Capstone Project
Applied Data Science Capstone by IBM/Coursera
Final Module

IMPLEMENTING ELECTRIC SCOOTER SHARING SYSTEM BUSINESS IN MANHATTAN
DATA REPORT
JUNE 17, 2020

TABLE OF CONTENTS

IMPLEMENTING ELECTRIC SCOOTER SHARING SYSTEM BUSINESS IN MANHATTAN
DATA REPORT
JUNE 17, 2020

1. INTRODUCTION:
Background and Business Problem description

Citi Bike is a privately owned public bicycle sharing system serving the New York City boroughs of the Bronx, Brooklyn, Manhattan, and Queens, as well as Jersey City, New Jersey. ***It officially opened in June 2013 with 332 stations and 6,000 bikes***. Annual expansions have brought the totals to 706 stations and 12,000 bikes as of October 2017, making the service the largest bike sharing program in the United States. Further expansions for Citi Bike are planned to extend its service area across the Bronx, Brooklyn, Manhattan, and Queens, and increase the number of bikes to 40,000. As of July 2017, there are 130,000 annual subscribers. Citi Bike riders took an average of 48,315 rides per day in 2018, and the system reached a total of 50 million rides in October 2017.

Bicycle Access to Office Buildings Law ("BAOB law") aims to increase bicycle commuting by providing cyclists with the opportunity to securely park their bicycles in or close to their workplaces. ***The law allows tenants to request for access from landlords in commercial buildings, and for DOT to monitor the requests for access and their status***. The law was enacted in 2013.

Amid Covid-19 emergency, and the subsequent foreseeable increase of sharing transportation system use by New Yorkers as well as the foreseeable reduction of the use of the subway system, New York City Department of Transportation is intending to issue a public bid for the provision of 200 electric scooters and the installation of 2 pick-up stations in Manhattan. The project, after its initial "Beta" phase, will follow the provision of other 2,000 scooters and the installations of new stations.

Scooz, Inc., a Colorado manufacturer of electric scooters, is intending to participate to the bid. The executive team, together with the commercial team, engaged us to provide a data report illustrating the best 5 Manhattan neighborhood ***areas where to install the 2 pick-up stations required by the bid***.

The CEO of the company has already released the following statement revealing the intention of the company to participate in the bid: "Cities around the world are embracing e-scooters as an environmentally-friendly, inexpensive way to get around, especially in transit deserts. New Yorkers are ready. We are encouraged with the overwhelming support we received from the legislature and from the environmental community. We look forward to bringing micro-mobility choices to New Yorkers next year."

The company has specifically pointed out that the pick-up stations are to be selected upon the following criteria:

(1) neighborhoods where most used City Bike stations in Manhattan in the month of June 2013 (when the City Bike project was launched) are located;

(2) commercial buildings that received major number of requests from tenant for bike inside parking; and

(3) the presence of restaurant, gym and coffee shop as most common venues in neighborhoods identified under (1) and (2). With regard to the latter point, the company wants to provide, on one side, New Yorkers with the opportunity to reach and/or move around these places through the electric scooters, and the other side, find partnership and advertising opportunities through such existing businesses in the areas.

2. DATA PREPARATION [DATA SECTION]:

1. ***Dataset 1:*** This New York City Neighborhood Names point File publicly available at https://geo.nyu.edu/catalog/nyu_2451_34572
2. ***Dataset 2:*** City Bike Trip Data History related to the month of June 2013 publicly available at https://s3.amazonaws.com/tripdata/index.html
3. ***Dataset 3:*** Bikes in Buildings Requests NYC File publicly available at https://data.cityofnewyork.us/Transportation/Bikes-in-Buildings-Requests/scjj-6yaf/data
4. ***Dataset 4:*** Foursquare Manhattan Most Common Venues Data through our corporate credentials.
5. ***Dataset 5:*** City Bike Trip Data History related to the months June 2013-June 2015 publicly available at https://s3.amazonaws.com/tripdata/index.html

Note: We have attached the link to each dataset used in our Notebook. We uploaded the same dataset from our internal serve in the Notebook for convenience.

