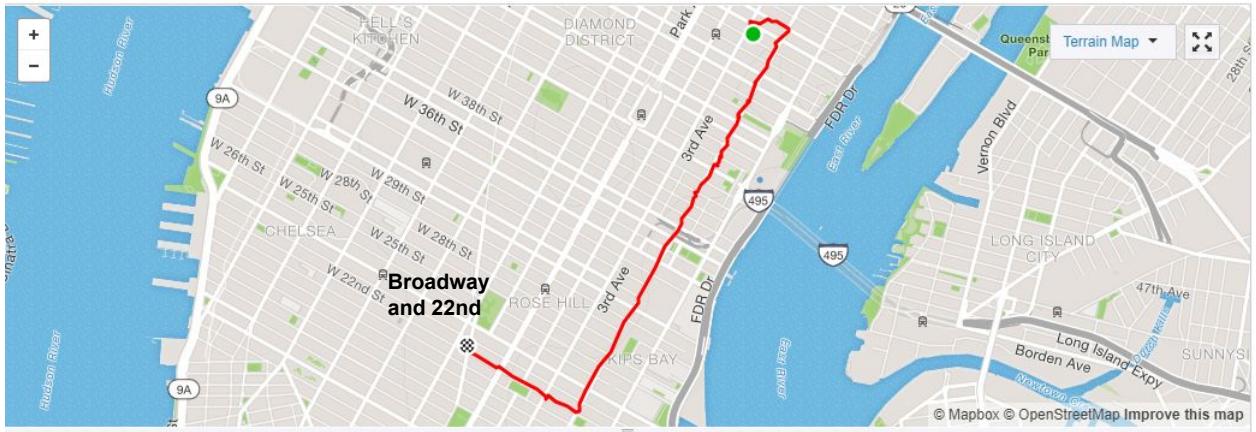
citi bike
citi bike
citi bike

2.5 mi	15:36	70 ft
Distance	Moving Time	Elevation

86 W	81 kJ
Estimated Avg Power	Energy Output

	Avg	Max
Speed	9.8 mi/h	26.2 mi/h
Elapsed Time	17:05	

Show More



IF THESE WHEELS COULD TALK

By Daniel Clark

WHAT DO WE HAVE TO WORK WITH?

- Comprehensive data of every Citibike ride in NYC dating back to July 2013.

Data includes information on the following:

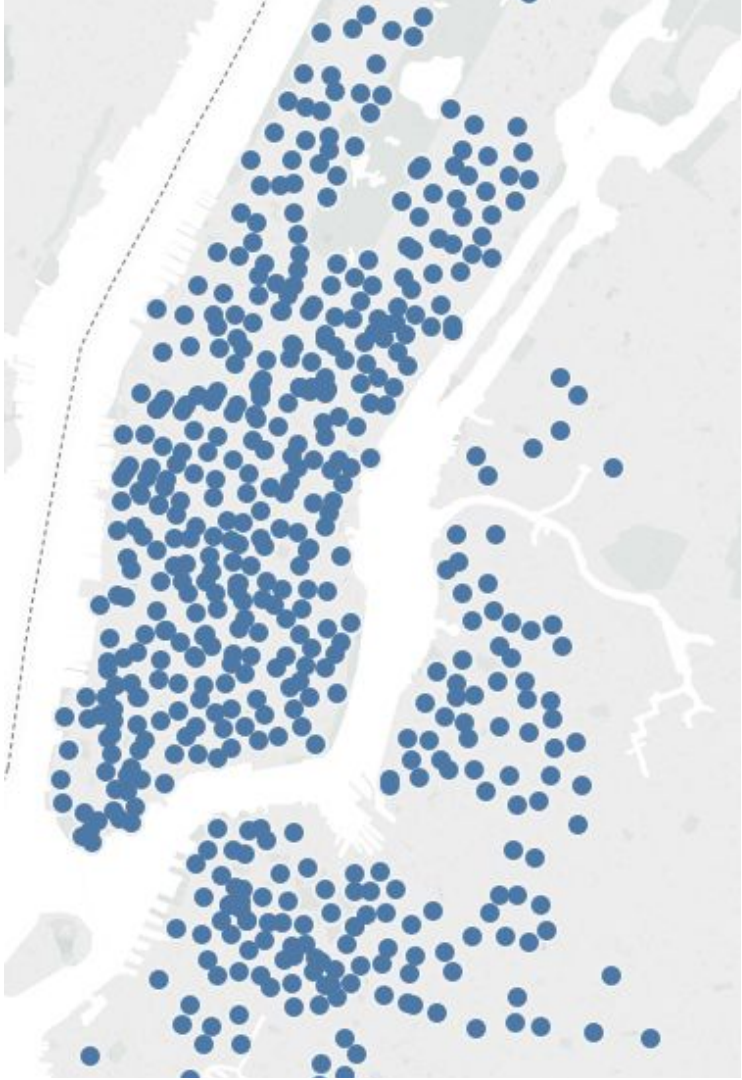
- Trip Start/End Time
- Start/End Location (Latitude / Longitude Coordinates)
- Trip Duration (Seconds)
- Bike ID Number
- Birth Year of Rider

With some Simple Math, we can figure out the following:

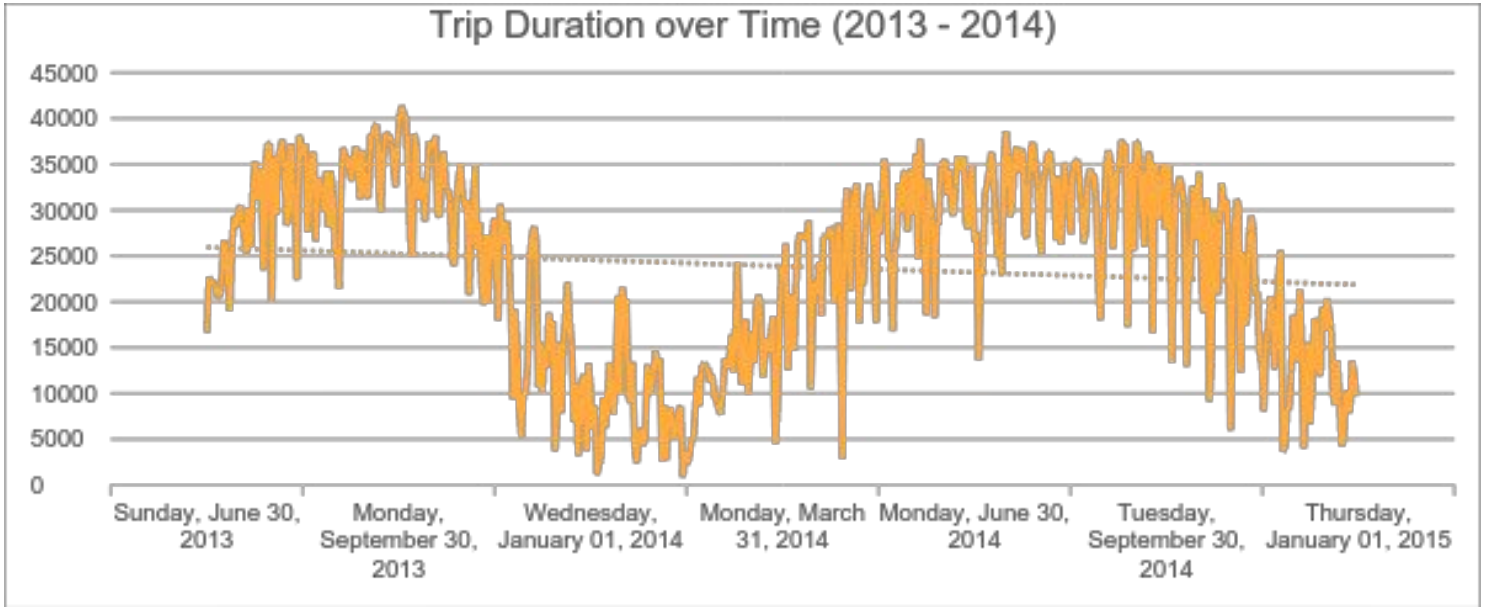
- Net Trip Distance (Miles)
- Trip Velocity (Miles Per Hour)
- Angular Direction
- Cardinal Direction
- Age of Rider

WHERE HAS IT BEEN?

- 5,121 Previous trips before me from July 1, 2013 through September 2017.
- The bike has gone a total of 5,491 net miles which amounts to about 1.07 Miles per trip. That's further than the distance from New York to Russia, going east.
- The bike's jurisdiction covers 40 square miles of Manhattan, Brooklyn and Queens



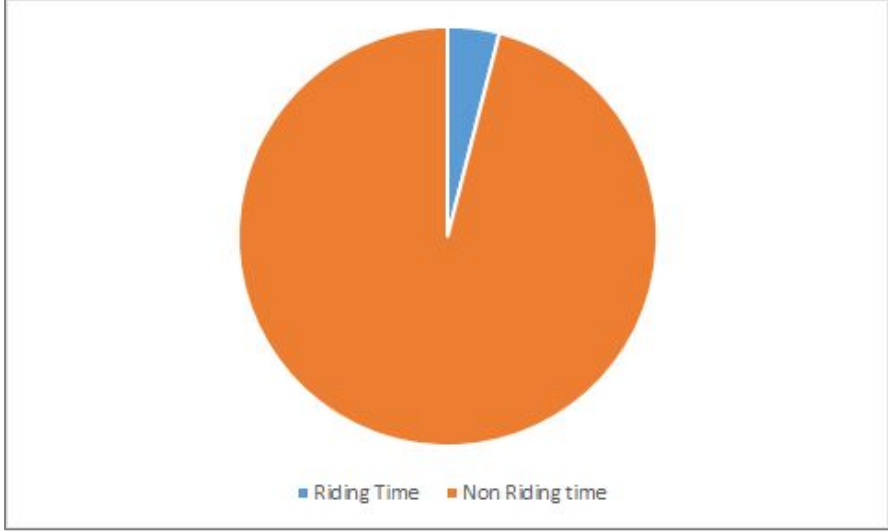
WHEN DO PEOPLE RIDE IT?

