

Improving Operation Efficiency and User Experience for Bike-sharing Services by Analyzing Trip Data

Project Proposal for The Data Incubator Data Science Fellowship

Kyle Guan

November 2019

Background, Motivations, and Project Scope

- Bike-sharing provides a green and affordable transportation alternative in many cities.
- Bike imbalance at docking station is one of the main operational challenges.
 - Bike rebalance: deploy a fleet of trucks to pick up and drop off bike; account for 30% - 50% of operating cost
- The project focuses on analyzing trip data to improve operational efficiency and user experience:
 - Data sources:
 - Historical and real-time trip data from service operators
 - Multi-modality: temporal and geo-spatial data
 - Evaluate the effectiveness of balancing approaches
 - Analyze traffic pattern and predict availability at a bike station
 - Suggest new bike station locations

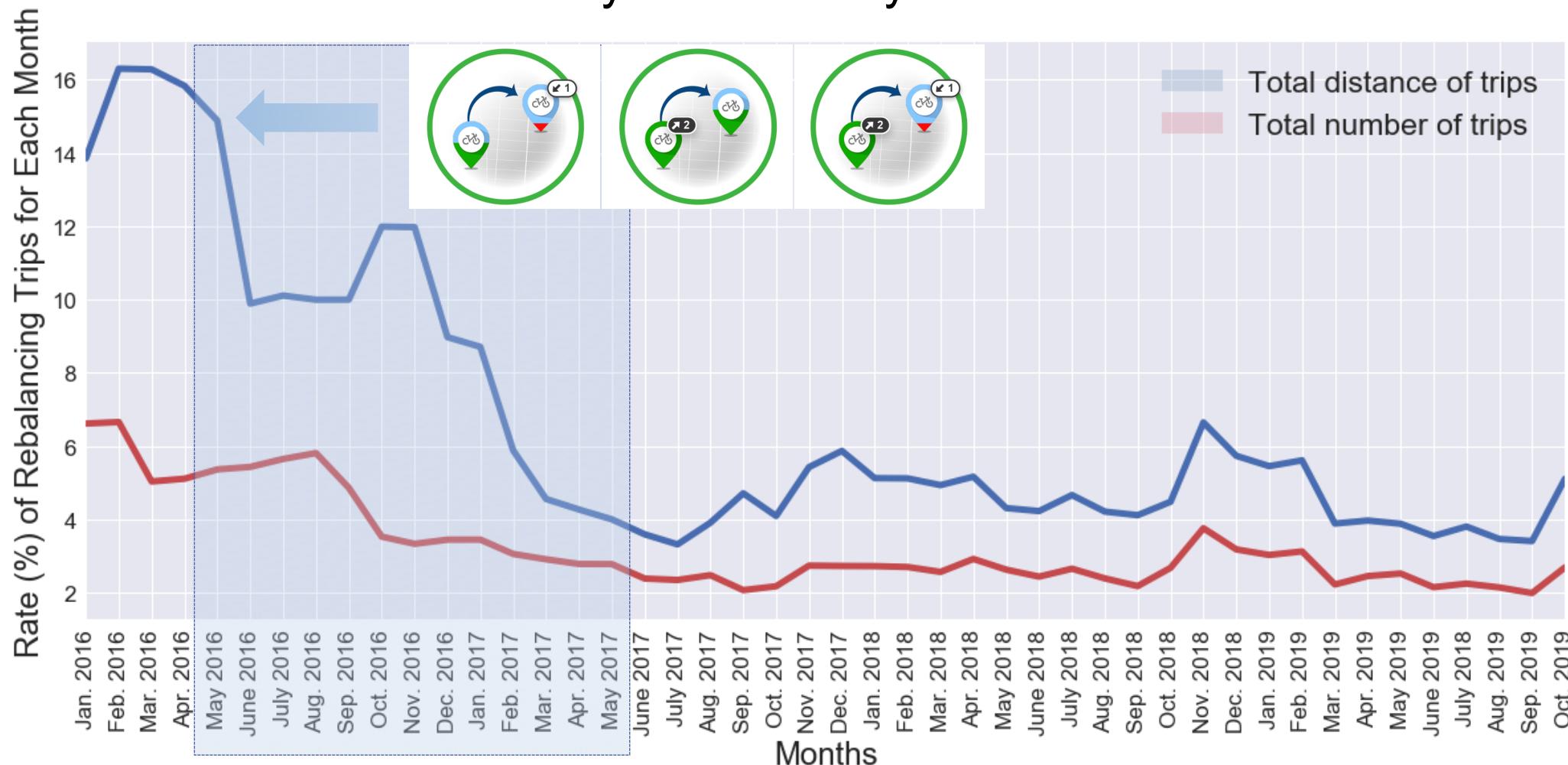


Evaluate the Impact of Bike Angels Program on Bike Rebalance

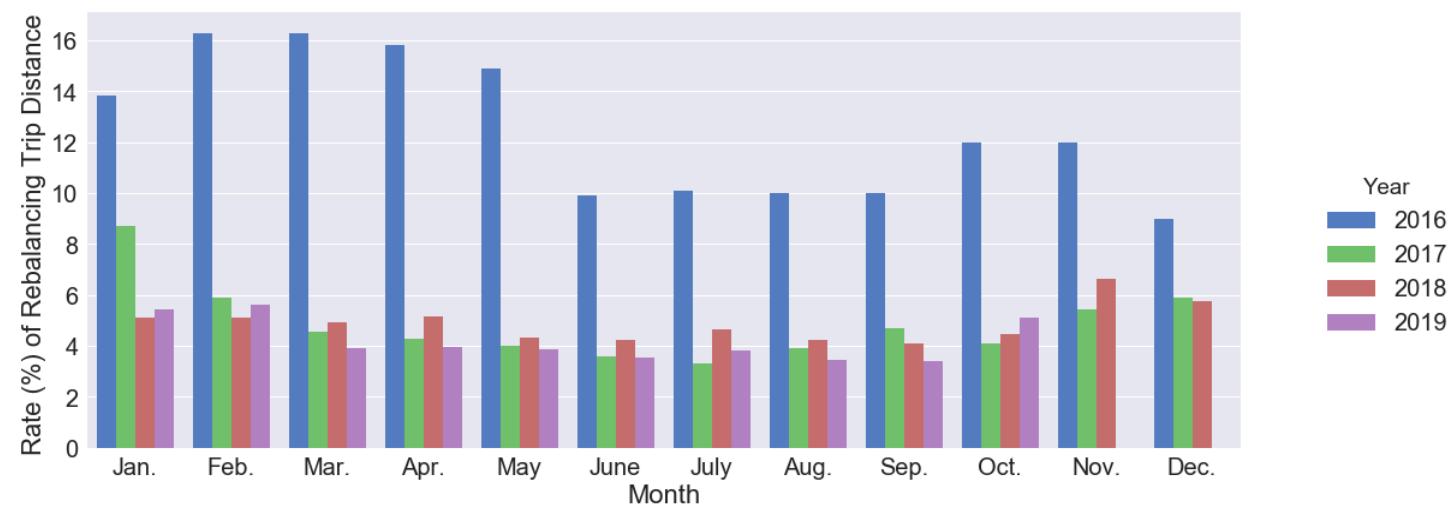
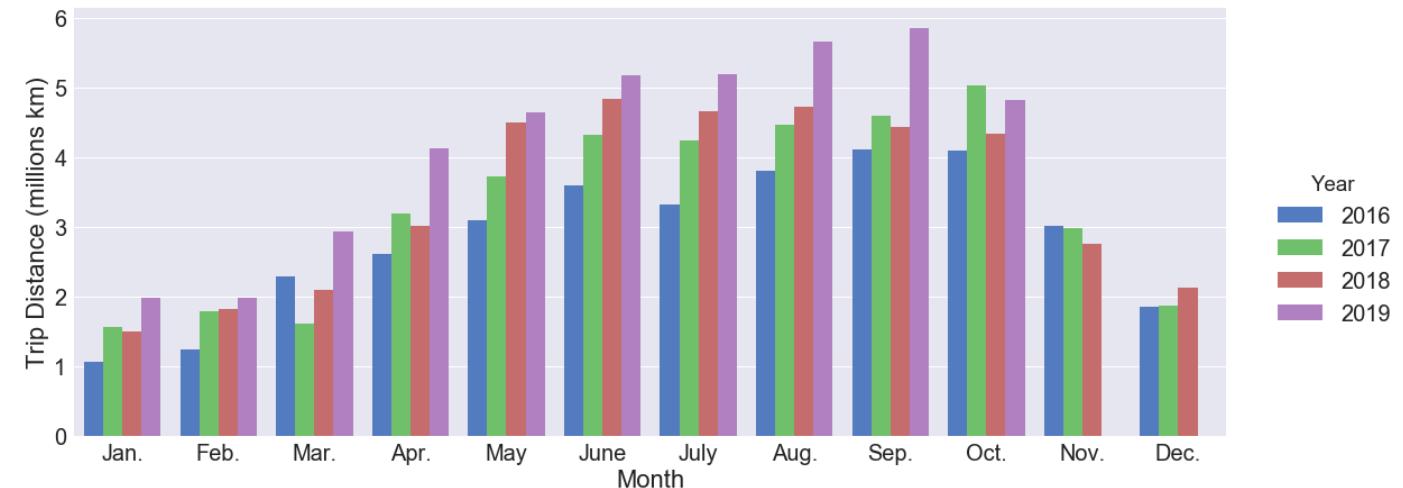
- In May 2016, Citi Bike initiated Bike Angels program to incentivize riders to redistribute bikes to in-demand areas.
- Have there been any noticeable changes in operational efficiency of bike rebalance?
- My approaches:
 - Analyzed the trips logs from Jan. 2016 to Oct. 2019 (~12 Gb CSV logs from Citi Bike)
 - Inferred and calculated the number rebalance trips and trip distances
 - Use two metrics:
 - Rate (%) of monthly rebalance trips: # of rebalance trips / # of total trips
 - Rate (%) of monthly rebalance trip distance: rebalance trip distance / total trip distance

