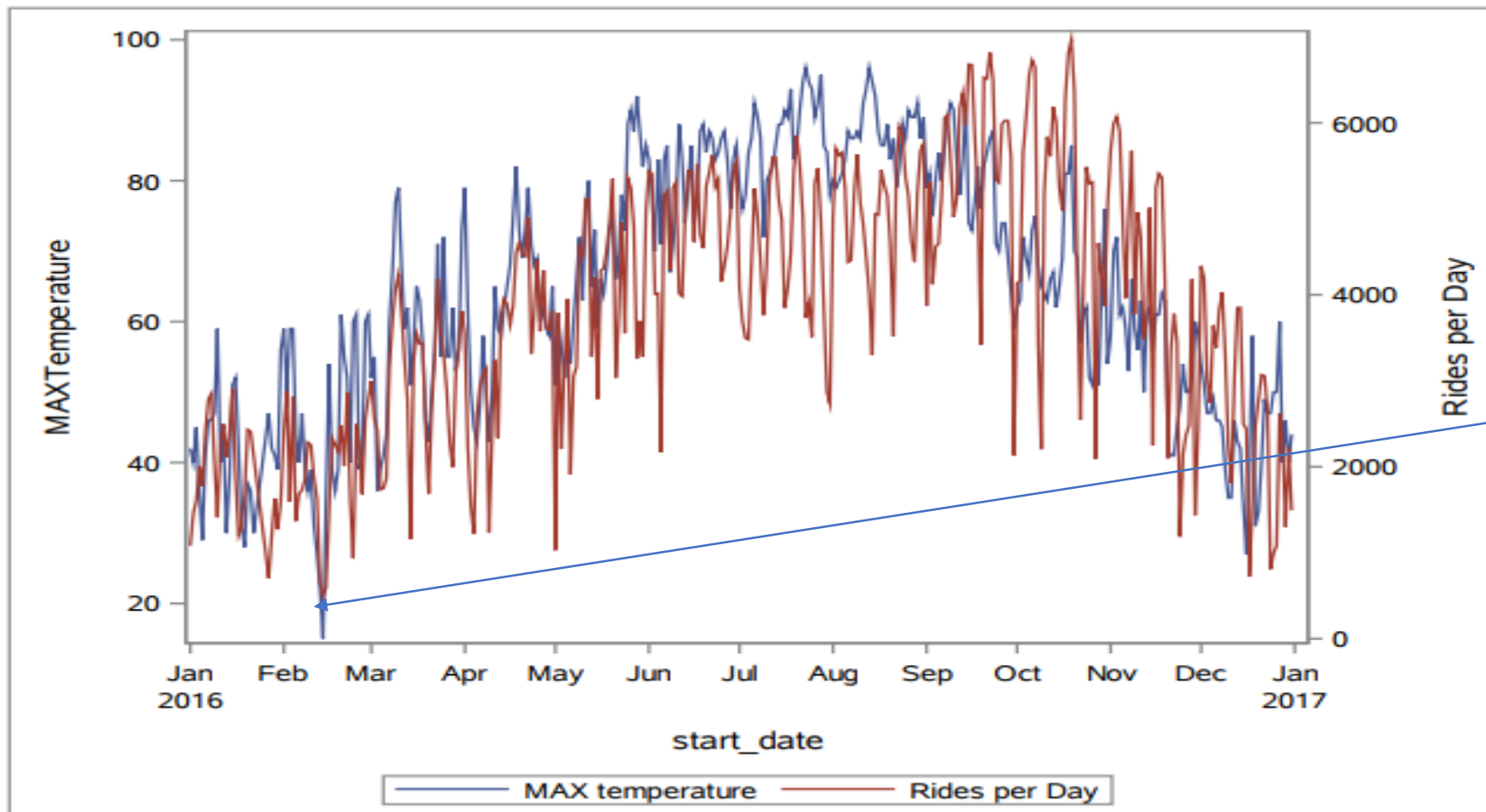
- Reason for Rise and Fall in the rides is due to the usage of bikes on weekends and weekdays.
- Rise in the rides for Customers mainly observed on the weekends and Holidays compared to weekdays.