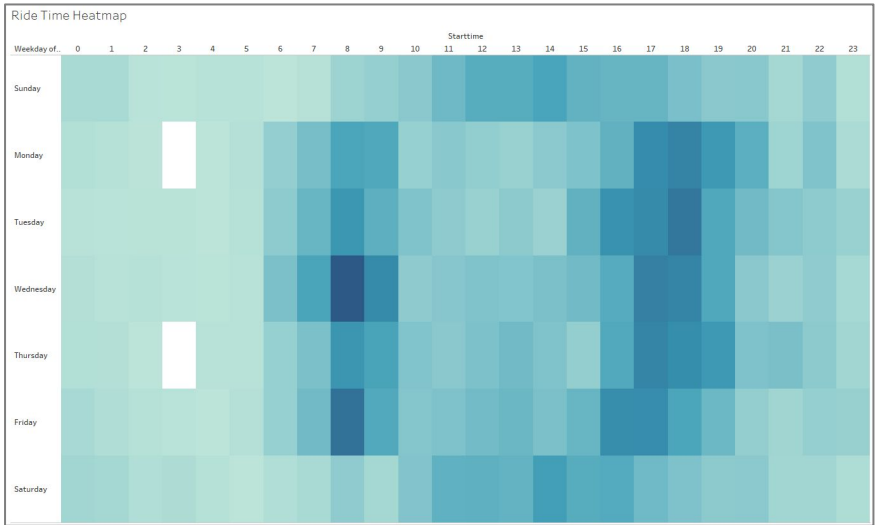


- Taking a quick chunk of data from the first couple of years of rides from Citibike, We can see that rides significantly drop in the winter time to recover in the Summer.
- What isn't quite as obvious is that over the past couple of years, the total number of Citibike rides is trending downward over time.

WHEN DO PEOPLE RIDE IT?

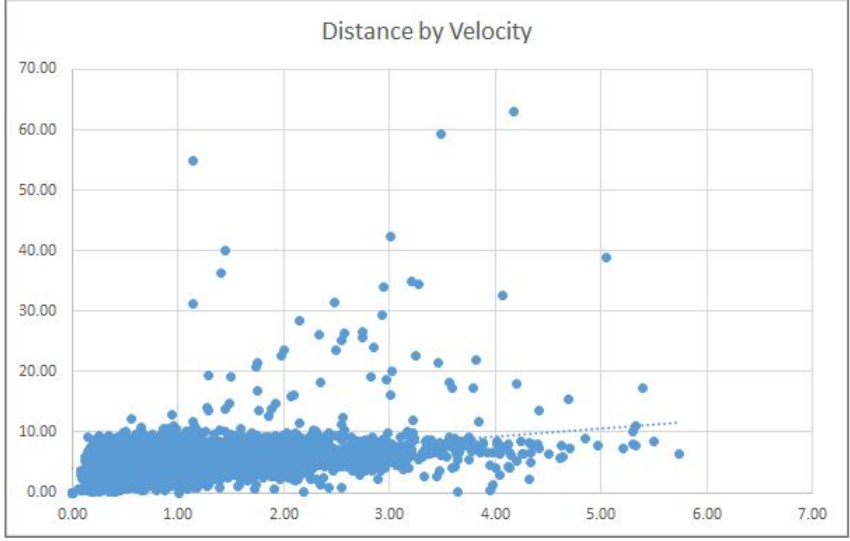
- During the bike's history, it spent 4% of time being ridden while 96% of time sitting around in it's station
- The majority of the bike rides take place between 7 am and 10 am on weekdays with a large increase in ridership during evening rush hours.
- On the weekends, ridership increases during the afternoons on Saturday's and Sundays

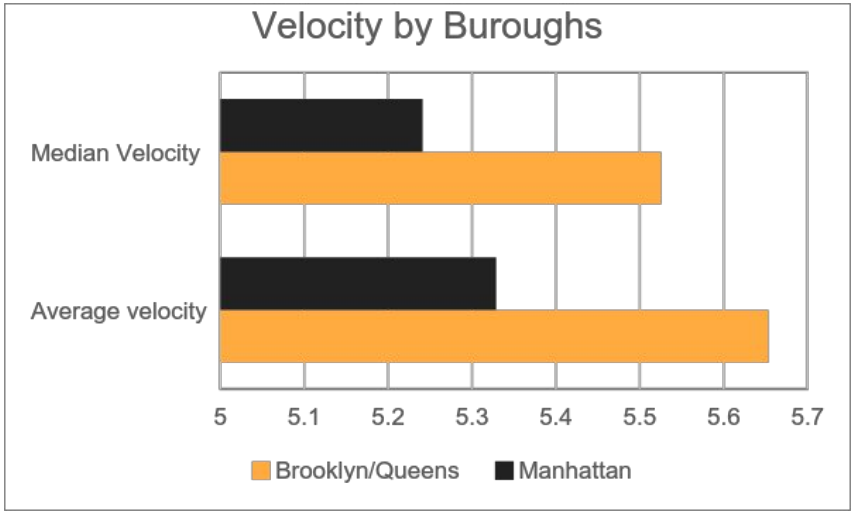




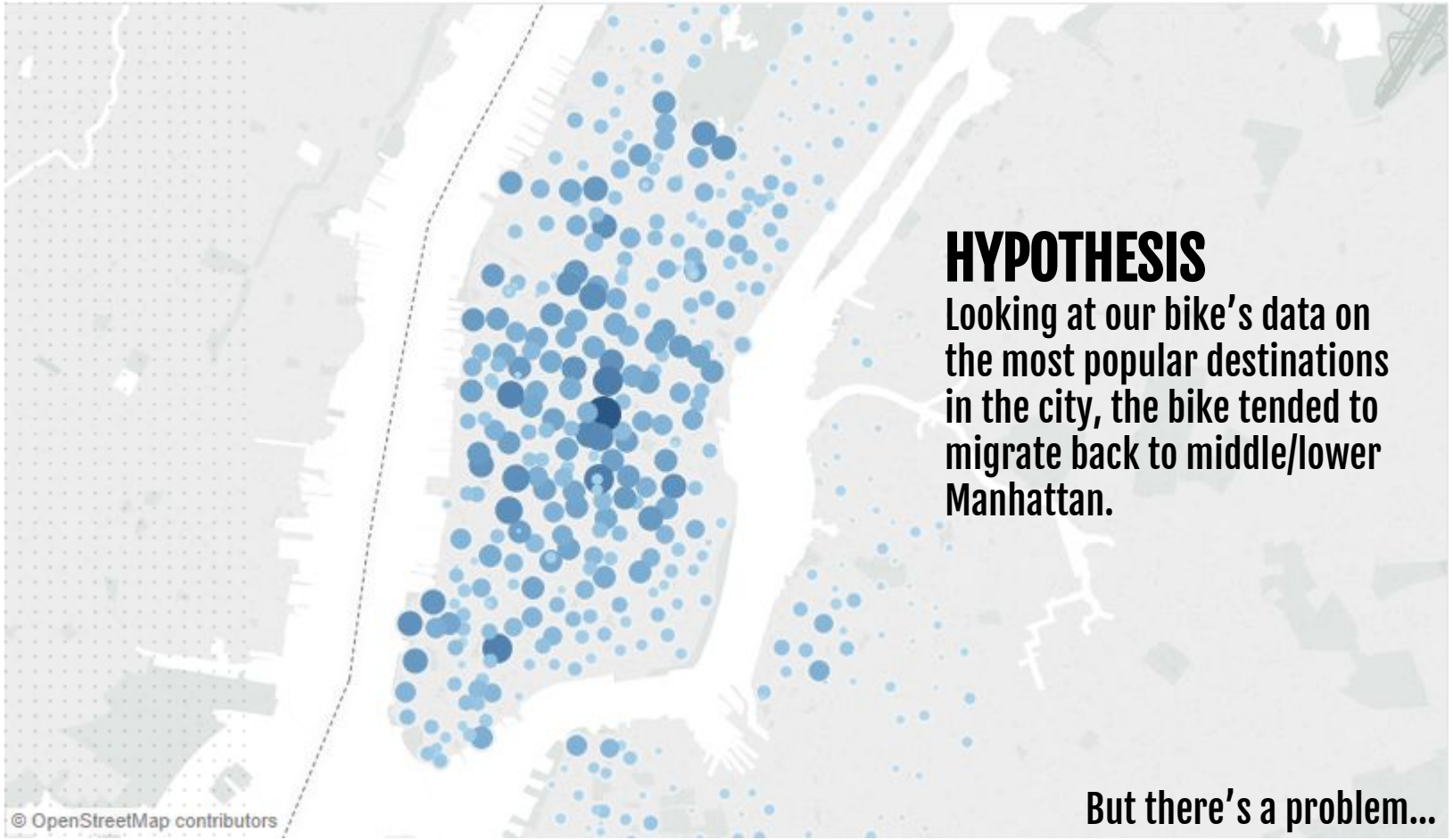
How Fast Has it Gone?

