**Data Mining Taxi Trips in Brooklyn**

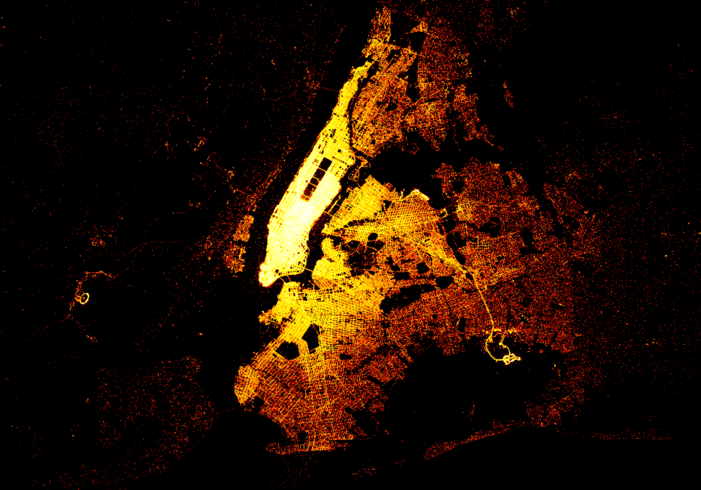
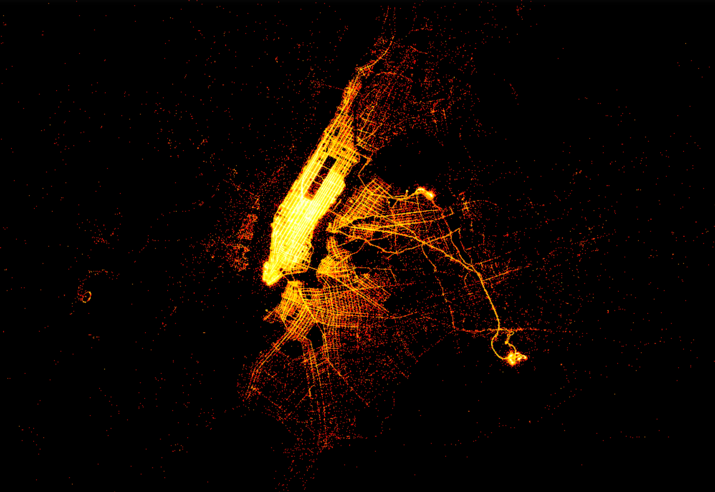
**Project Description:**

* To understand the usage and characteristics of yellow taxi in the borough of Brooklyn in October 2015.
* To understand the relationships between people’s behaviors of taking yellow taxis and different zoning features, including Residential Districts(R), Commercial Districts (C), Manufacturing Districts (M), Park Areas (P).
* To visualize the typical taxi trips and identify the populous/typical trips in Brooklyn.

**Analysis Data：**

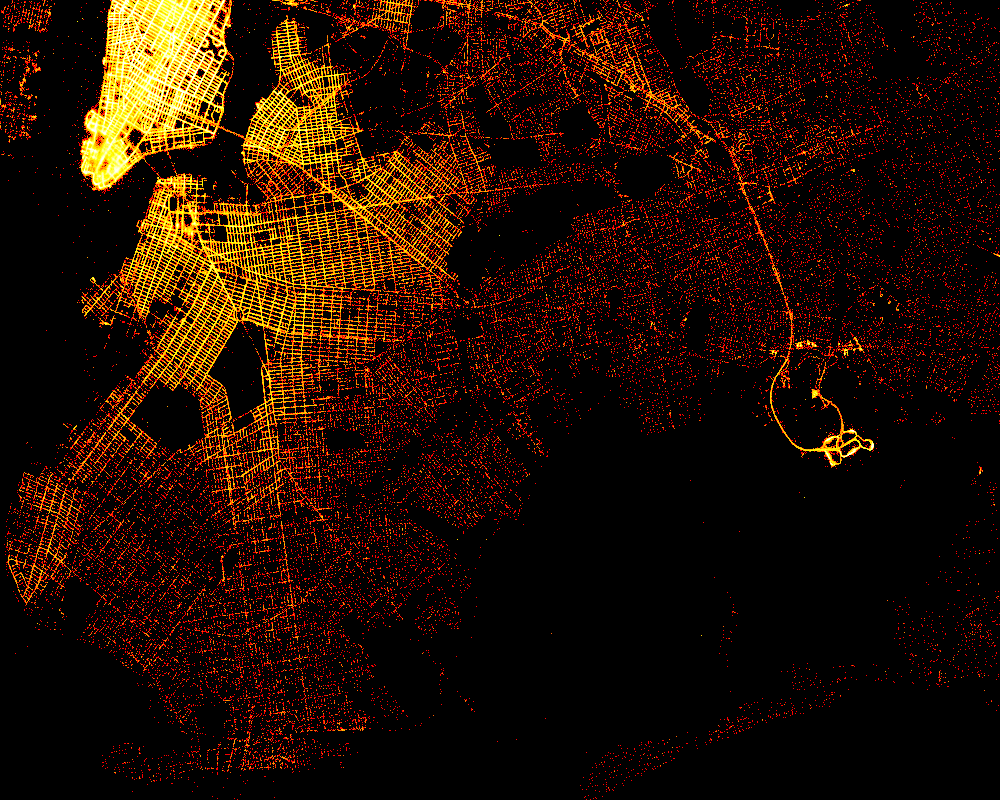
* Yellow cab’s data of the New York City in October 2015 consists of trip information like cab pick-up and drop-off time and coordinates, passengers count, trip distance, and total fare amount. (2GB)
* New York City Department of City Planning. 2018. NYC GIS Zoning Features
* Taxi Zones