Weather Data from National Climatic Data Center

- To provide additional insight for segmenting trips and riding days, a publicly available dataset of weather in the NYC area was merged by day with the trip data. This dataset included maximum and minimum temperatures, precipitation, snow, and wind for each day .
- It was taken from the national climatic data center. The data set included variables like temperature, snow, precipitation, and wind speed. These observations were merged on the latitude and longitude column to the Citi bike dataset.
- <https://www.ncdc.noaa.gov/climate-information/extreme-events>



Time series of Rides in a Day and Maximum Temperature

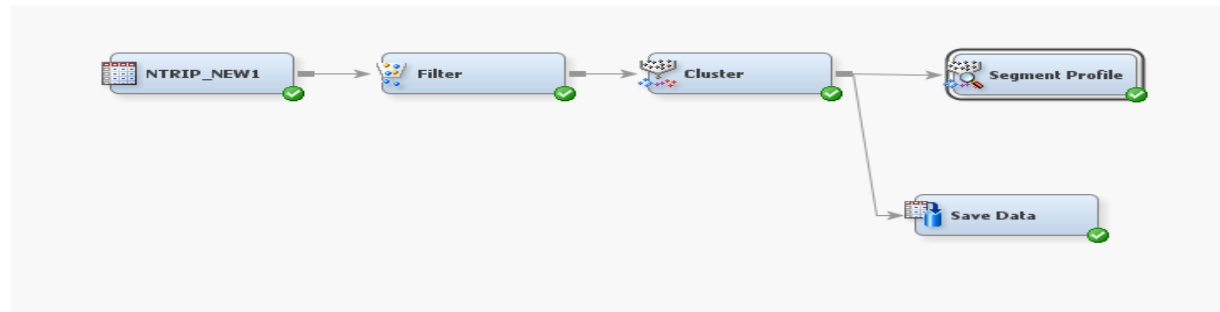


Number of rides decrease
in a day with dip in
temperature

Clustering to divide the dataset into certain distinctive characteristics

Clustering Criterion	Maximum Relative Change in Cluster Seeds	Improvement in Clustering Criterion	Segment Id	Frequency of Cluster	Root-Mean-Square Standard Deviation	Maximum Distance from Cluster Seed	Nearest Cluster	Distance to Nearest Cluster
0.439411	.0005528	.	1	13842	0.487013	9.880271	2	5.478329
0.439411	.0005528	.	2	666336	0.417928	9.258722	3	2.883323
0.439411	.0005528	.	3	426445	0.463381	8.990229	2	2.883323
0.439411	.0005528	.	4	63197	0.489502	11.44442	2	4.04417