- Bike averages a speed of 5.4 miles per hour and caps off at a little over 10 miles per hour per trip.
- Looking at end station data for the bikes that traveled over 15 miles per hour are located near bridges.
- Bike tended to go faster in the outer boroughs than it did in Manhattan.





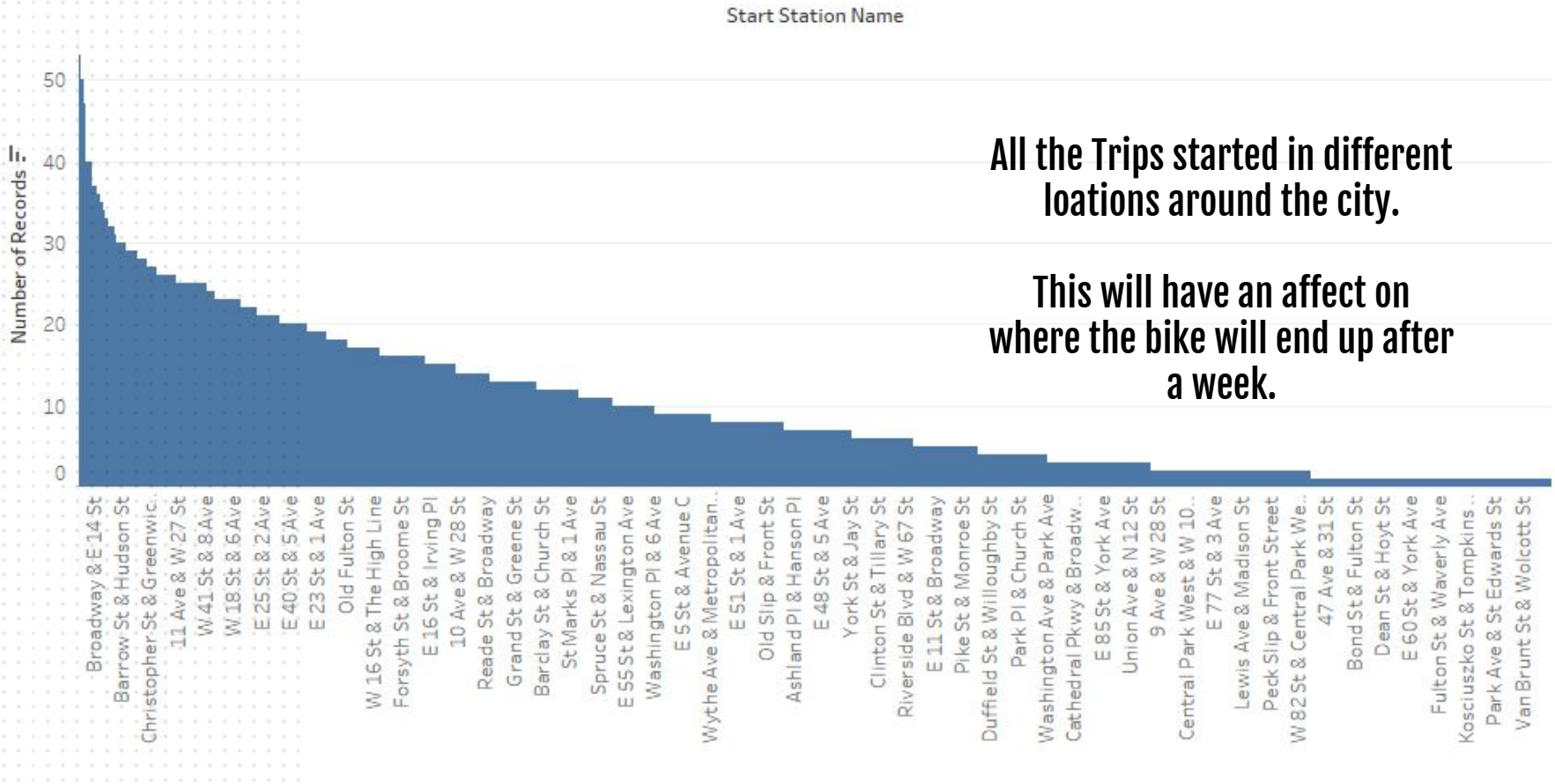
BASED ON THE AVAILABLE DATA
CAN WE PREDEICT
WHERE 15832
WOULD BE TODAY?



HYPOTHESIS

Looking at our bike's data on the most popular destinations in the city, the bike tended to migrate back to middle/lower Manhattan.

But there's a problem...



All the Trips started in different loations around the city.

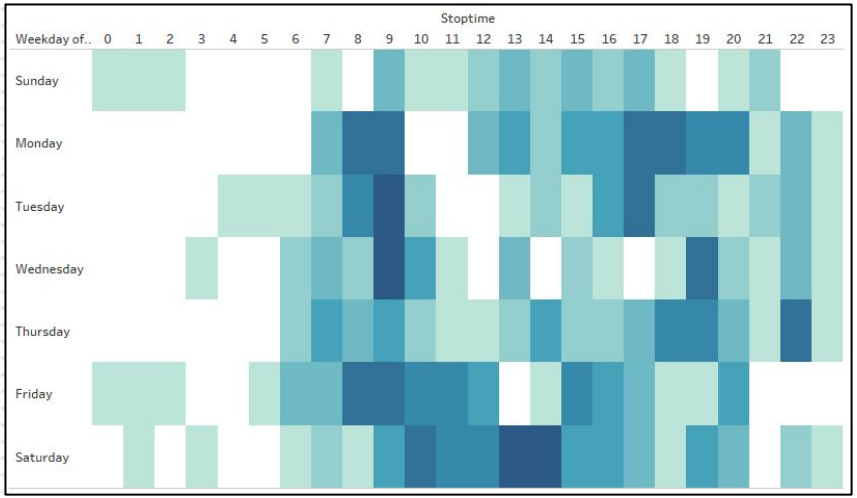
This will have an affect on where the bike will end up after a week.

METHODOLOGY

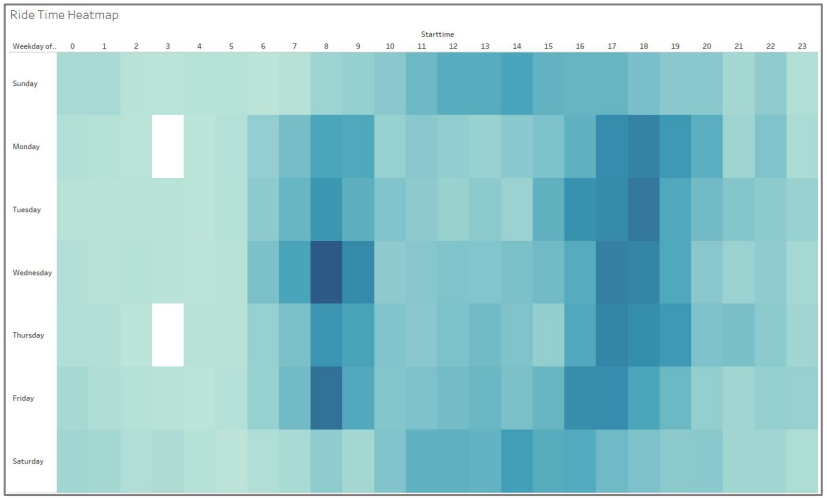
Take a representative sample of bikes with a start location on 22nd and Broadway during the first Thursday of November at 5:30 PM

2013	15463	16401	16767	17928	18313
2014	16120	16890	17260	18433	20112
2015	18257	20002	23140	23700	23940
2016	17219	21078	22153	23771	25064

HOW REPESENTATIVE IS THIS SAMPLE?



Sample Heat map



Bike 15832 Heat map

METHODOLOGY

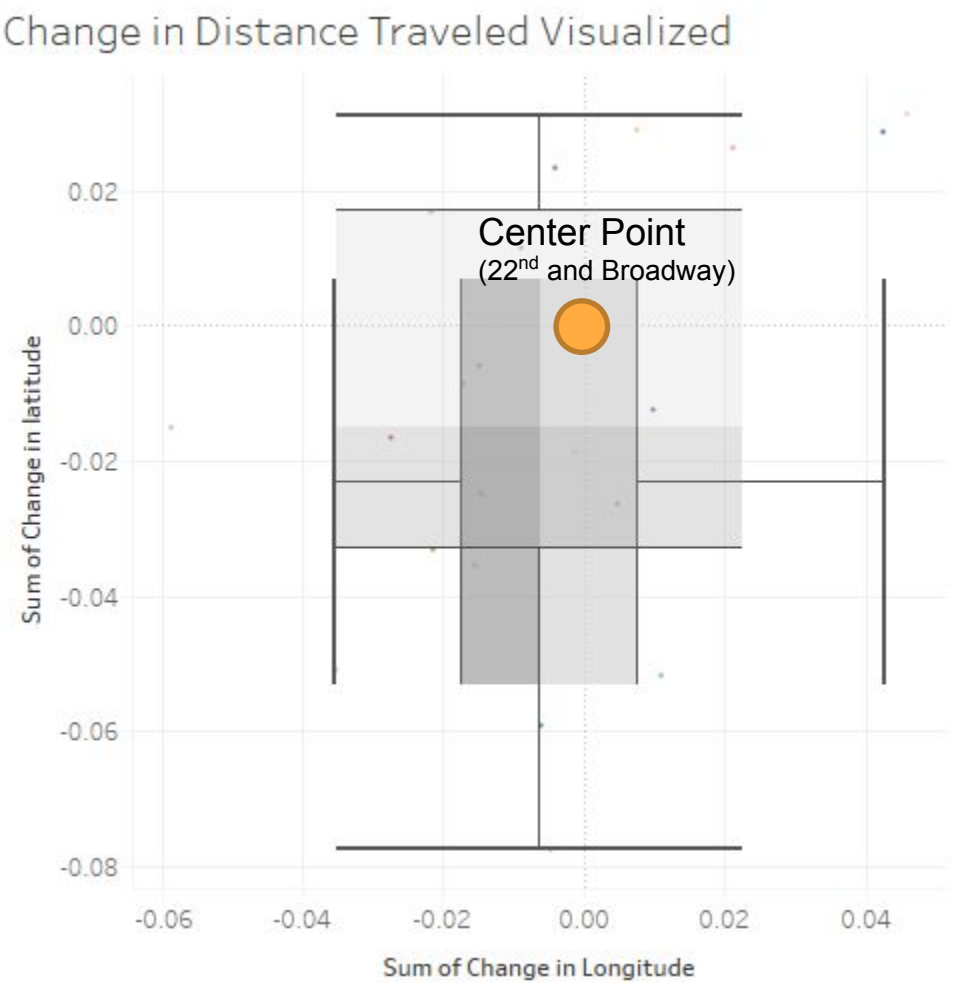
Calculate the Change in Latitude and Longitude after a week's worth of trips, aka net distance from center(or NDFC), and plot out what we find.