3. METHODOLOGY

For this report used different maps that could help the Company to locate two Manhattan neighborhood areas where to install the electric scooters pick-up stations in accordance with the NYC public bid:

1. We have used ***Dataset 1*** to build a preliminary dataset of longitude and latitude coordinates of each Manhattan neighborhoods to be used to build the map identified under par. 2 and 3 below.

2. We aimed to identify the most used start and end City Bike stations in the Manhattan neighborhoods using ***Dataset 2*** to build a map providing markers of such stations in the map.

3. NYC Open Data website offers a dataset that includes all requests made by commercial buildings tenants made under the BAOB law. We regard this dataset would be useful for our analysis. We analyzed this dataset identified as ***Dataset 3*** to locate the Manhattan neighborhoods where commercial buildings that received more tenants' request to park bikes inside under the in a five years time frame from 2013 are situated.

4. We have built a boxplot showing the flow of the requests made under BAOB law in Manhattan commercial buildings in connection to ***Dataset 3***.

5. We have used ***Dataset 4*** Foursquare data to display the current venues of each identified Manhattan neighborhoods under par. 2 and 3 above and understand if the most common venues

in such locations are restaurants, gyms and coffee shops in connection with the needs specified by the client (i.e. the client wants to install the stations in areas where the most common venues are such businesses).

6. We used ***Dataset 5*** to prepare a LSTM regression model showing the single accesses amount predictions. Note that we used the first 24 months of Citybike 24h accesses information in order to build our model since the "Beta" nature of the project. We indeed need to understand how the numbers will be after the first two years of project implementation, and particularly, if the accesses will maintain constant during a predicted 1 year time frame. Note that we cannot extend our Citybike model on which we build the predicting model as only partial information as to the year 2015 are available.

4. THE CODE

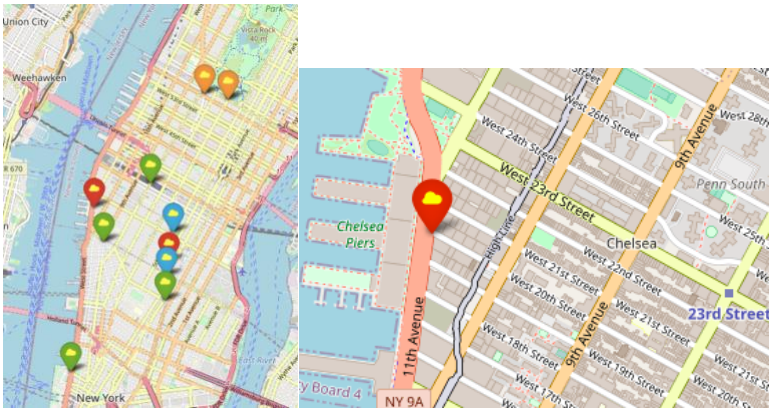
[Notebook showing the code used to perform the exploratory data analysis described under the Methodology Section above]

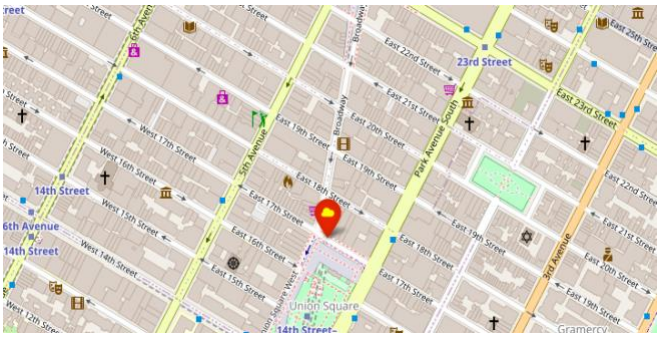
5. RESULTS

(A) Most used start and end Citybike stations in Manhattan during June 2013

We have identified the most used start and end Citybike stations during the month of June 2013 when the Citybike project was initially launched.