Cluster statistics

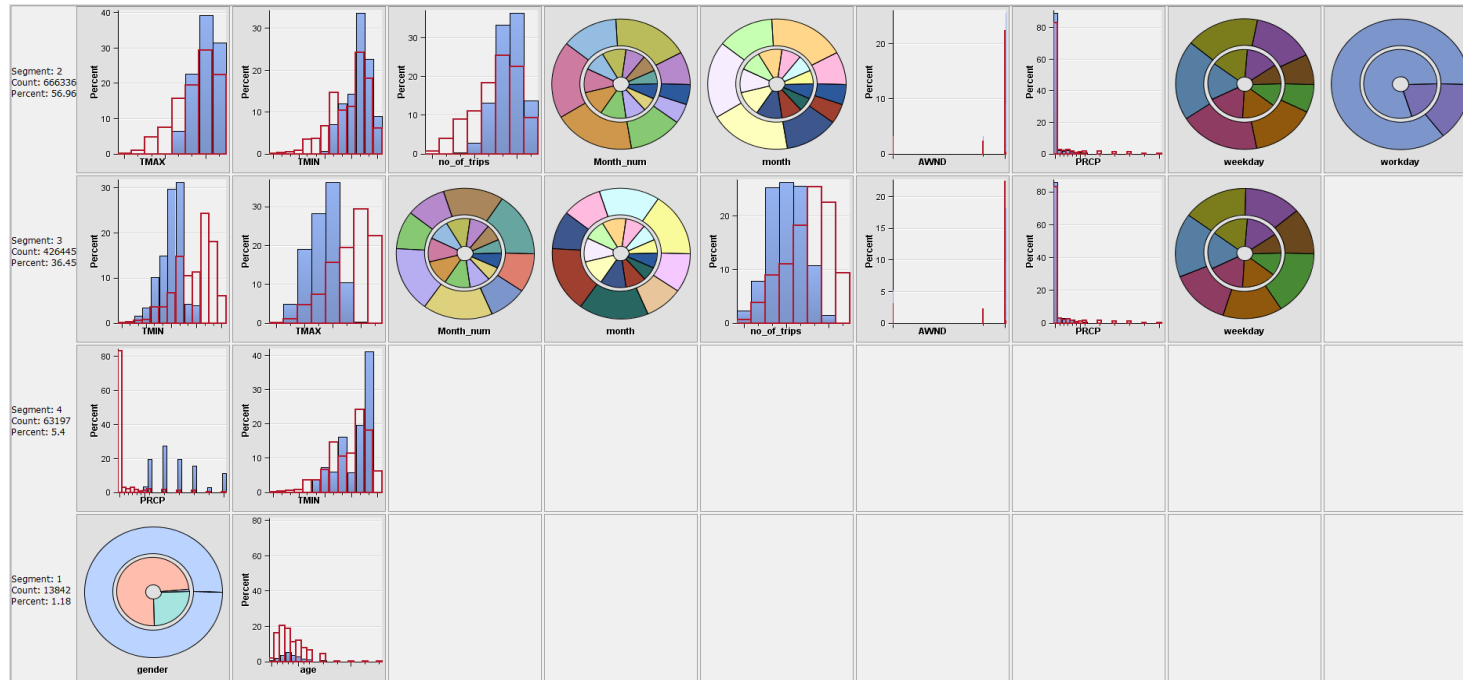


E.M Diagram Flow

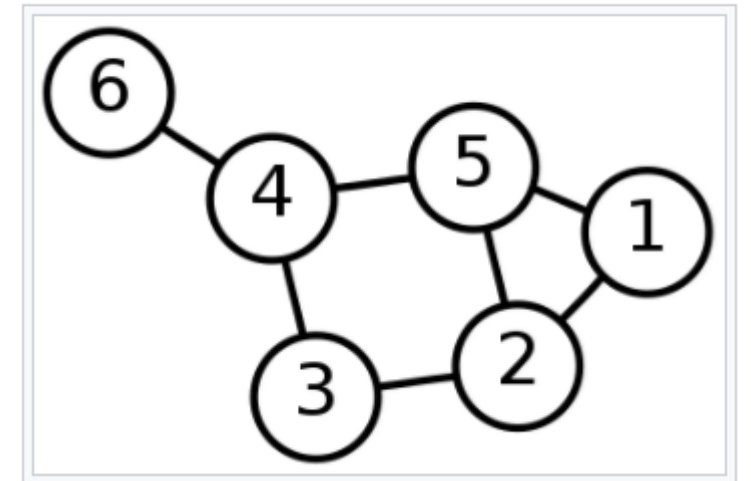
Variable	Variable Description
AWND	Average Wind Speed
HoD	Hour of Day
PRCP	Precipitation
SNOW	Snow (Indicator)
Tmax	Maximum Temperature
Tmin	Minimum Temperature
age	Age of rider
gender	Gender of Rider
month	Month of year
no_of_trips	No of trips per day
rush	nominal variable for rush hour of the day
trip_duration	Trip duration of each ride
weekday	Week day of the week
workday	Binary indicator for weekday/weekend