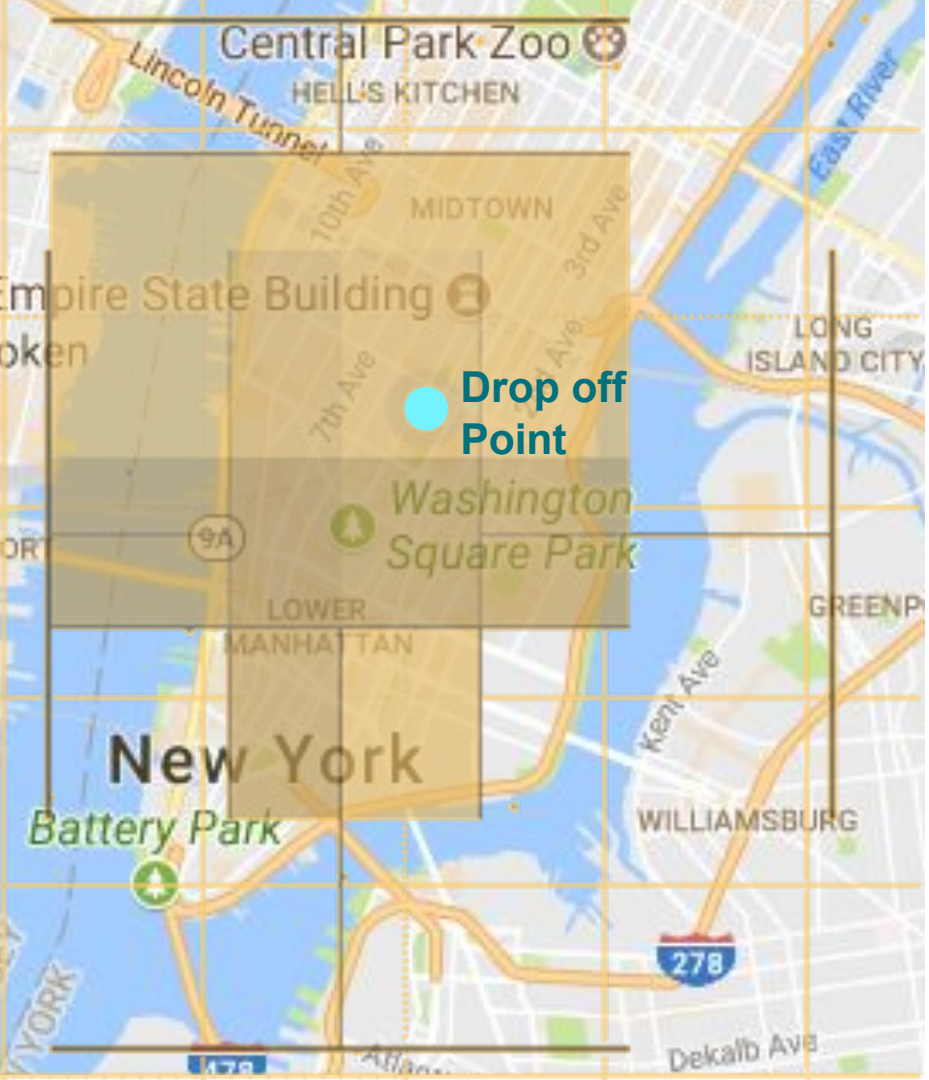
From here, you can calculate the average and variance of the NDFC to see patterns in where the bike has migrated after a week

Row Labels	Sum of Change in latitude	Sum of Change in Longitude
⊞<11/7/2013		
⊟2013		
15463	-0.012583792	0.00992175
16401	-0.07740873	-0.00474443
16767	0.02881185	0.00763268
16826	-0.00590172	-0.01479337
17928	-0.05905855	-0.00612938
18313	-0.01648654	-0.02729408
⊟2014		
16120	-0.01501777	-0.05854868
16890	0.01147205	-0.00901905
17260	-0.01870622	-0.0012262
18433	0.02635116	0.02125945
20112	-0.02491861	-0.01453084
⊟2015		
18257	-0.05161756	0.01101942
20002	-0.02641496	0.00462288
23140	-0.00861892	-0.01719327
23700	0.01683596	-0.02173603
23940	-0.035632634	-0.015342594
⊟2016		
17219	-0.033097394	-0.021283533
21078	-0.05096752	-0.03538435
22153	0.03125424	0.0458214
23771	0.023195184	-0.004077586
25064	0.028533111	0.04251012

RESULTS

- Using the Box-and-whisker plot in Tableau, we can be reasonably certain the bike likely went south and slightly west over a span of a week.

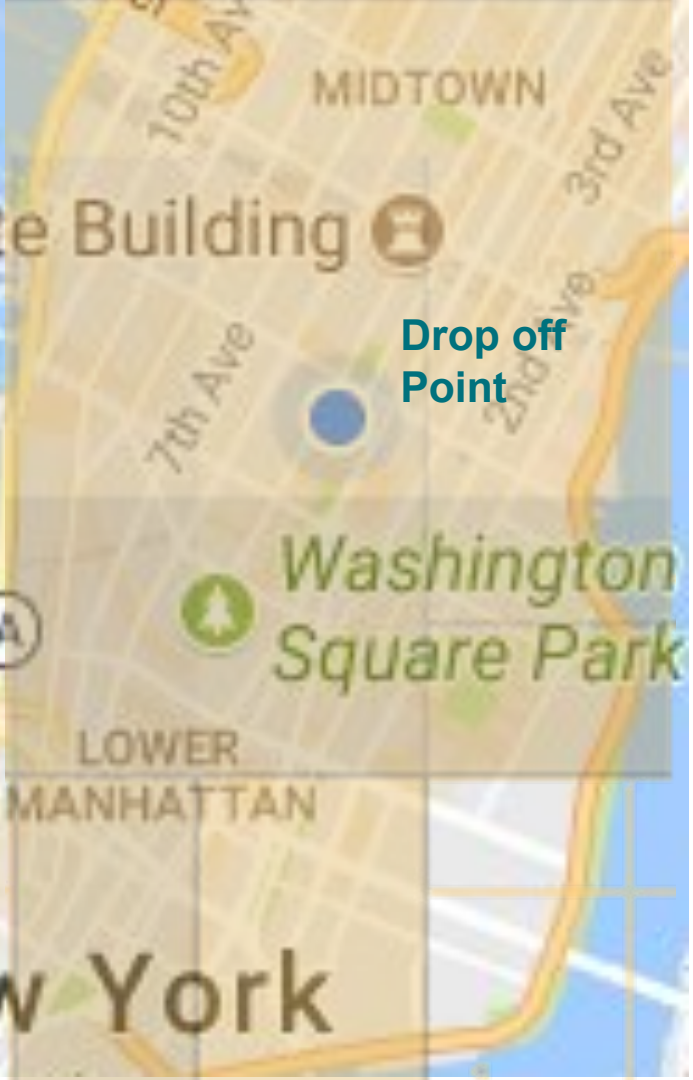




WHAT DOES THAT MEAN FOR CITY LOCATION?

- With outliers and rules about taking Citibikes across state lines considered, we can be most certain that the bike is somewhere South of Central Park, West of LIC, North of Prospect Park, and East of the Hudson River

- The locations of where our boxes converge represent the more likely places the bike may have ended up after a week.