**How people in New York travel across the city by yellow taxis in October 2015**

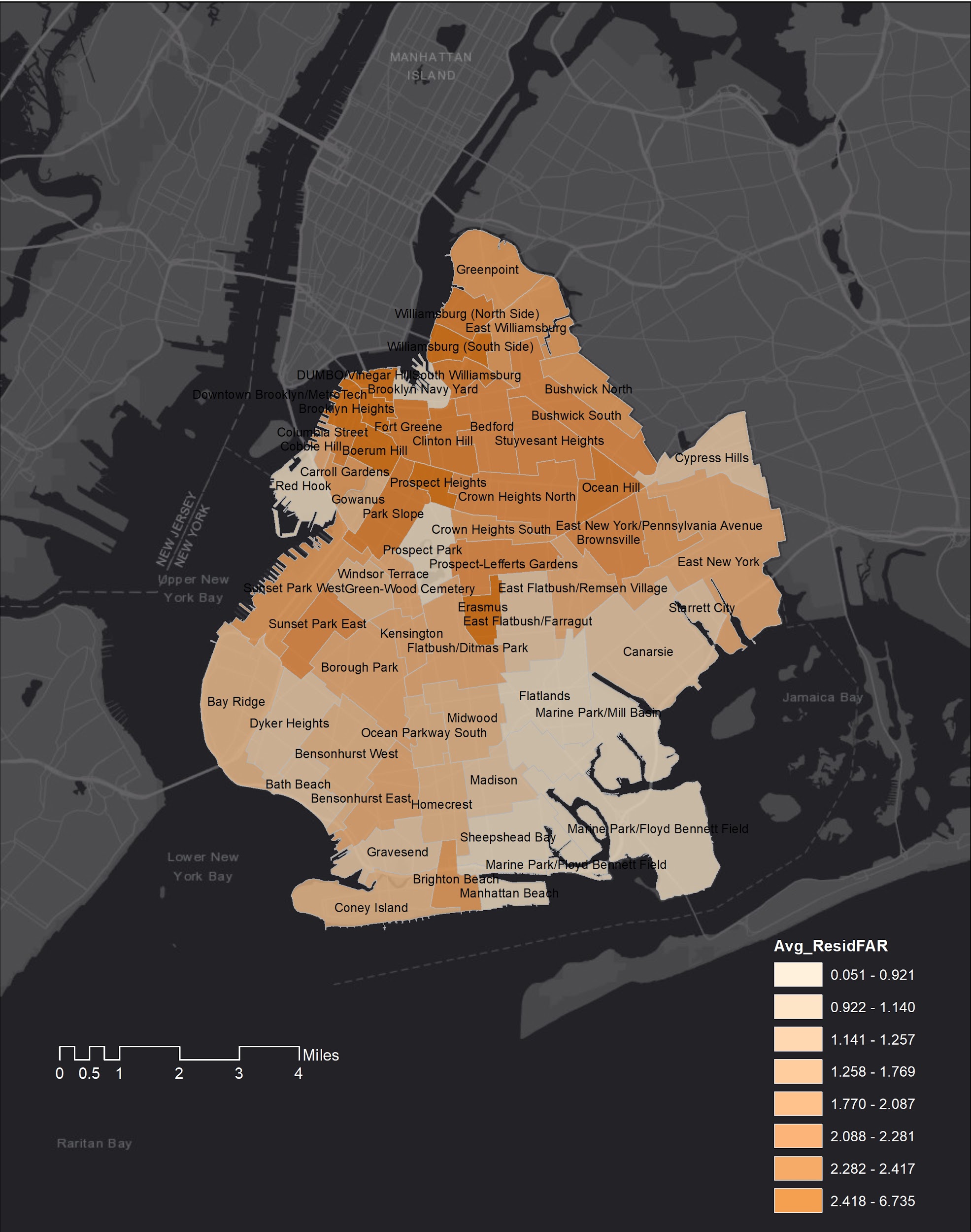
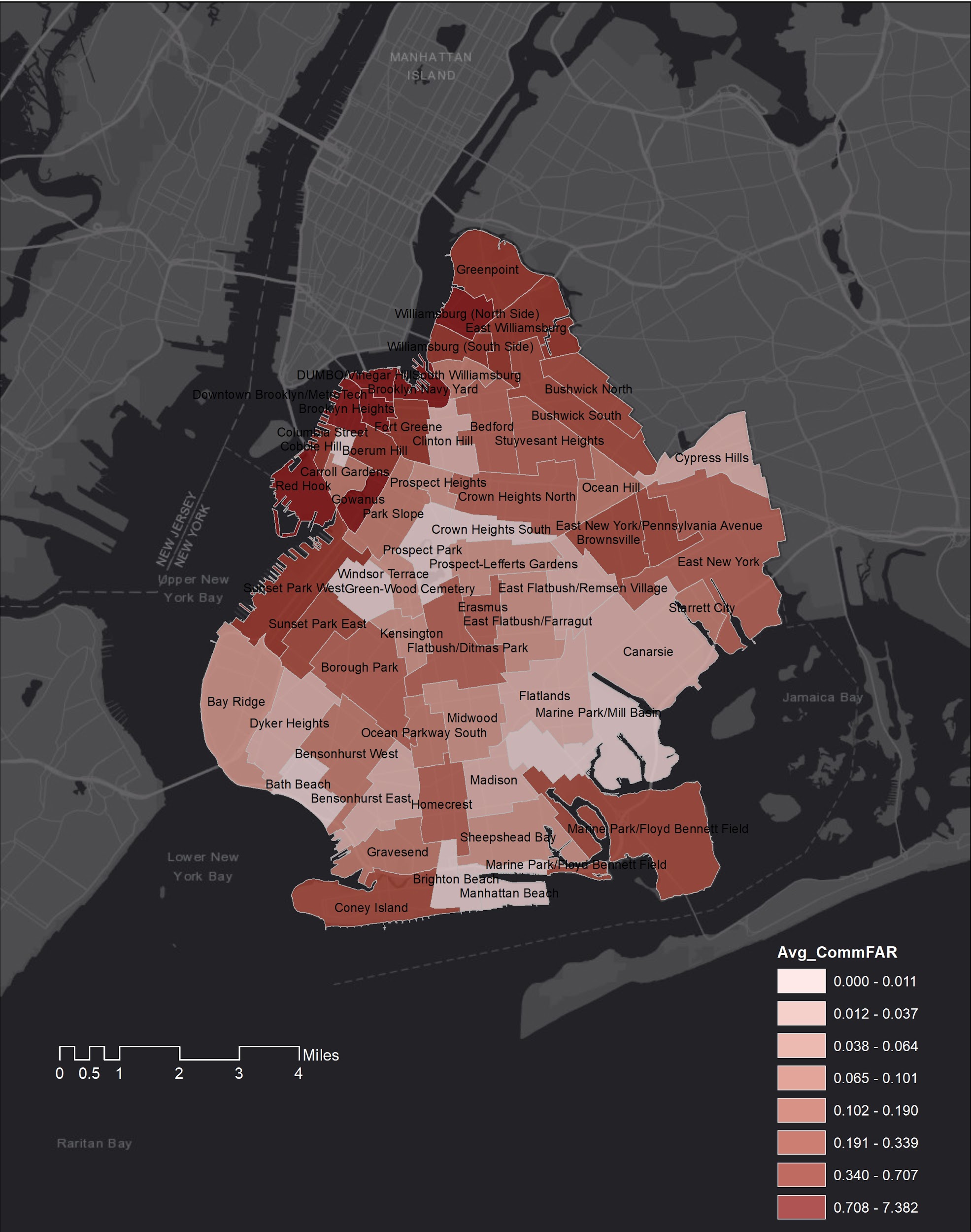