Variables used in clustering

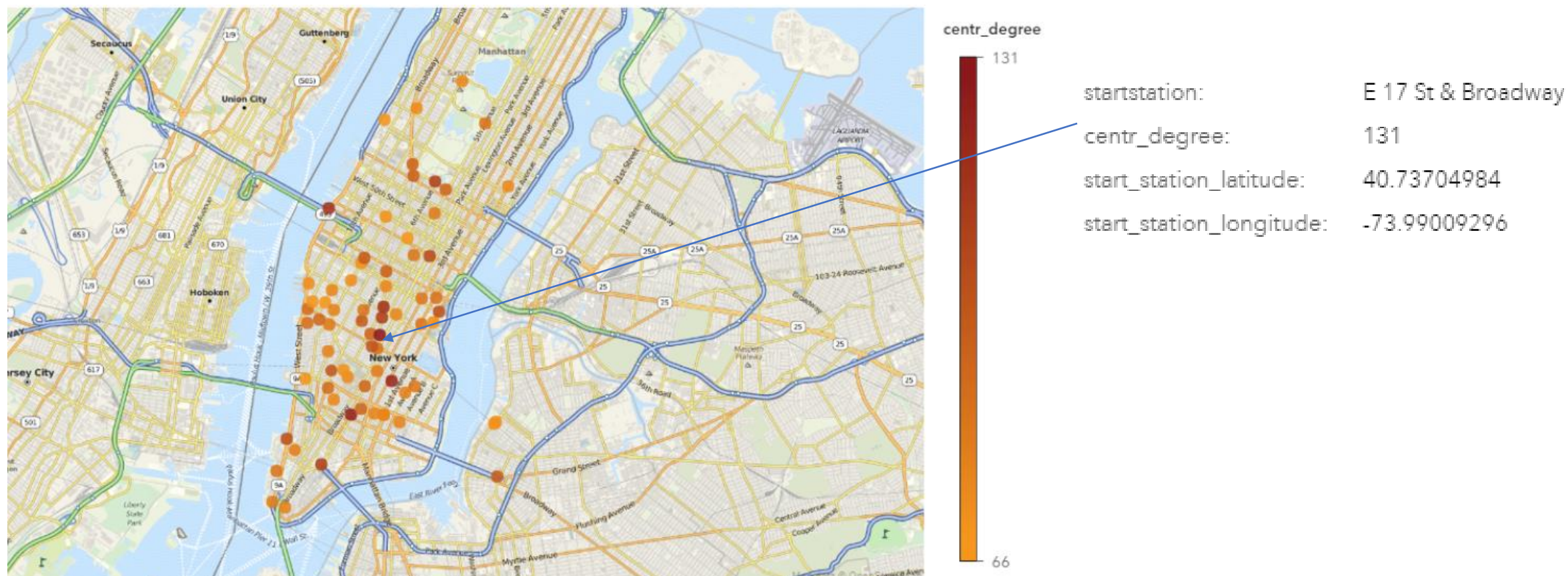
Segmentation Characteristics



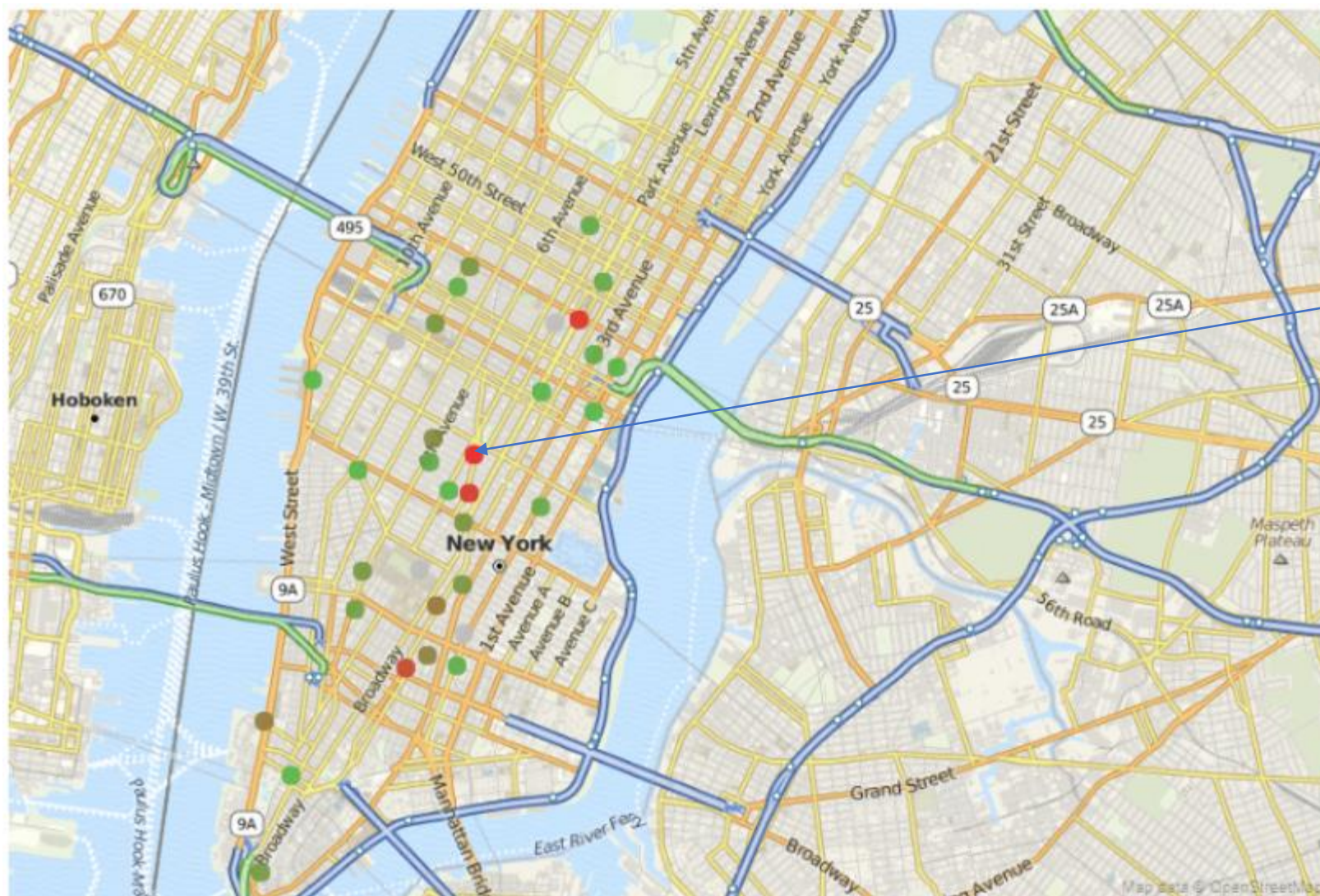
Segment Profile



Map of Busiest Stations by Degree Centrality, Citi Bike Trip Cluster 1



Map of Busiest Stations by Degree Centrality, Citi Bike Trip Cluster 2



centr_degree

□ (missing)