The two most used start stations were: (1) W 20 St & 11 Ave with 5983 counted starting trips; and (2) E 17 St & Broadway with 5621 counted starting trips.

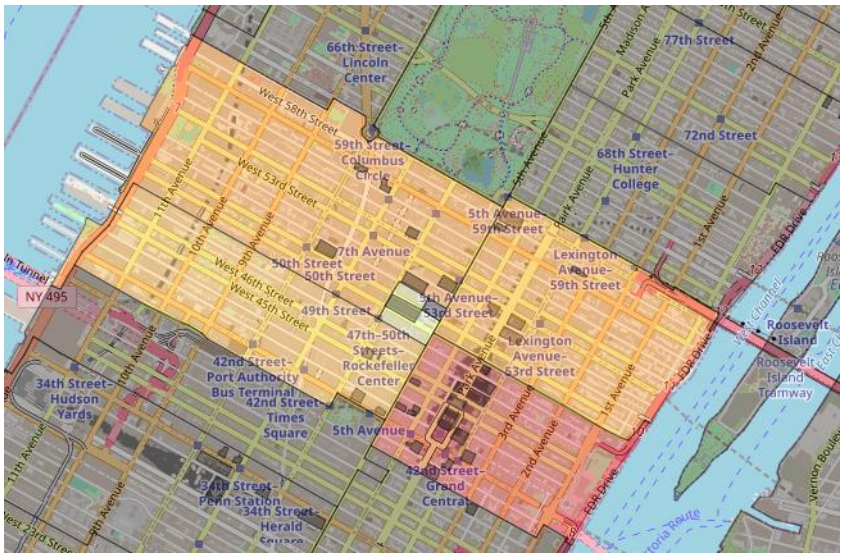




The two most used end stations were: (1) W 20 St & 11 Ave with 5742 counted starting trips [**Neighborhood: Greenwich Village**]; and (2) E 17 St & Broadway with 5530 counted starting trips [**Neighborhood: Flatiron**].

(B) neighborhoods where commercial buildings that received major number of requests as per BAOB law

We have identified neighborhoods where commercial buildings that received major number of requests from tenant to park bikes inside under the BAOB law are located. The result below shows that Turtle Bay, Tudor City, Midtown and Midtown South (*partially*) is the area mostly involved by commercial buildings requests to park their bicycles inside.

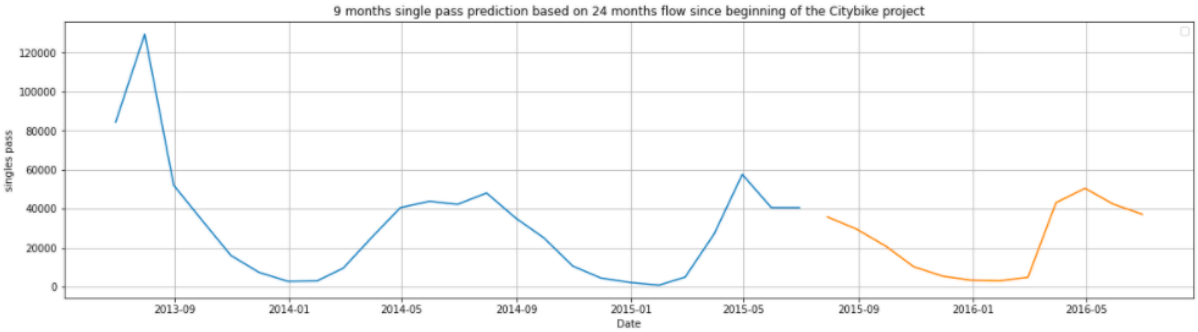


(C) Are restaurants, gyms and coffee shops the most common venues (under a top 5 most common venues scale) in neighborhoods identified under (A) and (B) above?