Manhattan, as an extremely compact, highly mixed-use and traffic-concentrated area, is too special for our analysis

**Analysis site zoomed to Brooklyn, NY**

*  Pick-up trip map and Drop-off trip map have very different patterns.
* Brooklyn has much more drop-off trips than pick-up trips - people are more likely to take taxis to Brooklyn, while few of them take taxis from Brooklyn: the yellow taxis are mainly concentrated in the borough of Manhattan
* Most of the pick-up trips in Brooklyn in October happen within 5 miles around Manhattan and area close to John F. Kennedy International Airport.
* Drop-off trips, instead of gathering in certain area, scatter around the entire Brooklyn.

**Commercial FAR and residential FAR**



* Generally speaking, commercial FAR and residential FAR show the same pattern in Brooklyn: taxi zones with higher commercial FAR are also tend to have higher residential FAR.
* Northern Brooklyn has higher commercial and residential FAR than any other area in Brooklyn: Northern Brooklyn is relatively denser area in Brooklyn
* Therefore, we assume that denser places with more population would generate more taxi trips. This is furthered analyzed in the next part when using Spearman's rank correlation analysis.

**Spearman's rank correlation analysis**

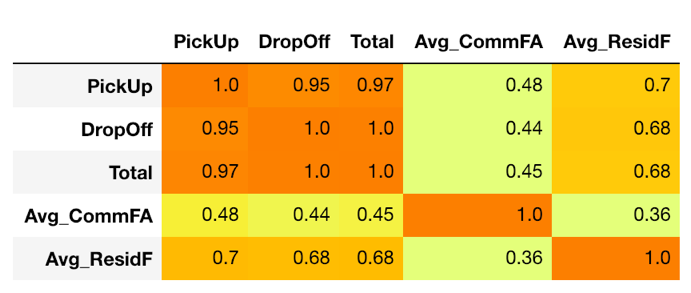


Figure Correlation\_Analytics\_BK\_all

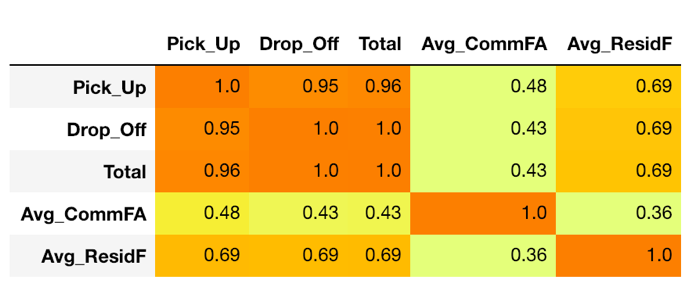


Figure Correlation\_Analytics\_BK\_weekend

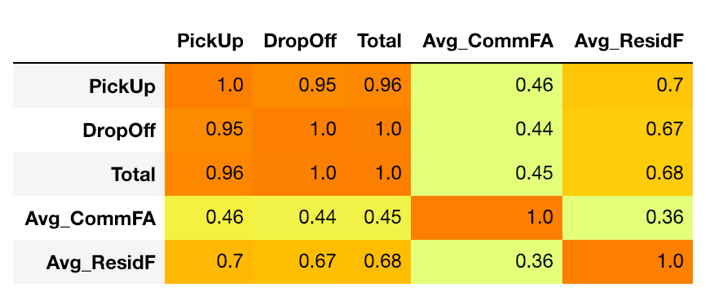
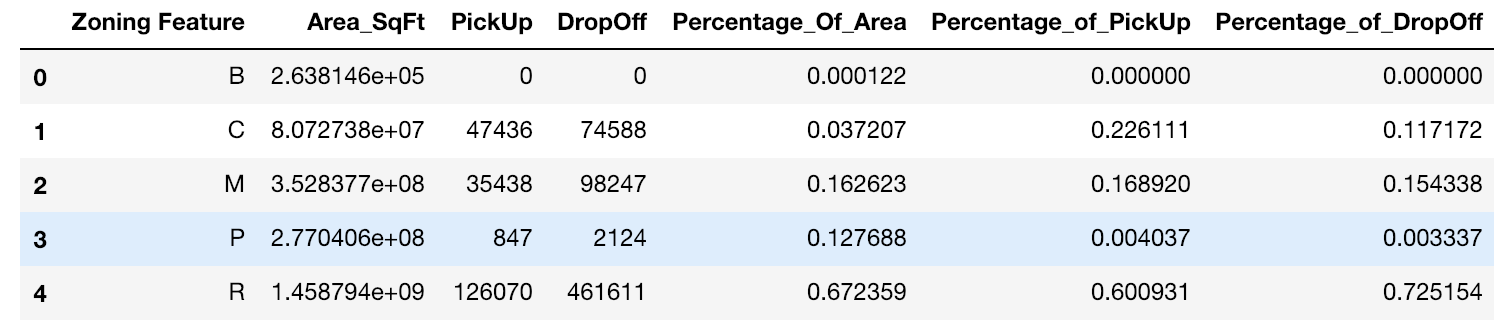