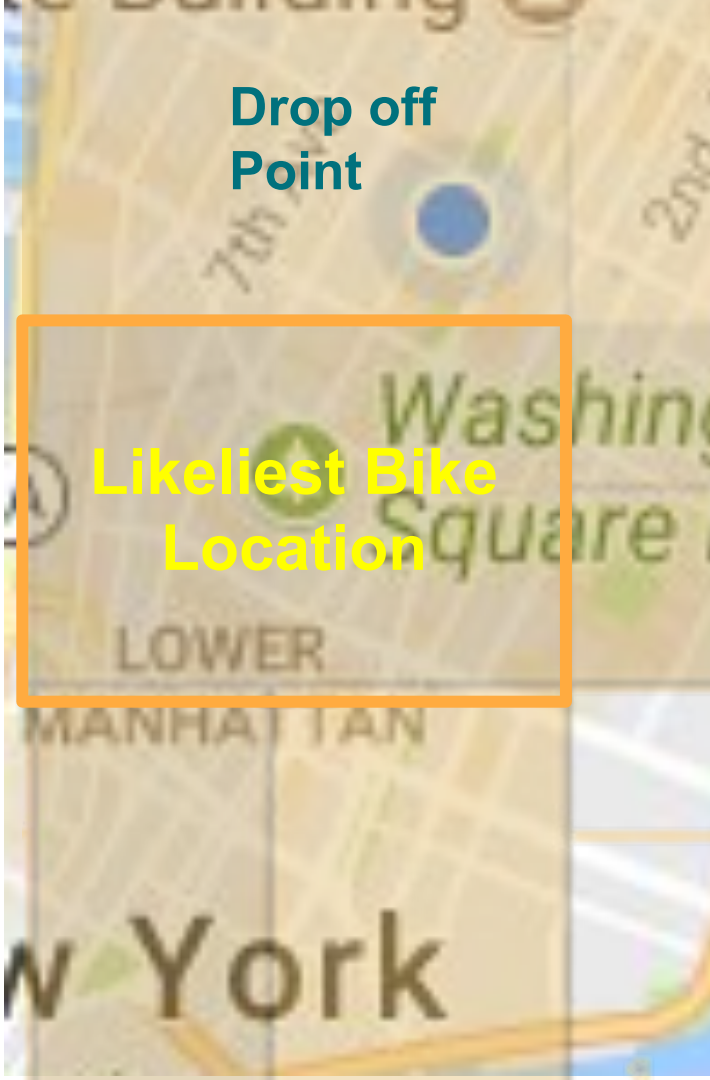
CAN WE GO DEEPER?

Weighting for outliers and looking into the area within 2 standard deviations, we can narrow our possible bike location to somewhere between Midtown and Lower Manhattan.

CAN WE GO EVEN DEEPER?

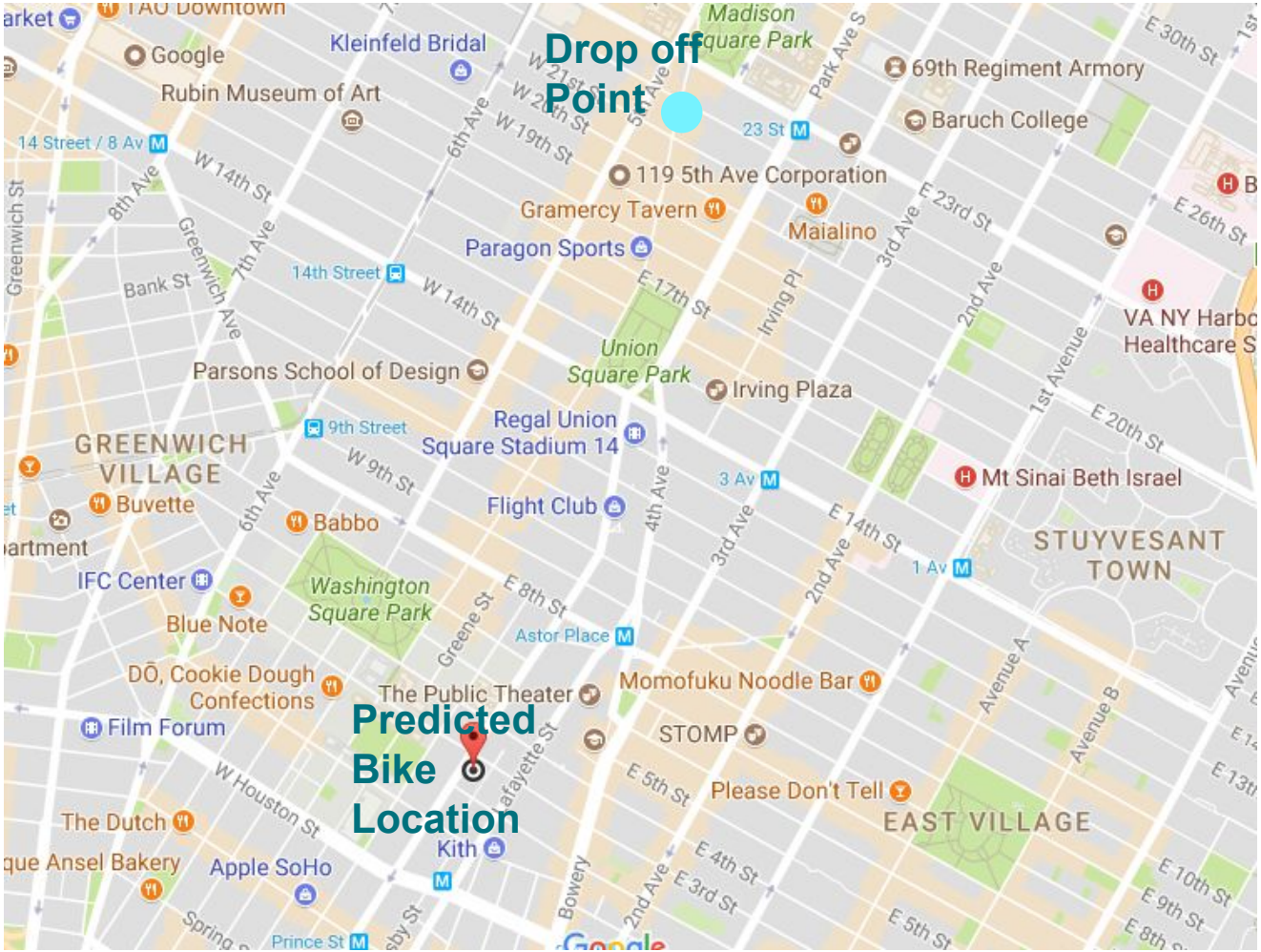
We can take the average change in longitude and latitude across all 970 bike trips among our samples during our specified timeframe and add that to our starting location on 22nd and Broadway.

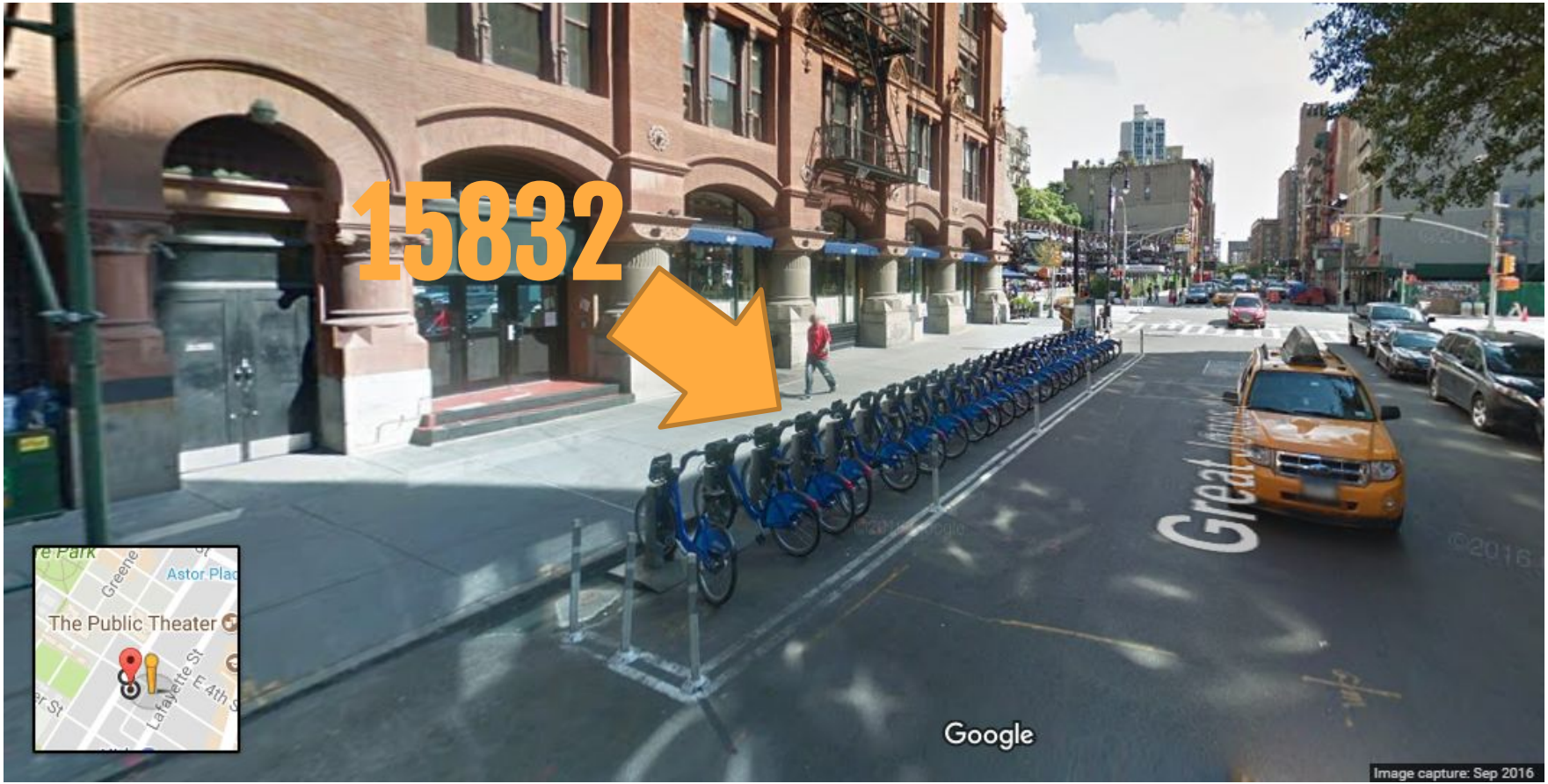
Coordinating that with the central zone of our box and whisker plot showing the most likely places bike number 15832 could be, this helps us conclude that...