789

startstationname: Broadway & E 22 St

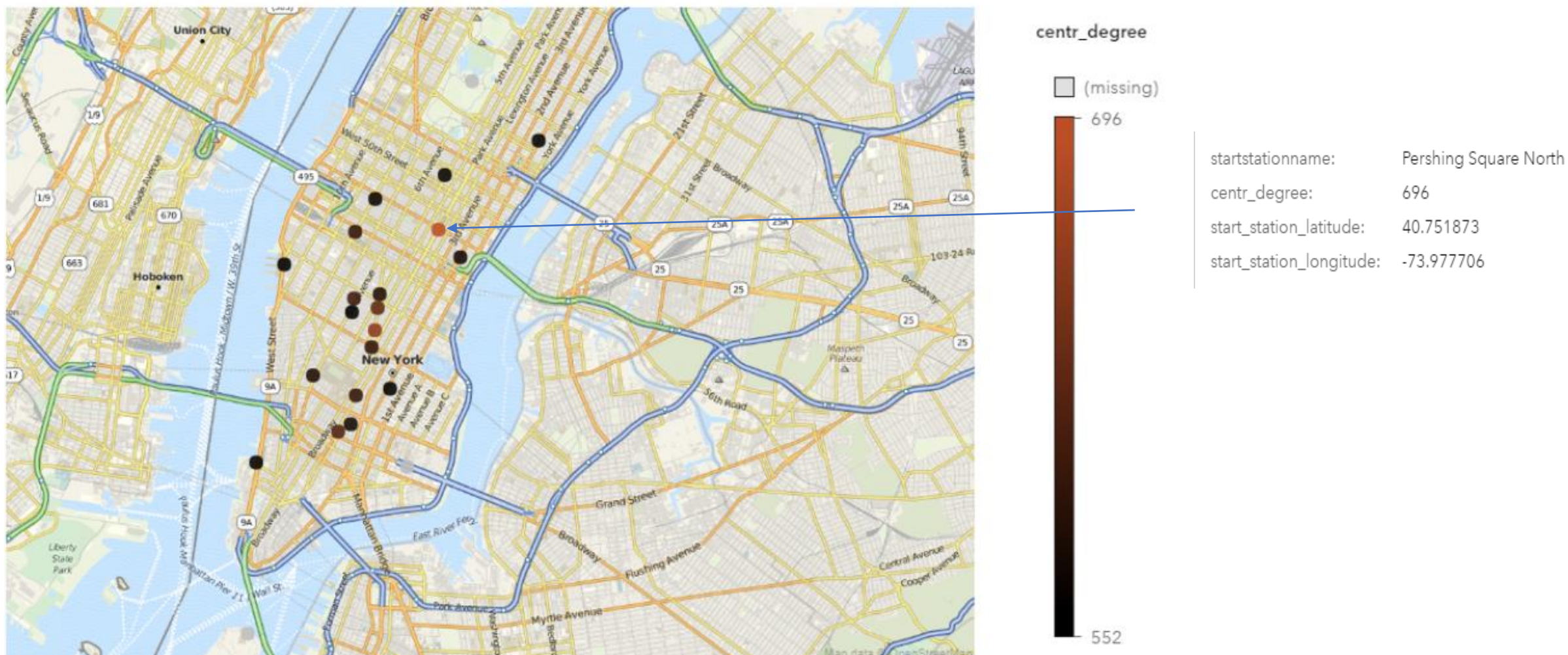
centr_degree: 789

start_station_latitude: 40.7403432

start_station_longitude: -73.98955109

653

Map of Busiest Stations by Degree Centrality, Citi Bike Trip Cluster 3



Map of Busiest Stations by Degree Centrality, Citi Bike Trip Cluster 4