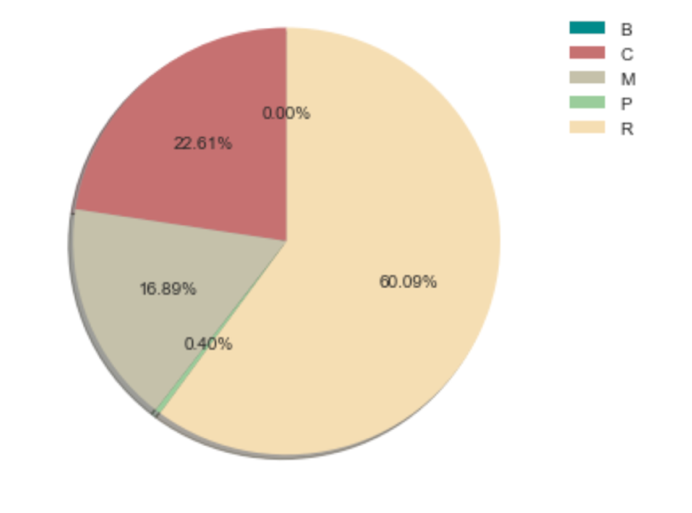
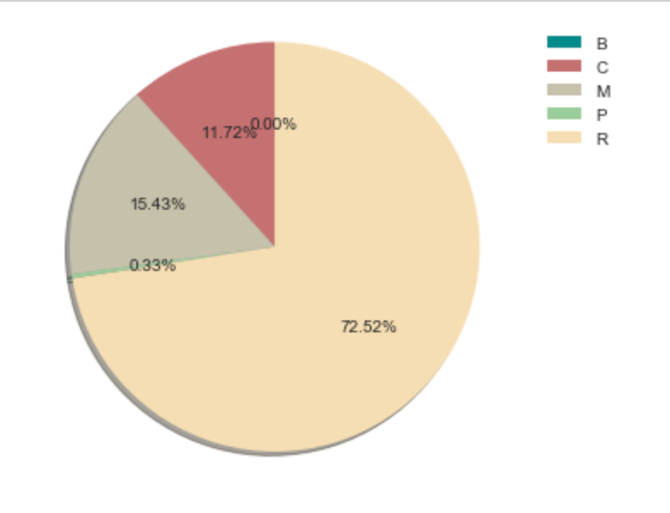
Figure Correlation\_Analytics\_BK\_weekday

* No matter for overall, weekend or weekday taxi trips in Brooklyn in October 2015, they all show positive relationships with the Average Commercial FAR and Average Residential FAR calculated based on taxi zone, with slight difference.
* High average residential FAR of a taxi zone is more related to high volume of taxi trips in Brooklyn than high average commercial FAR.

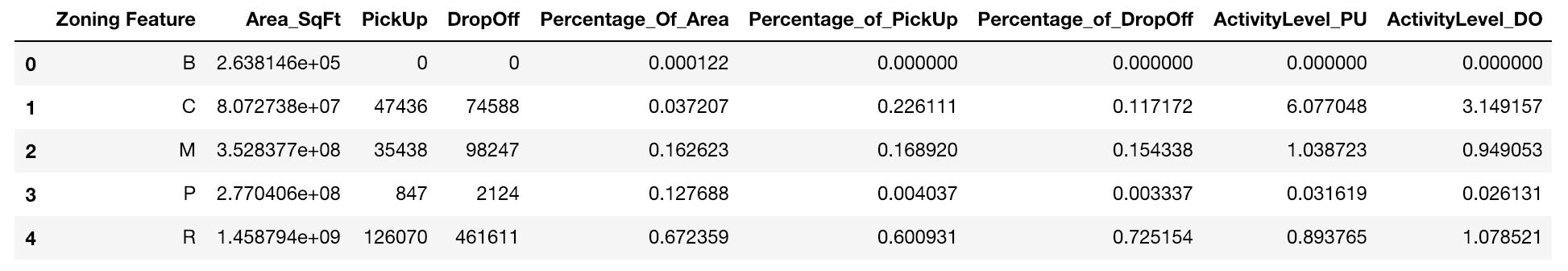
**Land Uses**

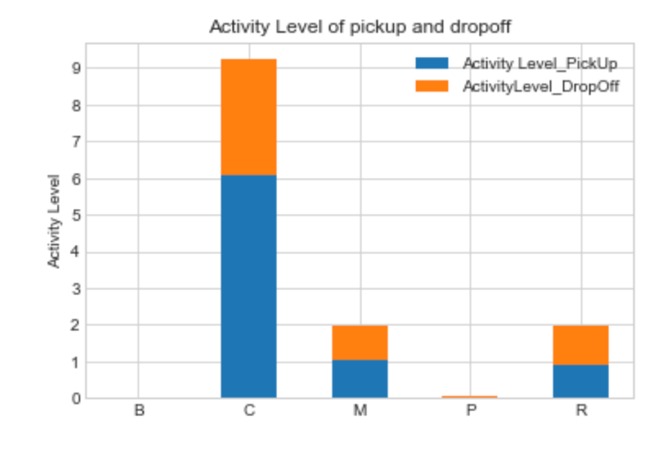


* 16.26% of Brooklyn are manufacturing districts, 12.77% are park area, 3.72% are commercial districts, while most of the area are residential districts that account for 67.23% of the total area in Brooklyn.