RIGHT NOW, BIKE # 15832 IS HERE:

40.72748713, -73.9947185


Drop off Point
Predicted Bike Location
Google
Kleinfeld Bridal
Rubin Museum of Art
Madison Square Park
69th Regiment Armory
Baruch College
23 St
119 5th Ave Corporation
Gramercy Tavern
Maialino
Paragon Sports
14th Street
W 14th St
W 19th St
W 21st St
6th Ave
5th Ave
Park Ave S
E 30th St
E 23rd St
E 26th St
3rd Ave
2nd Ave
1st Avenue
Irving Pl
E 17th St
VA NY Harbo
Healthcare S
Union Square Park
Irving Plaza
Parsons School of Design
Regal Union Square Stadium 14
9th Street
W 9th St
GREENWICH VILLAGE
Buvette
Babbo
Flight Club
3 Av
Mt Sinai Beth Israel
E 20th St
E 14th St
1 Av
STUYVESANT TOWN
IFC Center
Washington Square Park
Blue Note
E 8th St
Astor Place
Greene St
DŌ, Cookie Dough Confections
The Public Theater
Momofuku Noodle Bar
Film Forum
STOMP
Avenue A
Avenue B
W Houston St
E 5th St
Please Don't Tell
EAST VILLAGE
The Dutch
Kith
Apple SoHo
E 4th St
E 3rd St
E 10th St
E 9th St
Bowery
Prince St
Greenwich St
Bank St
Greenwich Ave
8th Ave
7th Ave
14 Street / 8 Av
Google


15832
The Public Theater
Astor Pla
Greene
Lafayette St
E 4th S
Google
Image capture: Sep 2016

CONCLUSION

- Bike number 15832 has been ridden, A LOT.
- Due to the centralized infrastructure of New York City, we can see patterns emerge from usage data showing that bikes tend to migrate to the center of the city over time.
- Based on the usage trends and hubs where Citibikes tend to matriculate to, could it be argued that Wash Sq. Park and NoHo are the true center of local New York ?

Thank you

Questions?

THE APPENDIX

GOALS

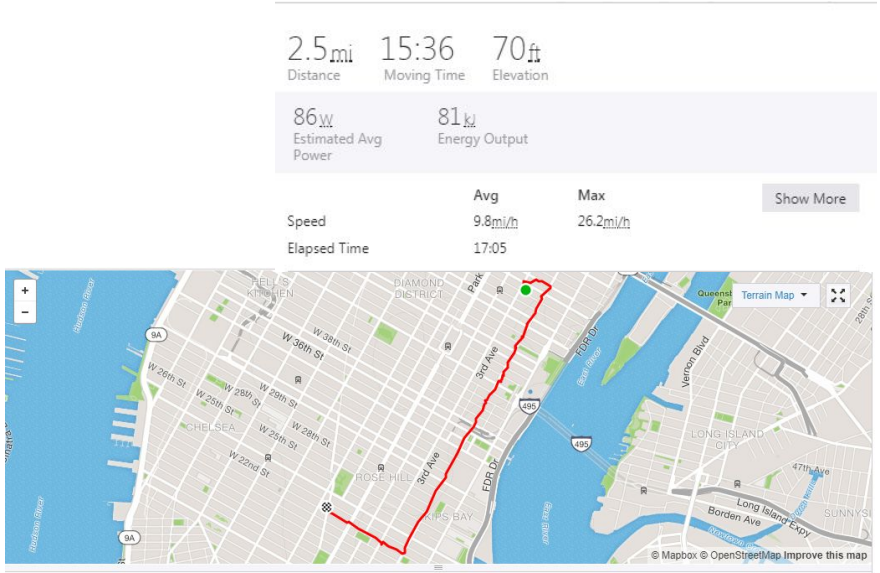
Get a full picture of what the average Citibike that you would ride throughout the city experiences and where it goes throughout it's lifetime.

See how that compares to other bikes who may have started in a different location. if that affects it's distance traveled.

METHODS

Rented a Citibike on Thursday, 11/2/2017, recorded it's bike number (15832), and rode it from 52nd and 3rd to Broadway and 22nd.

Then located the bike number within the comprehensive publicly available data of all the Citibike trips since the inception on July 1st, 2013



DATA CLEANING METHODS

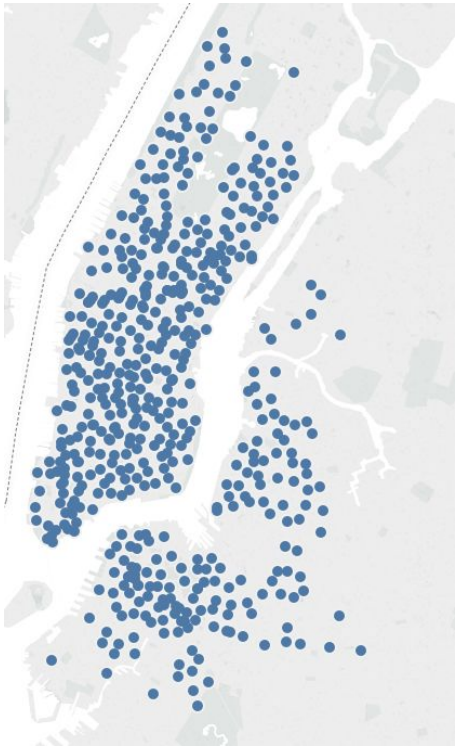
- Able to sort through each spreadsheet to pull specific bike number from each month's data. Compiled into a single spreadsheet to work from
  Pulled the sample data and compiled into separate spreadsheets
  Calculated distance of each trip (by way of the coordinates of the starting point and end point) with an excel formula (=6371*ACOS(COS(RADIANS(90-F2))*COS(RADIANS(90-J2))+SIN(RADIANS(90-F2))*SIN(RADIANS(90-J2))*COS(RADIANS(G2-K2)))/1.609)
  Used that distance to find average velocity per trip (=L2/(A2/3600))
  Removed data that had zero distance as a trip. Where essentailly someone bought a trip but didn't take the bike out. (happened 6 times)
  Calculated the age of the riders by way of (=2017-P2) and removed null values as well as people over the age of 80 since I am assuming that they are misrepresenting their age

FINDINGS

- 5121 trips before me for a total of 5,491 miles. (1.07 miles per trip)

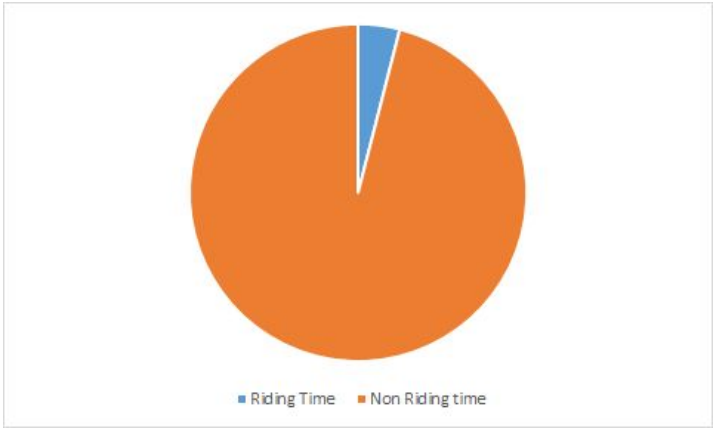
  Made a scatter plot of distance by average velocity and saw that the trend line shows that the further people go, the faster they tend to travel

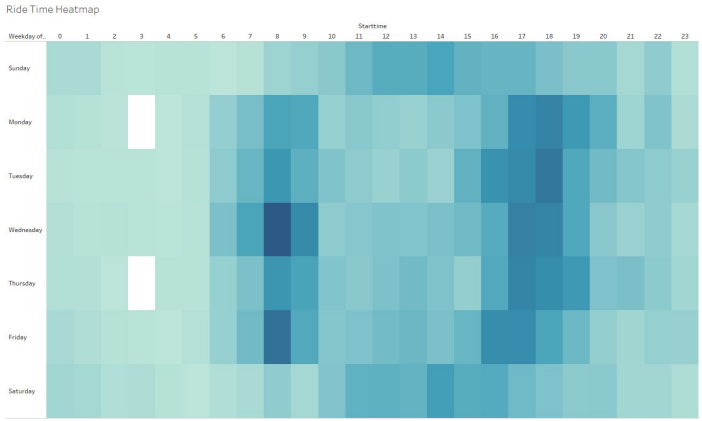
  Also saw that speeds on a city bike tend to max out at just above 10MPH. The bike covered an average of 5.40 mph per trip.