The answer is Yes for two selected neighborhoods. The Foursquare Data below shows that Flatiron and Tudor City are the closest with respect to the third criteria identified by the Company for the selection of the neighborhoods where to install the two pick-up stations. Gyms, Coffee Shops and Restaurants are indeed the most common venues in such areas.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Flatiron	Gym / Fitness Center	Café	Mediterranean Restaurant	New American Restaurant	Coffee Shop	Italian Restaurant	Gym	Park	Vegetarian / Vegan Restaurant	Japanese Restaurant
1	Greenwich Village	Italian Restaurant	Café	Sushi Restaurant	Clothing Store	Ice Cream Shop	Chinese Restaurant	French Restaurant	Bar	Pilates Studio	Bubble Tea Shop
2	Midtown	Hotel	Coffee Shop	Clothing Store	Bakery	Theater	Cuban Restaurant	Pizza Place	Steakhouse	Indian Restaurant	Tailor Shop
3	Midtown South	Korean Restaurant	Hotel	Japanese Restaurant	Café	Lounge	Burger Joint	American Restaurant	Cocktail Bar	Spa	Hotel Bar
4	Tudor City	Café	Park	Mexican Restaurant	Sushi Restaurant	Gym	Garden	Dog Run	Diner	Thai Restaurant	Deli / Bodega

(D) The LSTM prediction model shows a potential constant regression of the single accesses to the electric scooters. The result is however partial for reasons explain under Section 3 par. 6 above.



6. DISCUSSION

We have been asked to identify pick-up stations for electric scooters based on the following criteria: (1) neighborhoods where most used Citybike stations in Manhattan in the month of June 2013 (when the Citybike project was launched) were located; (2) neighborhoods where commercial buildings are inclined to have inside bikes parking; and (3) the presence of restaurant, gym and coffee shop as most common venues in neighborhoods identified under (1) and (2).

We noted that under (1) the pick-up stations should be located in Greenwich Village and Flatiron which the most used Citybike stations were located in the month of June 2013 when Citybike project initially started.

We noted that under (2) we could use the Dataset 3 which include the list of commercial buildings which received tenants' request to park inside the buildings under the BAOB law of NYC. We then identified that the neighborhoods in which commercial buildings that received most of such requests were located in Midtown, Midtown South, Tudor City, and Turtle Bay. We also noted that after the enactment of the BAOB law, during a time-frame of 5 years, such buildings received most requests during the second year after the enactment of the law (2010). If a law similar to the BAOB law would be enacted by NYC for electric scooters, this information should be evaluated by the Company in two main perspectives: (i) installation of further stations in the areas identified under (2) to the extent that the company should expect an issuance of tenants request for parking electric scooters inside the building; and (ii) potential electric scooters private lease agreements with the commercial buildings of such areas.

We finally noted under (3) that Tudor City and Flatiron are the neighborhoods where gyms, restaurant and coffee shops are collectively included in the top 5 common venues.

We used the Time Series Prediction with LSTM Recurrent Neural Networks to predict single passes access to the electric scooters based on 24 months Citybike 24h accesses in the first 21 months after the launch of the Citybike project (from June 2013 to June 2015). We noticed that the predictions look pretty much constant with respect to the previous 24 months accesses with certain peaks in the spring and summer months.

7. CONCLUSION

In conclusion, Tudor City and Flatiron are most suitable areas for stations installation according to the criteria given by the company to prepare our data exploratory analysis.

8. STEPS FORWARD

A. Accuracy of the models shown above has room for improvement in terms of identification of more specific areas in the maps and/or identification of neighborhood borders.

B. Mapping Manhattan bike lanes routes and merged this map with each map shown above with transparency mode to give further insights on the most convenient locations for pick-up stations.

C. Analyze MTA data for commuting density data from Upper East Side and Upper West Side (identified as typical residential areas) and compare with eventual Citybike trips in these areas to show potential installation of stations in these areas.

D. Business-plan preparation: identify price range for different types of membership in comparison to Citybike most recent applied prices.

E. Applied different regression models to predict future single passes subscriptions based on Citybike available data or similar scooter sharing models