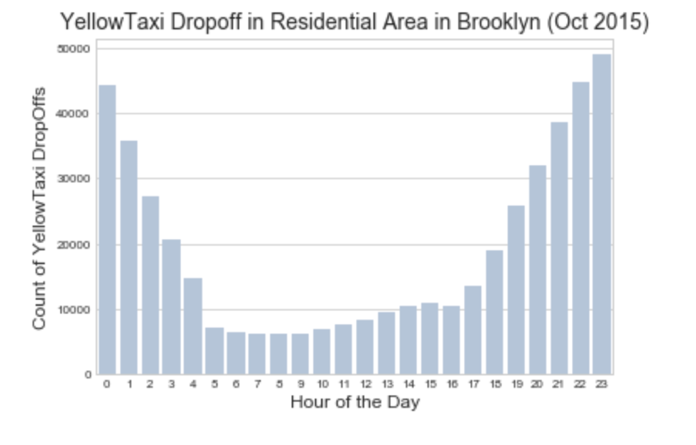
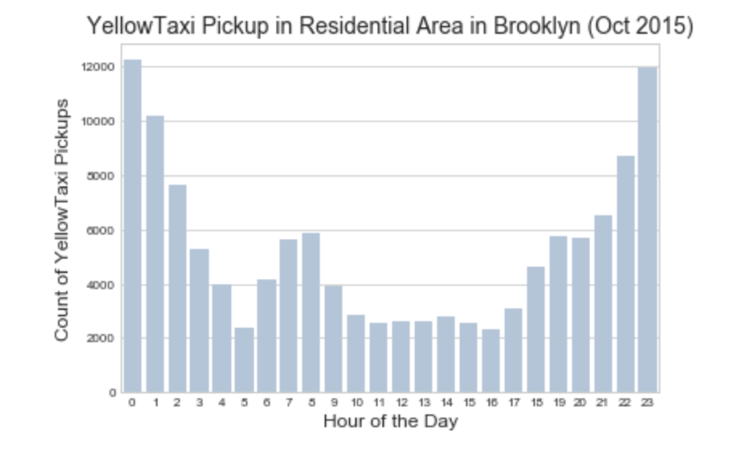
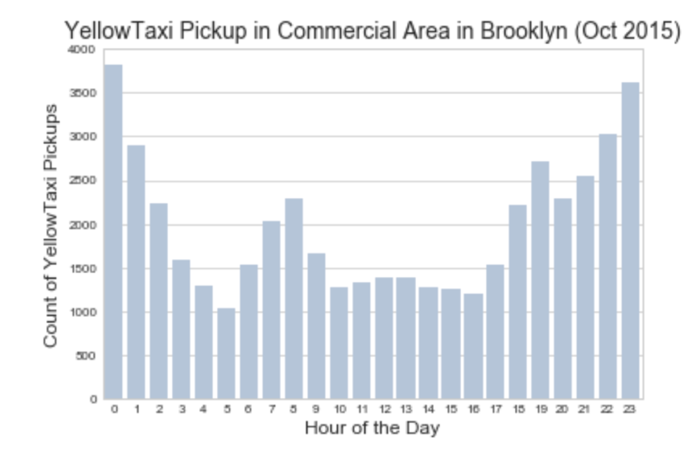
 

* For the pick-up trips, rare percent of trips departed from park areas or Battery Park, while about 16.89% of the pick-up trips departed from manufacturing districts, 22.61% of the pick-up trips departed from commercial districts, and most of the pick-up trips (60.09%) departed from residential districts.
* For the drop-off trips, also rare percent of trips arrived in park areas or Battery Park while about 15.43% of the pick-up trips arrived in manufacturing districts, 11.72% of the pick-up trips arrived in commercial districts, and most of the pick-up trips (72.52%) arrived in residential districts.



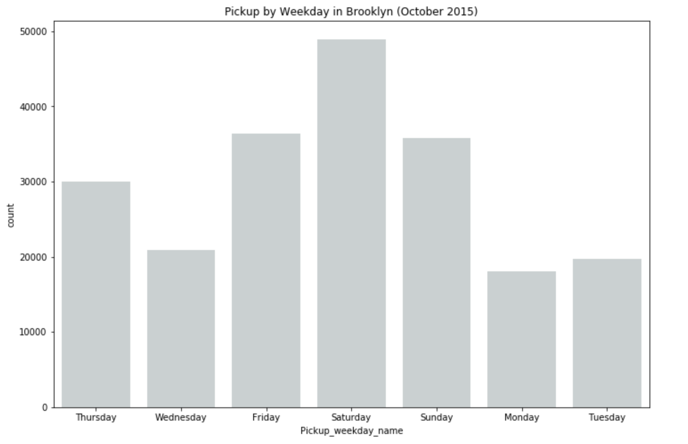
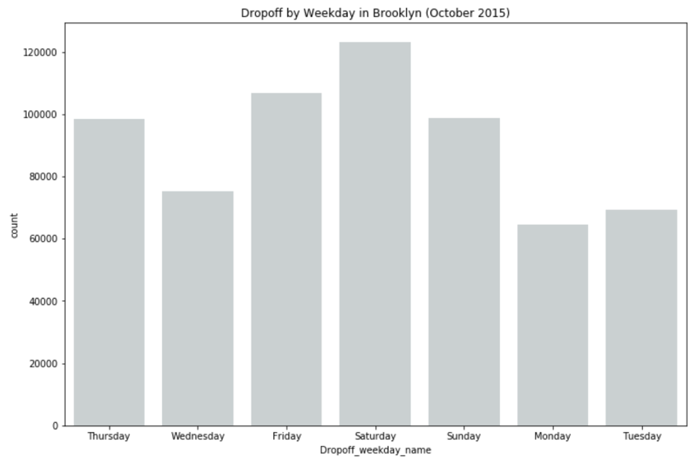


* the activity level of different zoning districts - dividing percentage of pick-up trips by the percentage of zoning district area, and percentage of drop-off trips by the percentage of zoning district area.
* The commercial districts have the highest level of taxi pick-up activity
* The commercial districts also have the highest level of taxi drop-off activity
* More people hail taxis in the commercial districts to somewhere else, and more people take a taxi to the commercial district through a taxi, compared with the taxi activities in residential districts and manufacturing districts.
* Comparing the level of taxi pick-up activity (6.08) and the level of taxi drop-off activity in the commercial districts (3.14), we can conclude that people are more likely to hail taxis when they leave the places in the commercial districts such as shops, offices, theaters, restaurants than they come to the commercial districts.