Monthly Rate of Rebalance Jan. 2016 - Oct. 2019

- There was significant reduction in monthly rate of rebalance trips and distances from May 2016 to May 2017.

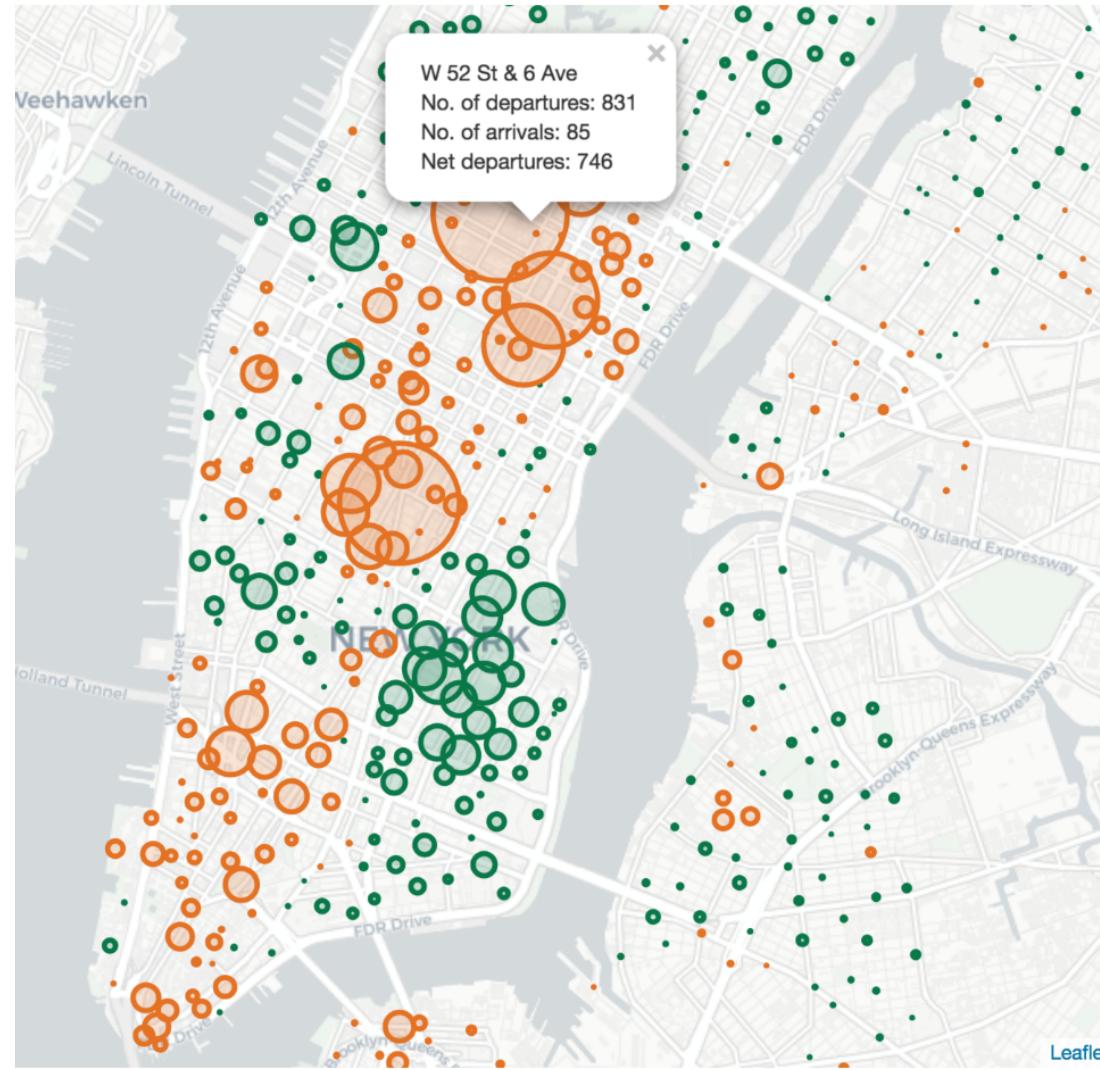


Month-to-month Trip Distance Comparison



Need to analyze other possible contributing factors, e.g. new docking stations

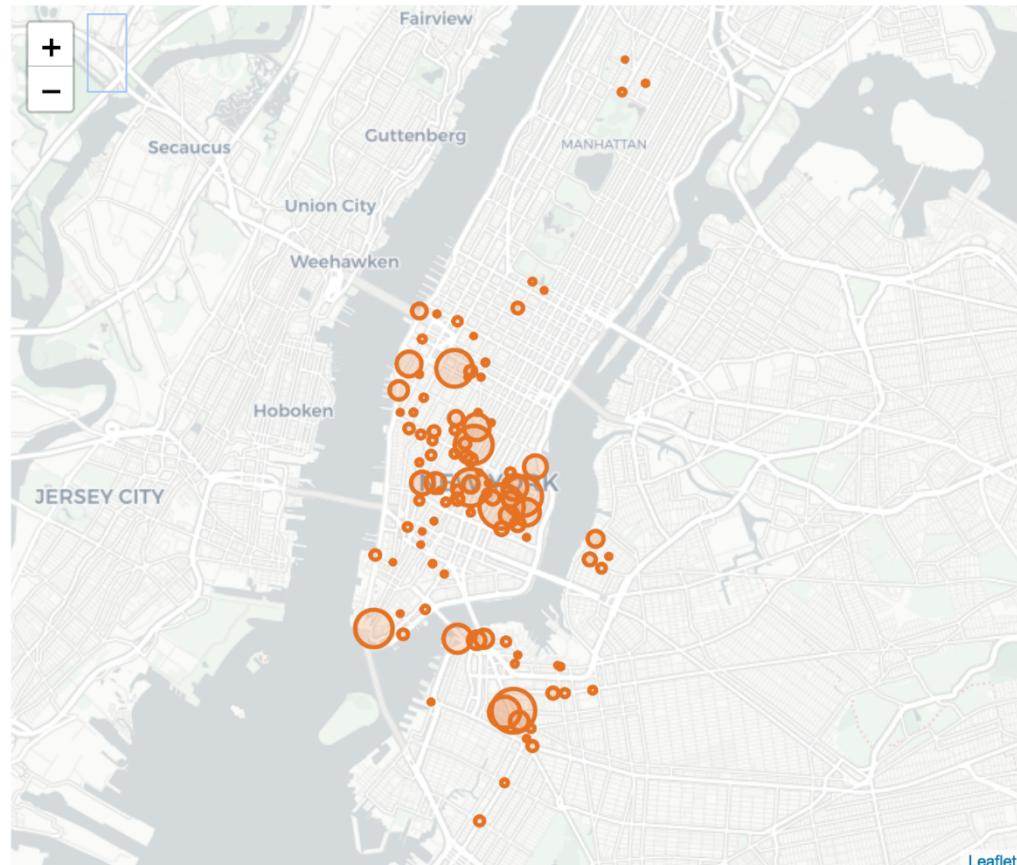
Backups



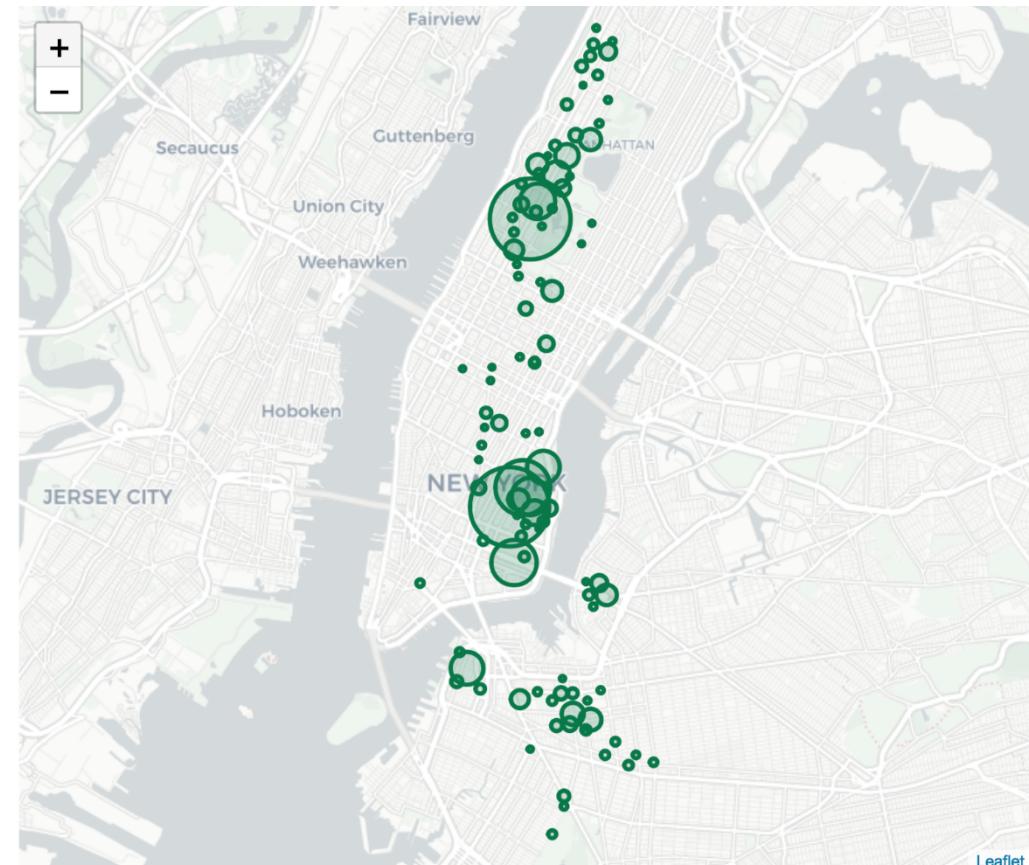
Leaflet

Rebalancing in March 2019

Top 100 start stations



Top 100 end stations



Balancing Trips Identification

- Analyze the trips of each bike and identify those where the start station is different from previous end station

Bike ID	Prev End Station ID	Prev End Station Name	Start Station ID	Start Station Name	Prev Stop Time	Start Time	Prev End Station Latitude	Prev End Station Longitude	Start Station Latitude	Start Station Longitude	
11	14529	72	W 52 St & 11 Ave	261	Johnson St & Gold St	8/1/2016 22:42:18	8/11/2016 11:09:55	40.767272	-73.993929	40.694749	-73.983625
66	14529	3168	Central Park West & W 85 St	3165	Central Park West & W 72 St	8/17/2016 13:34:04	8/18/2016 07:37:55	40.784727	-73.969617	40.775794	-73.976206
82	14529	3140	1 Ave & E 78 St	72	W 52 St & 11 Ave	8/19/2016 15:59:43	8/2/2016 07:33:39	40.771404	-73.953517	40.767272	-73.993929
101	14529	531	Forsyth St & Broome St	3140	1 Ave & E 78 St	8/2/2016 20:17:09	8/20/2016 11:53:25	40.718939	-73.992663	40.771404	-73.953517