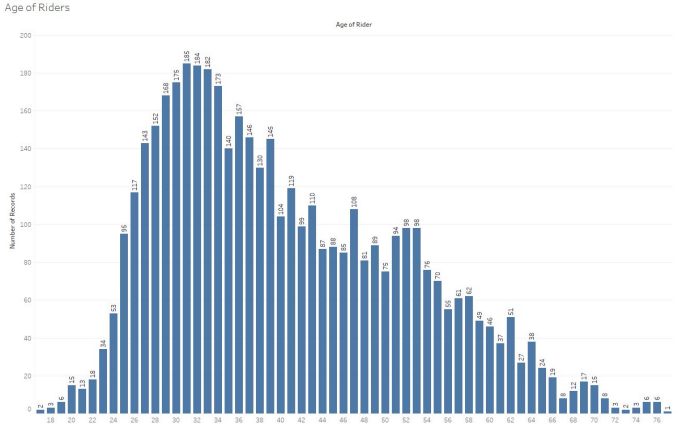


FINDINGS

- Over a span of 4 years and 2 months, the bike spent about 5285725 seconds being ridden (1,468.2 Hours)
  Average Age of our rider is 40 with a median of 38 years old
  Bike is primarily ridden on the week days during the morning and afternoon rush with a spike on the weekends during the early afternoons

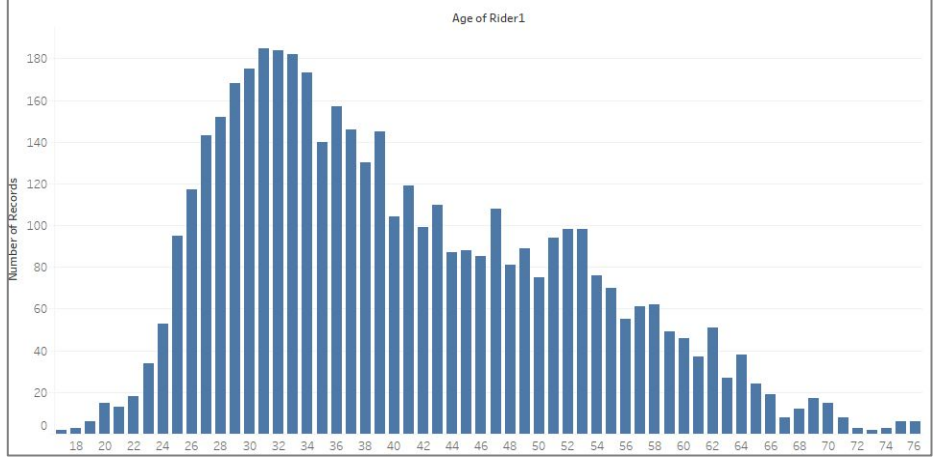


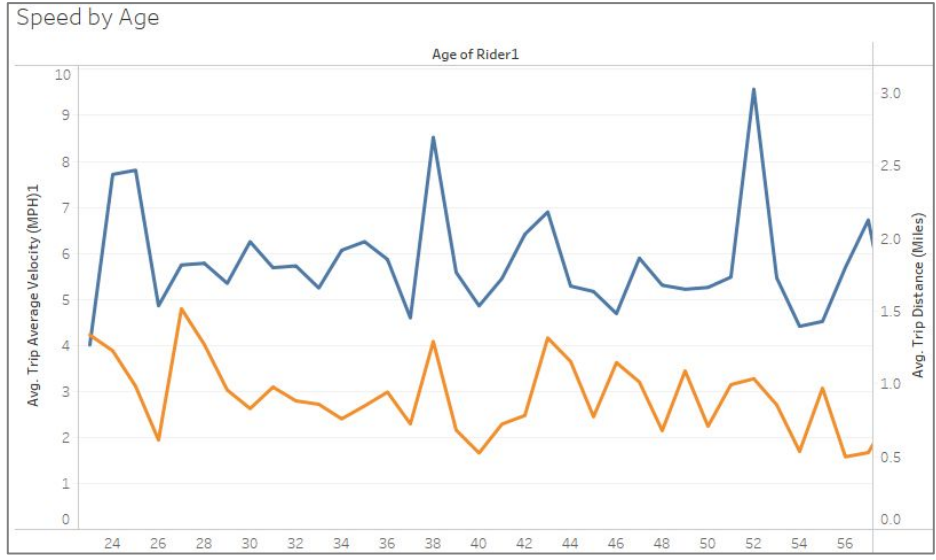




HOW FAST HAS IT GONE?

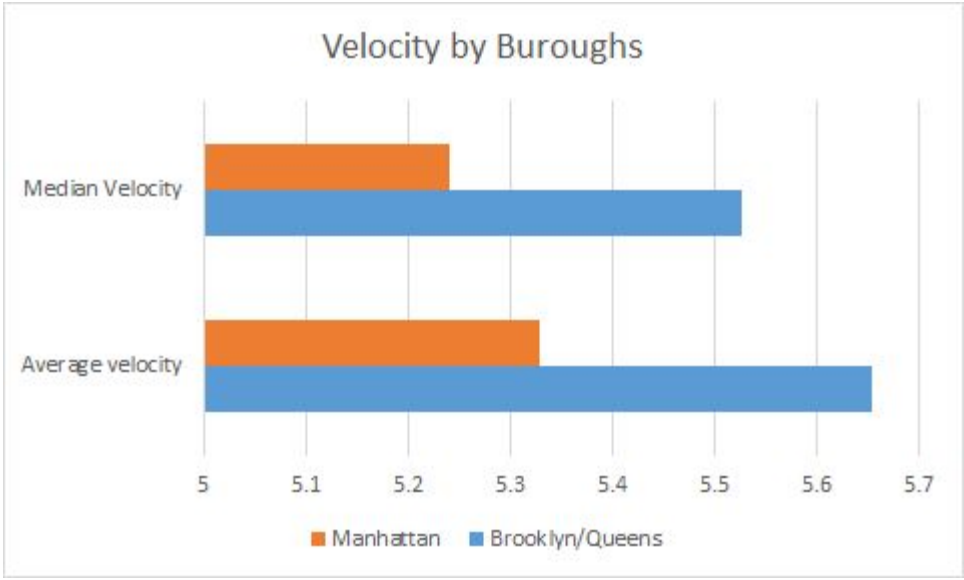
- After a quick analysis of the age breakdown of the rider, we see that the most common age range of Citibike riders falls between 28 and 33 years old.
- As people grow older, surprisingly, they do not get much slower, but maybe they will go shorter distances.





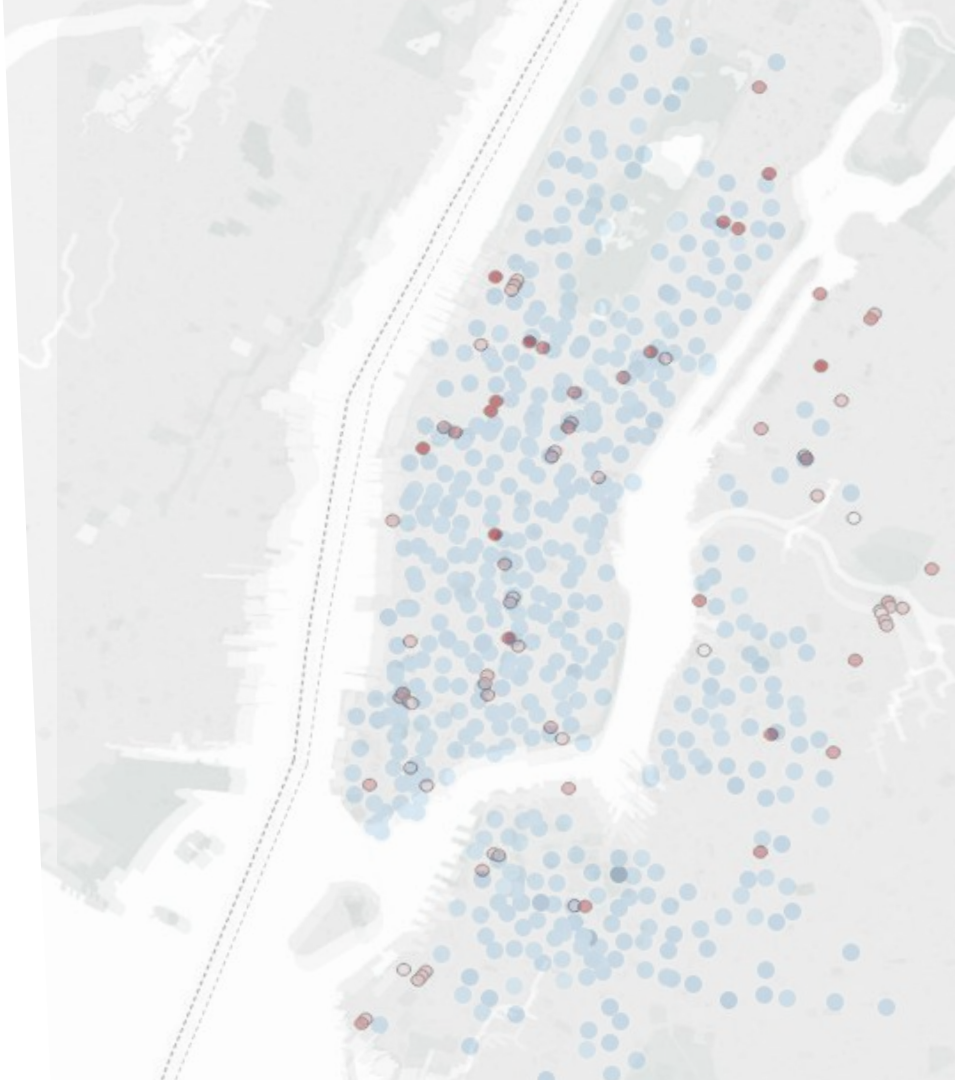
FINDINGS

- Bike is ridden faster in the outer Burroughs than it is in Manhattan



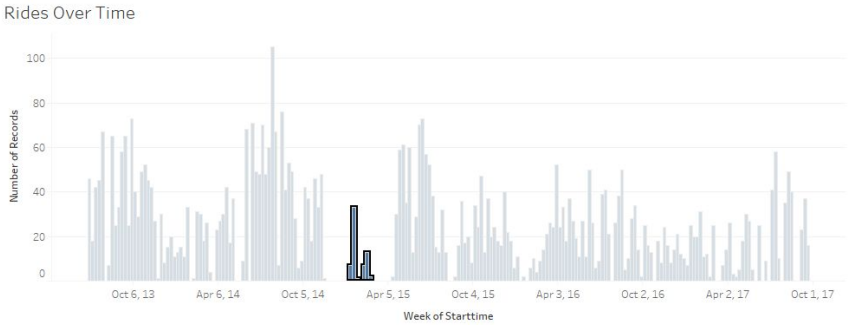
ROAD CONDITIONS

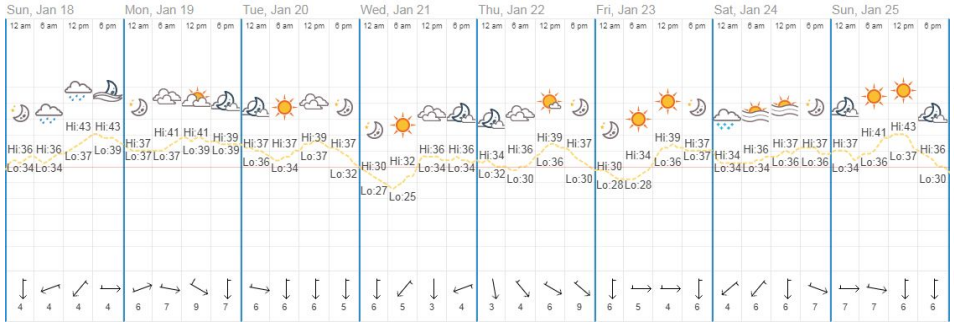
- Combining the Citibike history data with the NYC Road Condition Survey, a few of the trips have been taking place on poorly paved roads throughout the City.

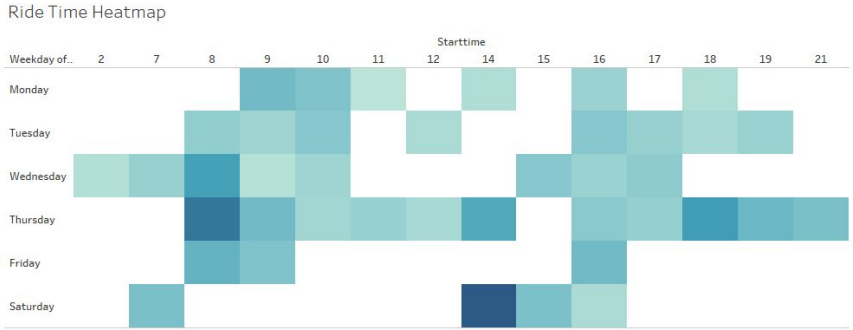


What Happened in December 2014 and January 2015

- I took a look at the lull in use over the span of December 2014 - February 2015, and there was a direct correlation between the amount of rides and the weather data over the period of time







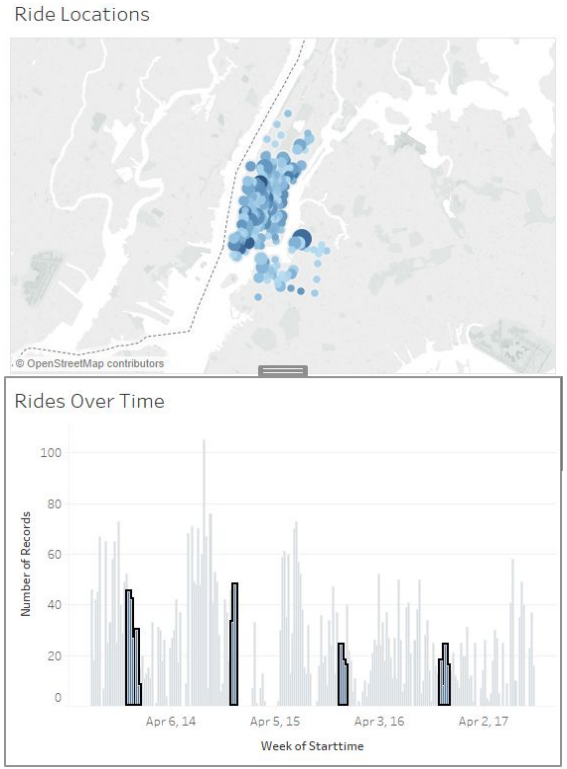
RECOMMENDATIONS

Where do we think this bike is today after a week has passed since riding it on 11/2.

First off, I did an analysis of the number of rides per week that the bike experiences and saw that the average is 27.3 with a median of 24.

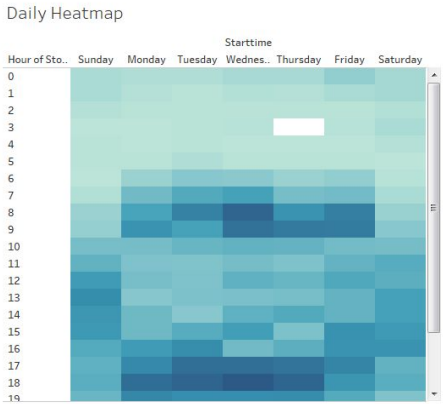
Considering that the ride took place in early November, let's factor in the amount of rides that took place during that week in previous months (29.4 average with a median of 28.5 and standard dev of 14)

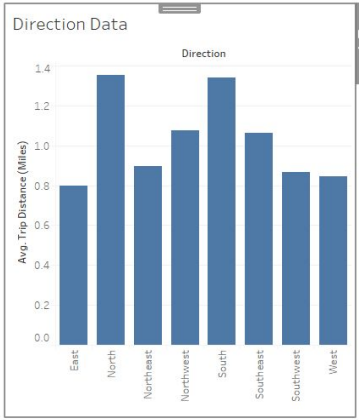
During previous years on the same time frame, the majority of the trips took place in Manhattan (167 in Brooklyn, 372 in Manhattan)

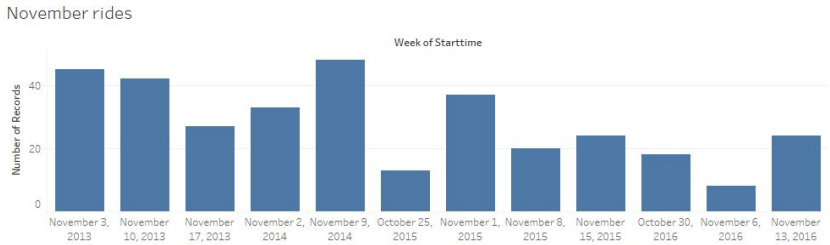


RECOMMENDATIONS

I then did some triganometry to calculate the bearing(=DEGREES(ATAN2(COS(RADIANS(F2))*SIN(RADIANS(J2))-SIN(RADIANS(F2))*COS(RADIANS(J2))*COS(RADIANS(K2-G2)),SIN(RADIANS(K2-G2))*COS(RADIANS(J2))))) Then angular bearing(=IF(N2<0,360-(-N2),N2)) then used that to calculate direction (=VLOOKUP(O2,'Trips from broadway'!$V$4:$W$11,2,TRUE)) of each trip, then I ran that in tableau daily and timely heatmaps to get a sense of when and what direction people are riding at certain points of the day.







I then created a filtered dashboard that makes a timeframe heatmap that calculates the direction and average distance traveled in each over the time frame.

RECOMMENDATIONS

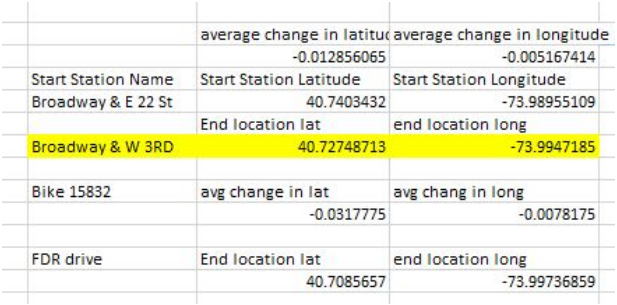
From Here, you can measure the total migration during an early November week of the bike using the average trip data by way of measuring the distance between the directions

South - 75.2 Miles / 339 trips = 0.2 Miles or 4 city blocks,
West - 3.1 Miles / 339 trips = 0.01 Miles or 0 city blocks

This would suggest that my bike is on 17th street and Broadway

standard deviation between too broad to make a confidence interval. the proportion of the bikes that went south is roughly equal to the proportion of hte bikes that went norh.

trying to predict the confidence interval, maybe focus on bikes that go shorter distances, and focus on distance from the center of the city.



	average change in latitu	average change in longitude
	-0.012856065	-0.005167414
Start Station Name	Start Station Latitude	Start Station Longitude
Broadway & E 22 St	40.7403432	-73.98955109
	End location lat	end location long
Broadway & W 3RD	40.72748713	-73.9947185
Bike 15832	avg change in lat	avg chang in long
	-0.0317775	-0.0078175
FDR drive	End location lat	end location long
	40.7085657	-73.99736859