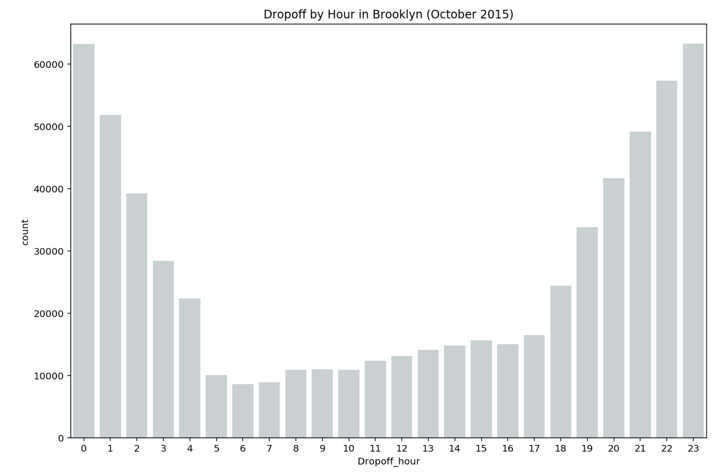
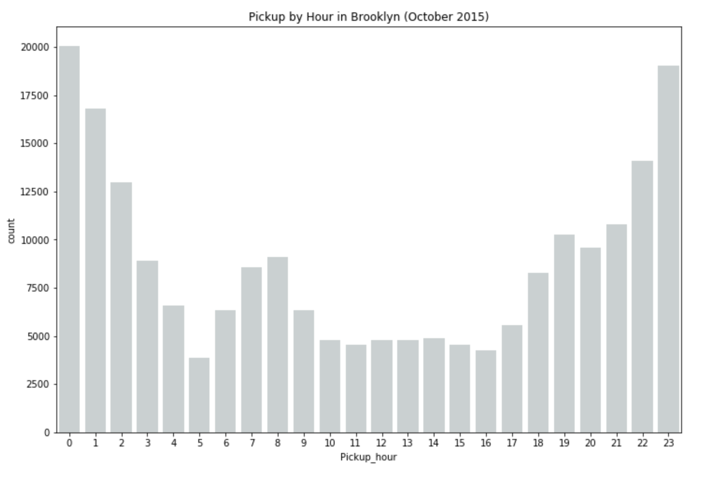


* In the mid of the day, there is a trend of increasing pick-up and drop-off trips in the commercial area, compared with the same time period in the figure showing the pattern of all pick-up and drop-off trips in the residential districts.

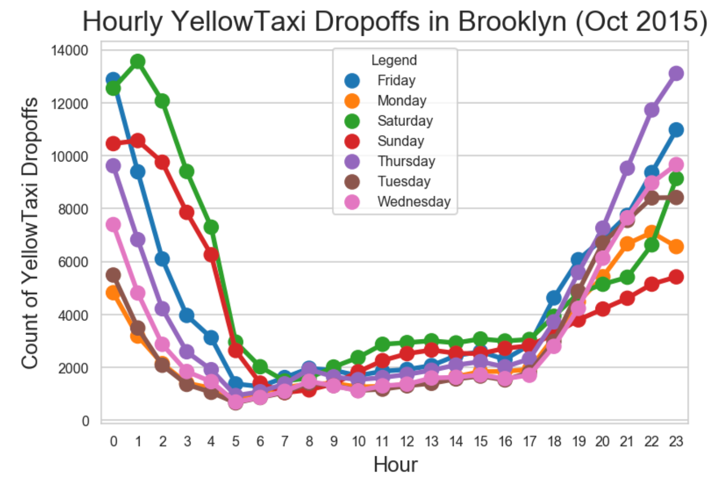
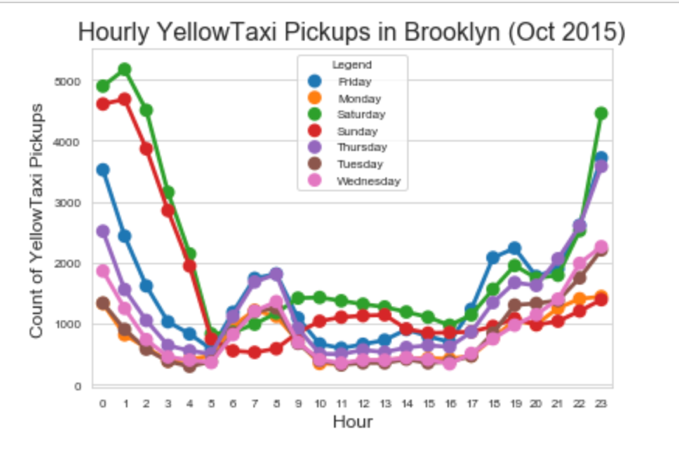
**Overall Descriptive Analysis**



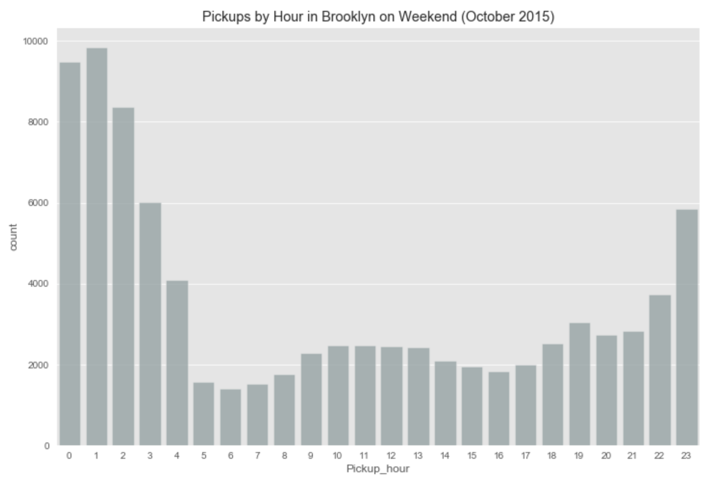
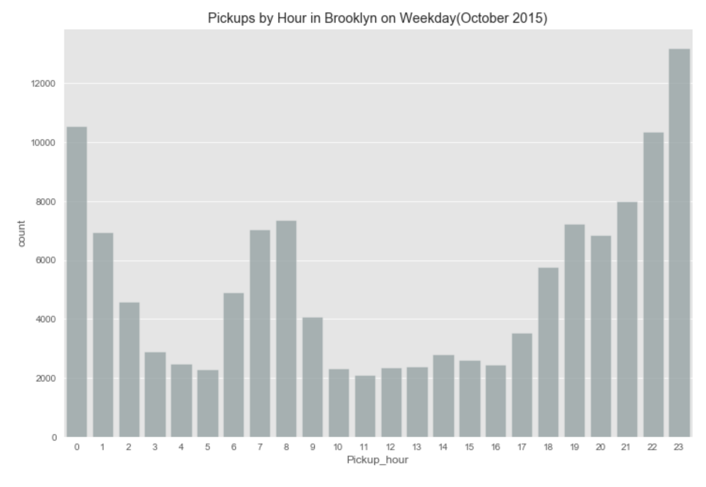
* More taxi trips on Saturday than other days of a week, no matter for drop-off trips or pick-up trips.

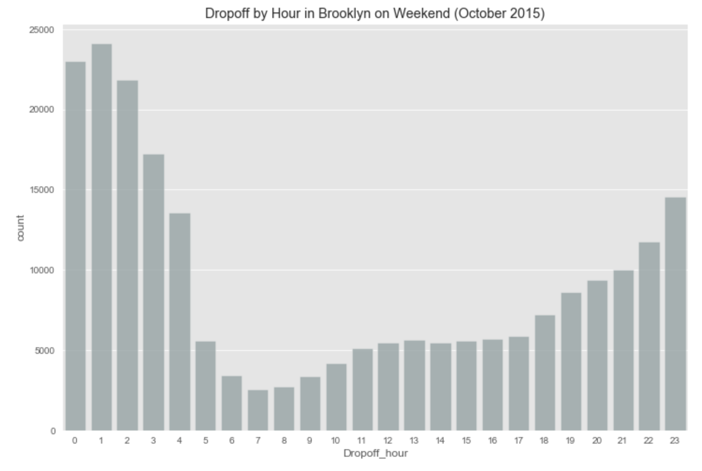
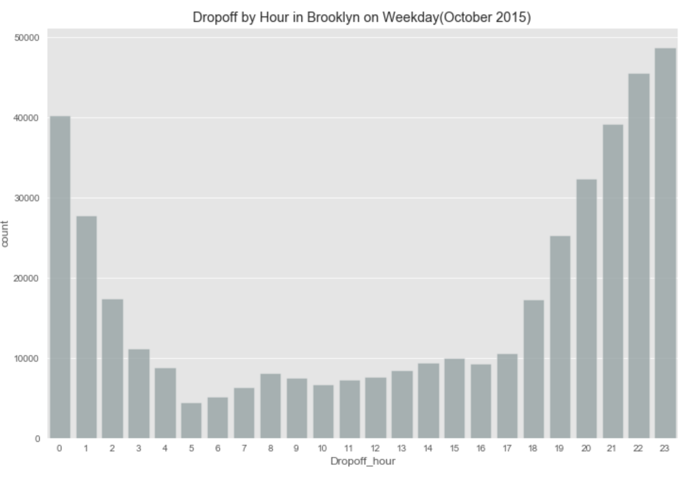


* Much more drop-off trips in Brooklyn from 10:00 pm to 0:00 am and more pick-up trips from 11:00 pm to 1:00 am.
* Much more drop-off trips than pick-up trips in Brooklyn, which matches what we found in previous analysis.



* Taxi travel patterns are different on weekdays and on the weekend.





* On weekdays, the peak hours of both pick-up and drop-off trips in Brooklyn is from 10:00 pm to 0:00 am.
* On weekend, the peak hours of both pick-up and drop-off trips in Brooklyn is from 0:00 am to 2:00 am,
* The volume of drop-off trips is much larger than that of pick-up trips no matter when.

**Typical pick-up and drop-off trips**

* Clustering all pick-up and all drop-off trips into 50 groups on weekdays and on weekend respectively in four maps.
* Pink points represent originations while blue points represent destinations.

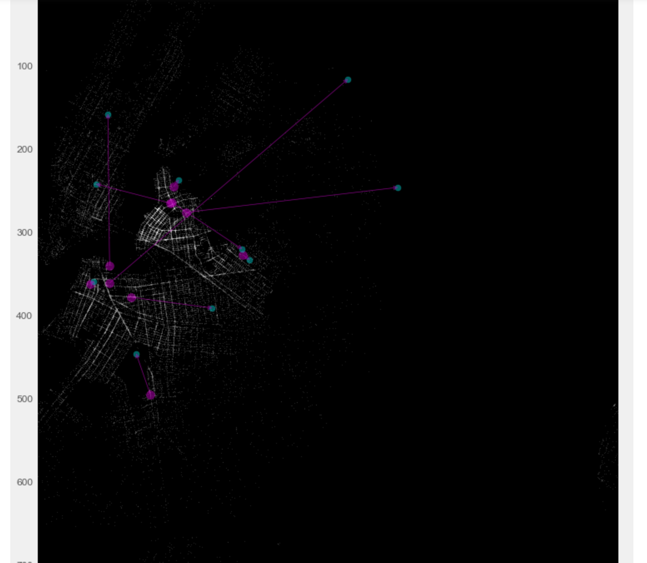


Figure 4 weekend typical pick-up trips Figure 5: weekend typical drop-off trips

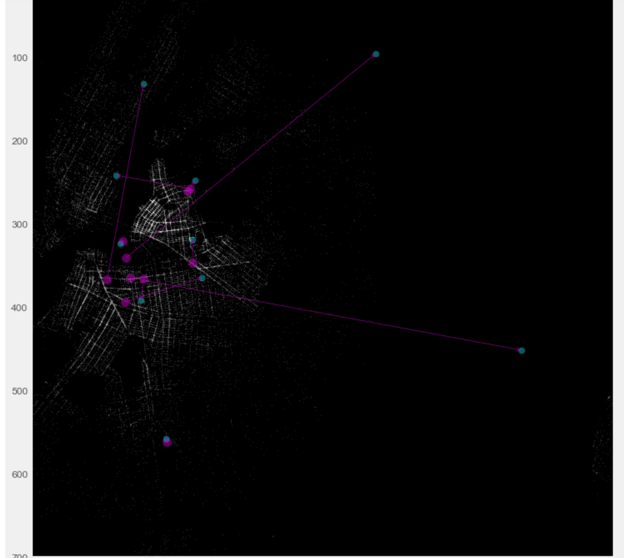


Figure 6 weekday typical pick-up trips Figure 7 weekday typical drop-off trips

* Typical pick-up trip maps on weekdays and on weekend
  + Commons: the originations of most typical pick-up trips, no matter on weekdays and on weekend, are concentrating in west Brooklyn around Williamsburg and Brooklyn Heights.
  + Difference: more people take taxis from Brooklyn to the airport on weekdays.
* Typical drop-off trip maps on weekdays and on weekend
  + Common: the destinations of most typical drop-off trips, no matter on weekdays and on weekend are concentrating in northwest Brooklyn around Bedford-stuyvesant, Red Hook, Williamsburg and Brooklyn Heights.
  + Difference: more people take taxis from Manhattan to Brooklyn on weekdays.

**Popular trips**

* Grouping pick-up and drop-off trips respectively according to their pick-up (origination) longitude, pick-up latitude, drop-off (destination) longitude, and drop-off latitude.

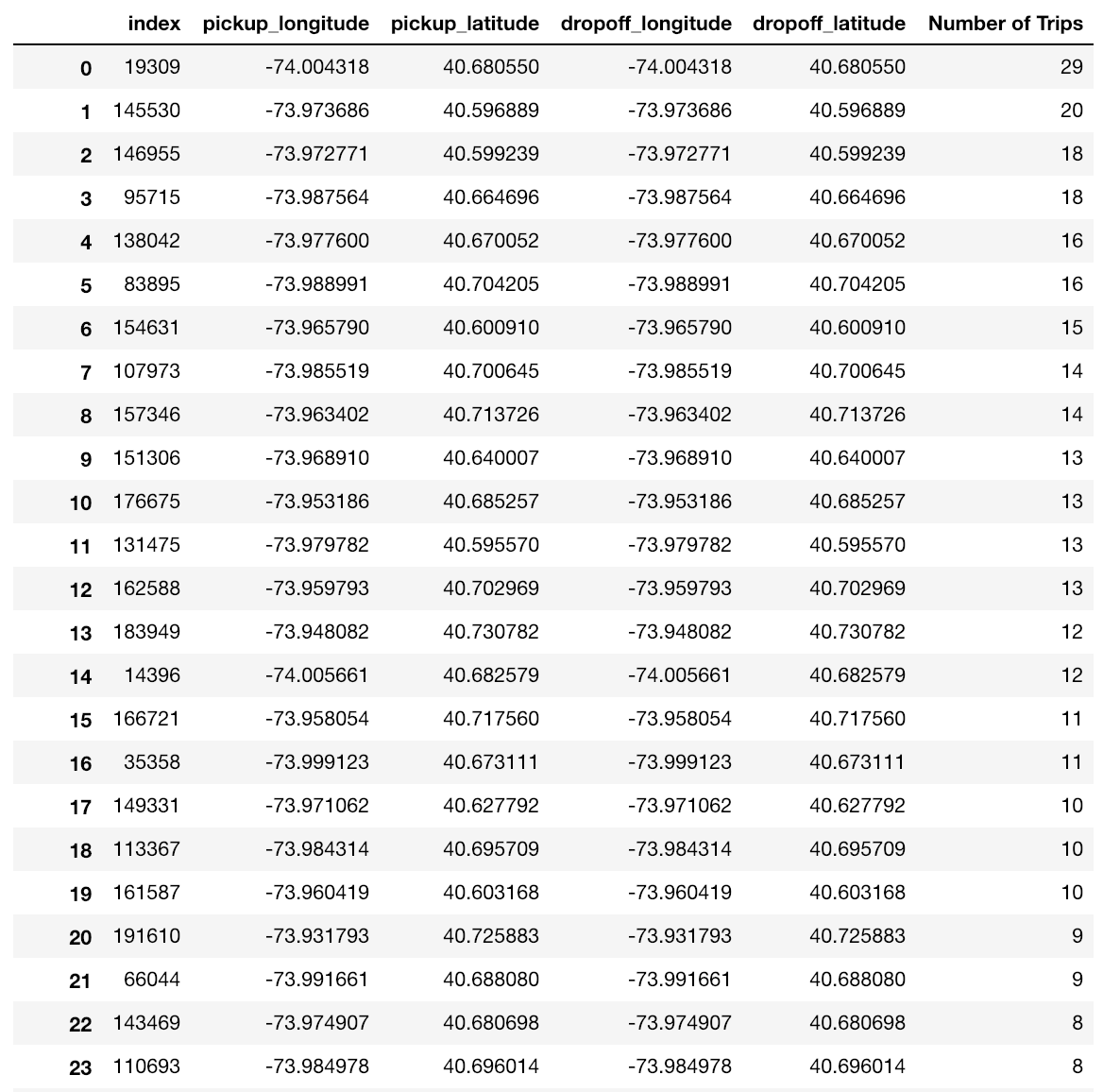
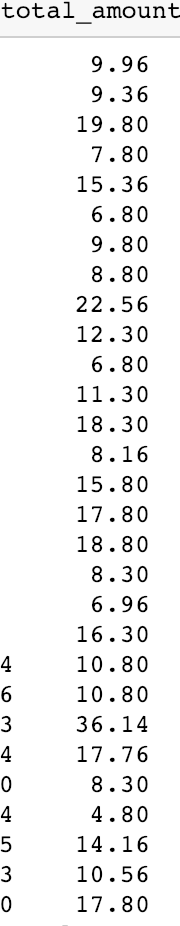


Figure 8 The most popular pick-up/drop-off trips in Brooklyn.

* The most popular pick-up and drop-off trips are those without trip distance.



* The trip cost of those popular trips without any trip distance are not zero
* This might because of the disordered taxi meters or the frauds.

**Conclusion**

* More people hail taxis in the commercial districts to somewhere else, and more people take a taxi to the commercial district through a taxi, compared with the taxi activities in residential districts and manufacturing districts.
* People are also more likely to hail taxis when they leave the places in the commercial districts.
* Characteristics of the yellow taxi’s data
  + Much more drop-off trips than pick-up trips
  + More taxi trips on Saturday than any other day of a week, no matter for drop-off trips or pick-up trips.
* The typical trips visualized and the populous trips identified provide a more narrative story of yellow taxi activity in Brooklyn.

**Data Refences**

New York City Department of City Planning. 2018. NYC GIS Zoning Features Release [shapefile]. Retrieved from <https://www1.nyc.gov/site/planning/data-maps/open-data.page>.

Calgary, O. (n.d.). 2017 Yellow Taxi Trip Data | NYC Open Data. Retrieved December 10, 2018, from [/Transportation/2017-Yellow-Taxi-Trip-Data/biws-g3hs, /Transportation/2017-Yellow-Taxi-Trip-Data/biws-g3hs](file:////Transportation/2017-Yellow-Taxi-Trip-Data/biws-g3hs,%20/Transportation/2017-Yellow-Taxi-Trip-Data/biws-g3hs)

NYC Taxi & Limousine Commission - Trip Record Data. (n.d.). Retrieved December 10, 2018, from <http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml>