**Data Mining Taxi Trips in Brooklyn**

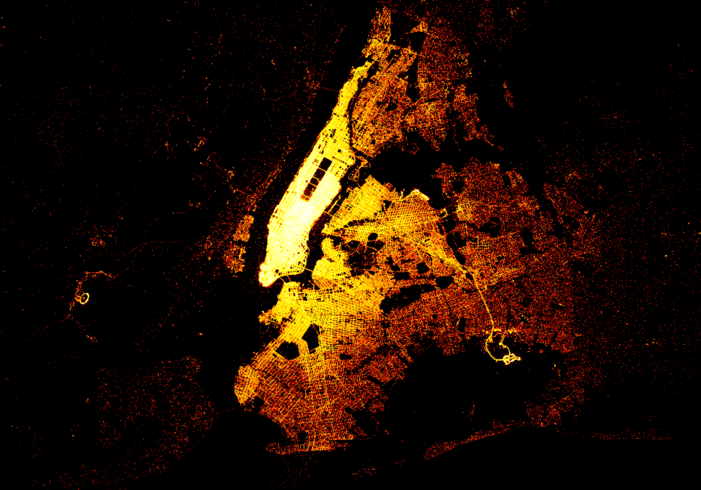
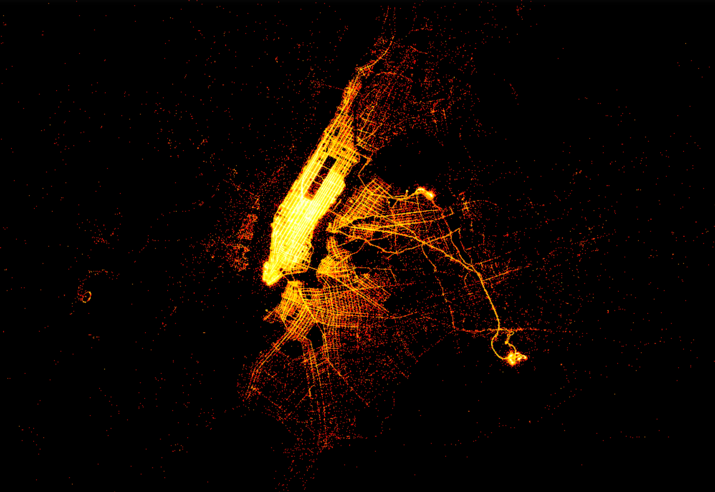
**Project Description:**

* 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.

**Data：**

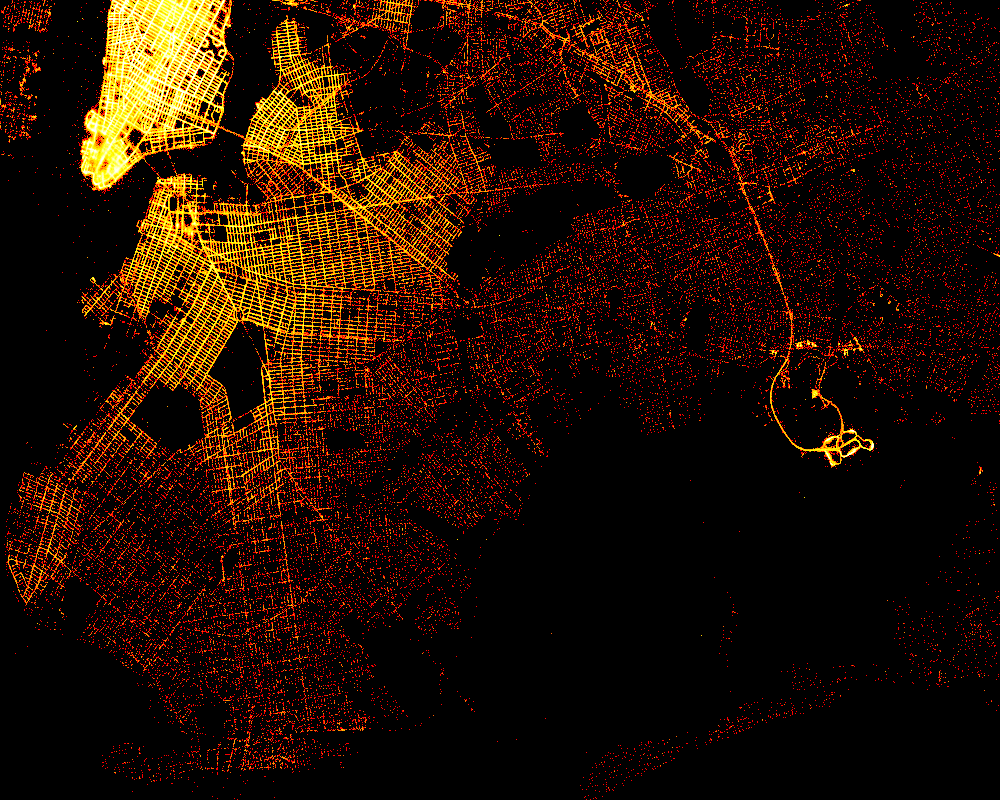
* Yellow cab’s trip data of the New York City in October 2015
* NYC GIS Zoning Features
* NYC Commercial/Residential FAR Dataset
* Taxi Zones Shapefiles

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



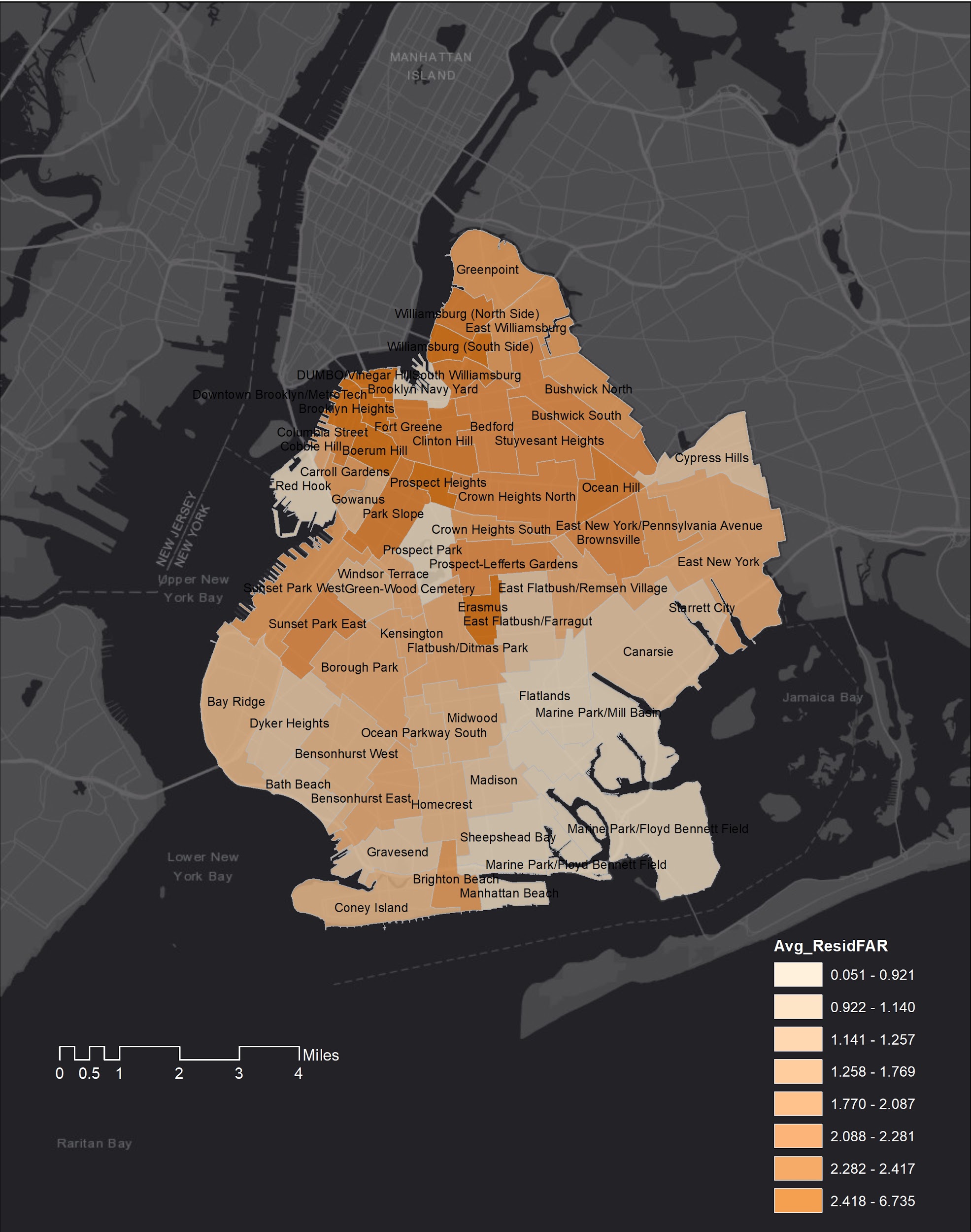
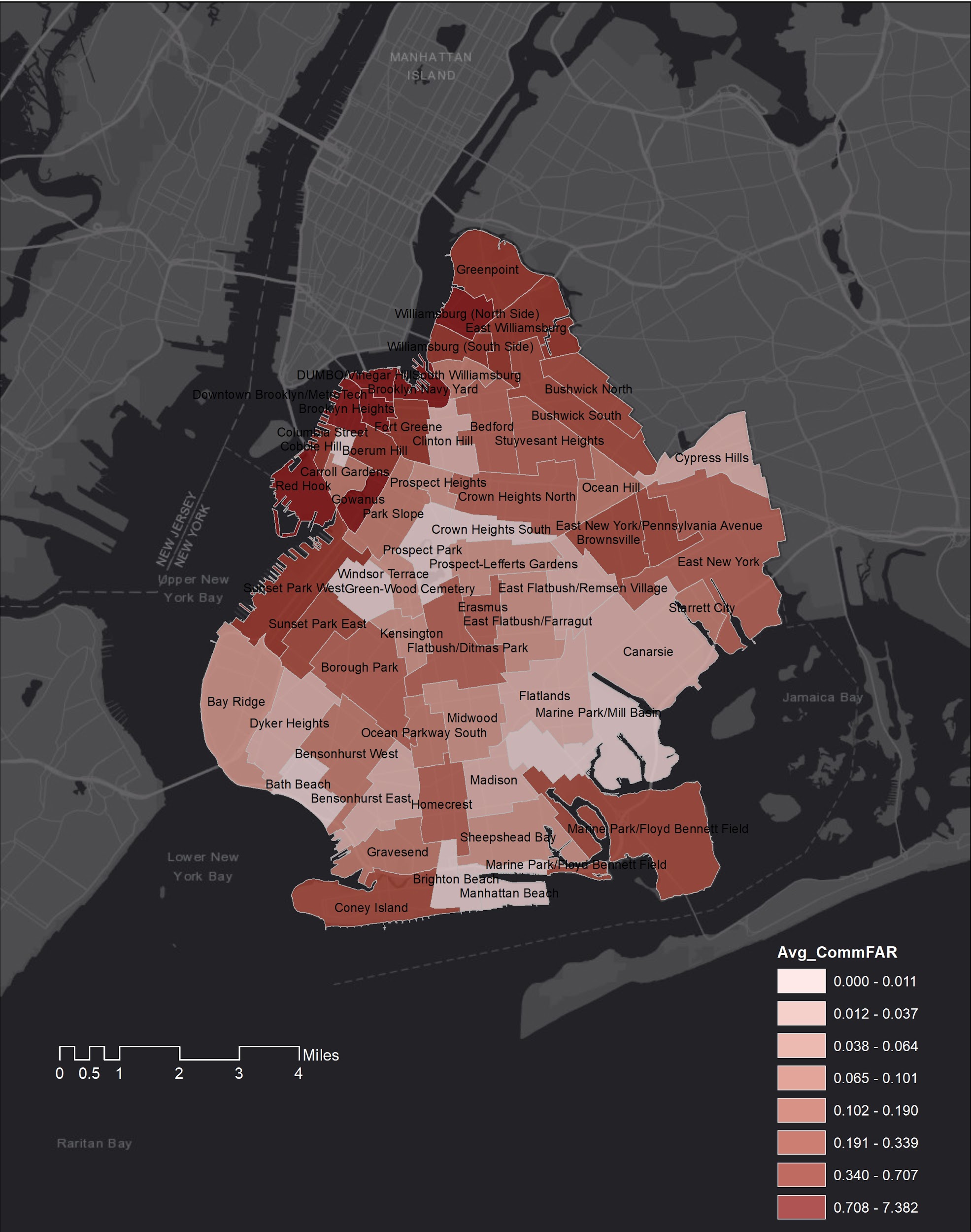
Manhattan, as an extremely compact, highly mixed-use and traffic-concentrated area, is too special for our analysis

**Analysis site zoomed to Brooklyn, NY**



* Pick-up trip map and Drop-off trip map have very different patterns.
* BK. has more drop-off trips than pick-up trips - people are more likely to take taxis to BK, while few from there: the yellow taxis are mainly concentrated in the MN.
* Most pick-up trips in BK in October happen within 5 miles around MN and areas close to JFK.
* Drop-off trips scatter around the entire Brooklyn.

**Commercial FAR and residential FAR**



* Commercial FAR & residential FAR show the same pattern in Brooklyn: taxi zones with higher commercial FAR are also tend to have higher residential FAR.
* Northern BK has higher commercial and residential FAR, it is also a relatively denser area in BK.
* Hypothesis: Denser places with higher population would generate more taxi trips.

**Spearman's rank correlation analysis**

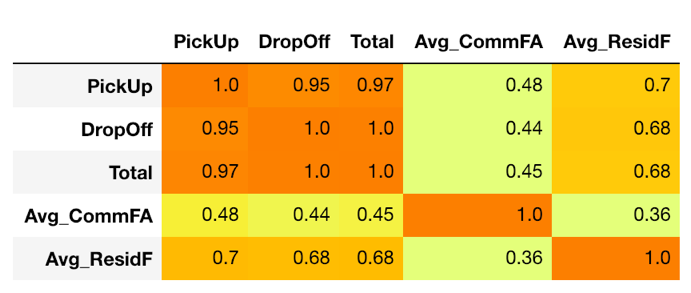


Figure 1 Correlation\_Analytics\_BK\_all

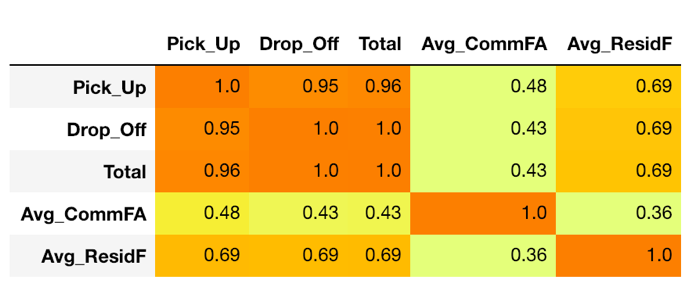


Figure 2 Correlation\_Analytics\_BK\_weekend

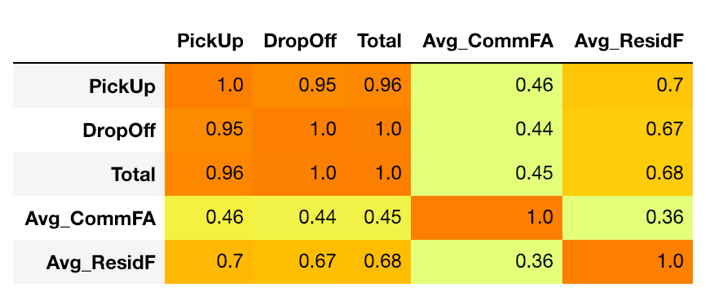
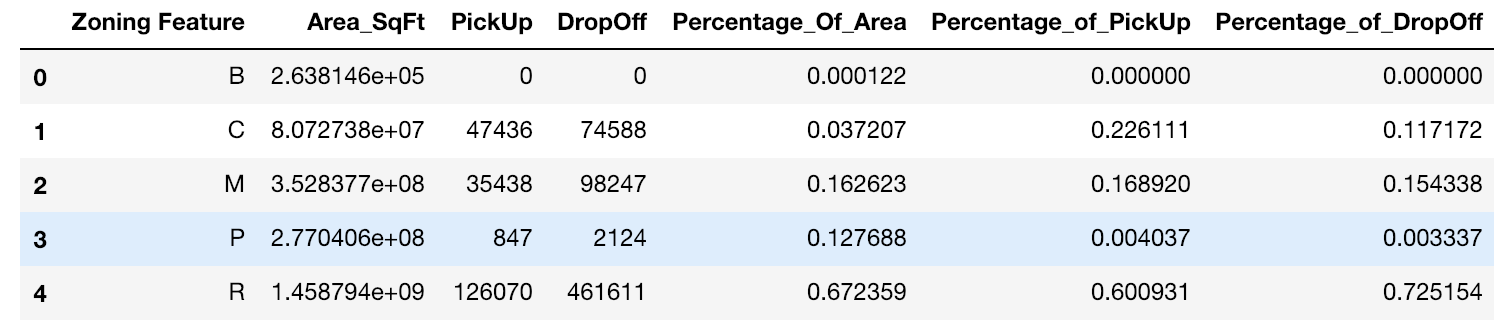


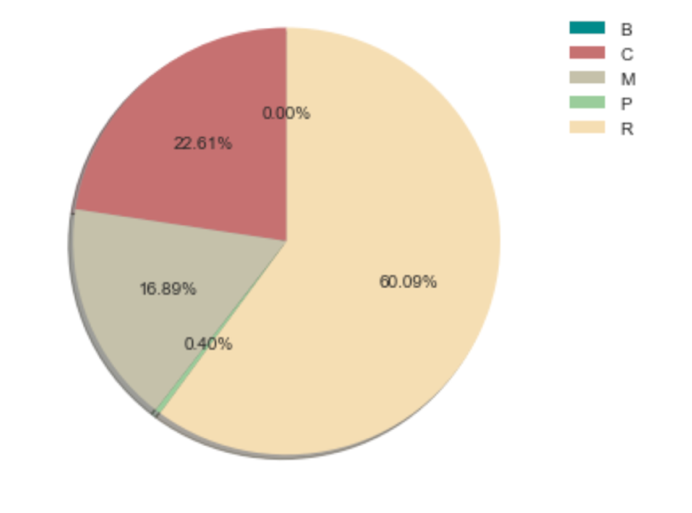
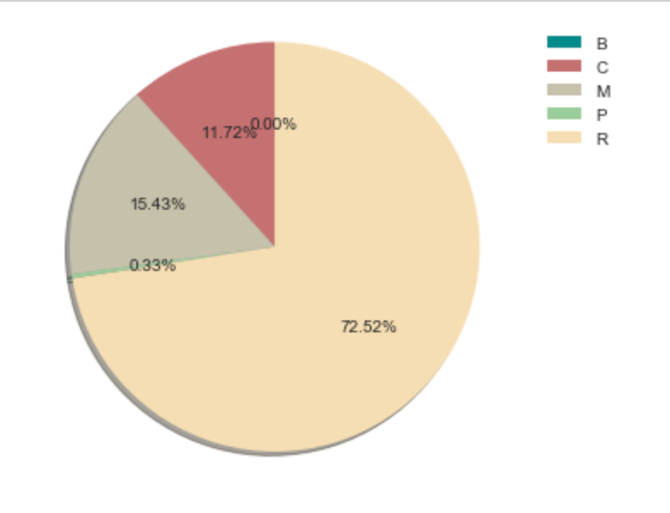
Figure 3 Correlation\_Analytics\_BK\_weekday

* There are positive relationships between taxi trips and the Average Commercial FAR/Average Residential FAR calculated based on taxi zone, with slightly difference across different time. Hence, High average residential FAR of a taxi zone is more related to high volume of taxi trips in Brooklyn than high average commercial FAR.

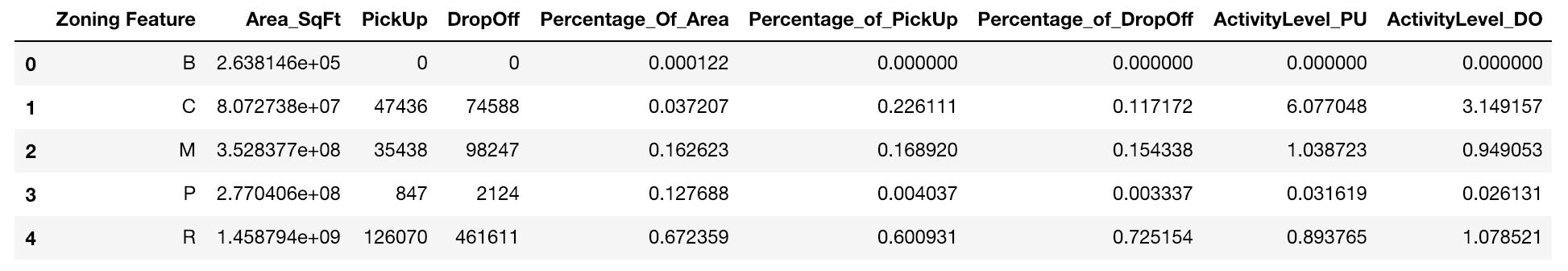
**Land Uses**

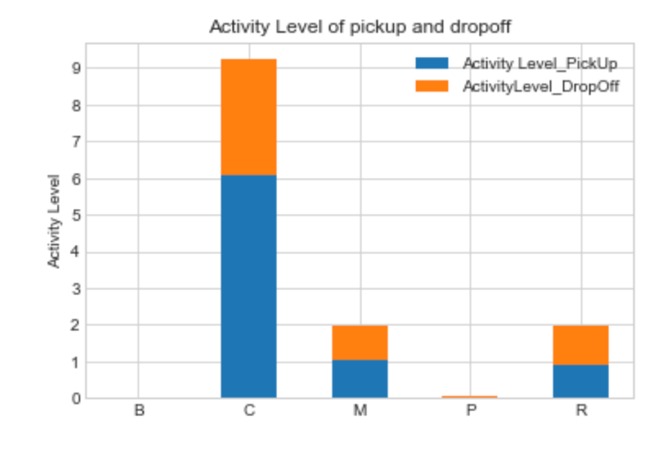


* BK land distribution: 16.26% manufacturing districts, 12.77% park area, 3.72% commercial districts, 67.23% residential districts.

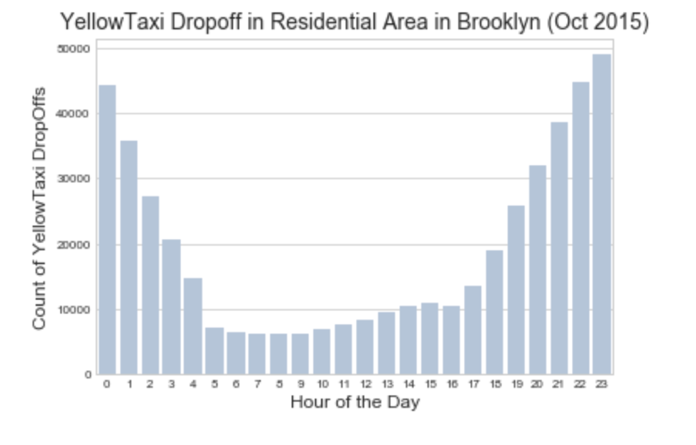
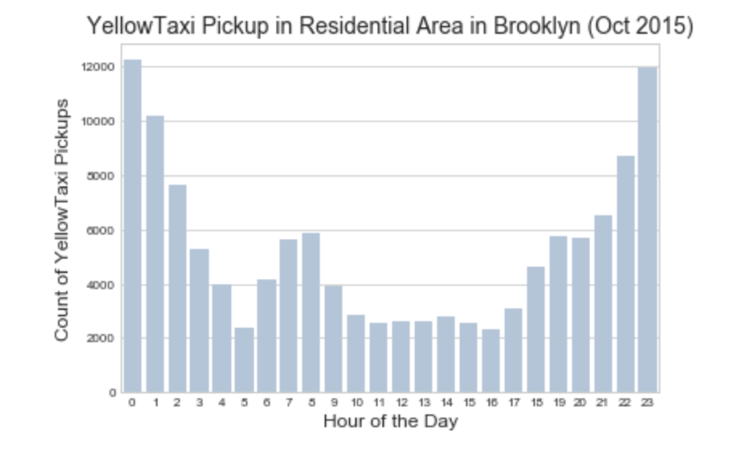
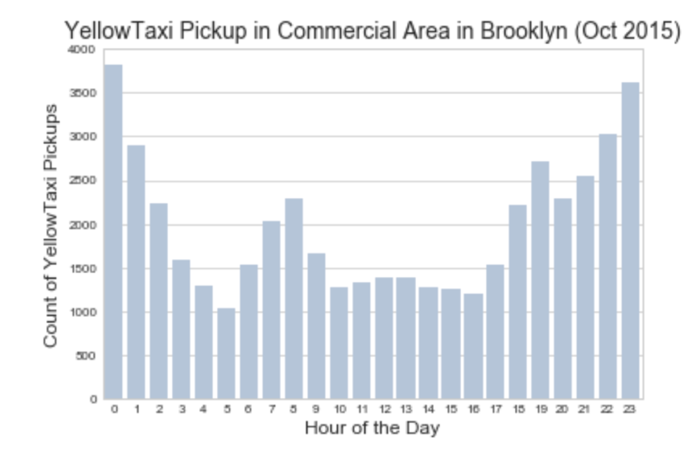
 

* Pick-up trips 16.89% from manufacturing districts, 22.61% from commercial districts, and most of the pick-up trips (60.09%) from residential districts.
* Drop-off trips: 15.43% arrived in manufacturing districts, 11.72% arrived in commercial districts, and most of the pick-up trips (72.52%) arrived in residential districts.



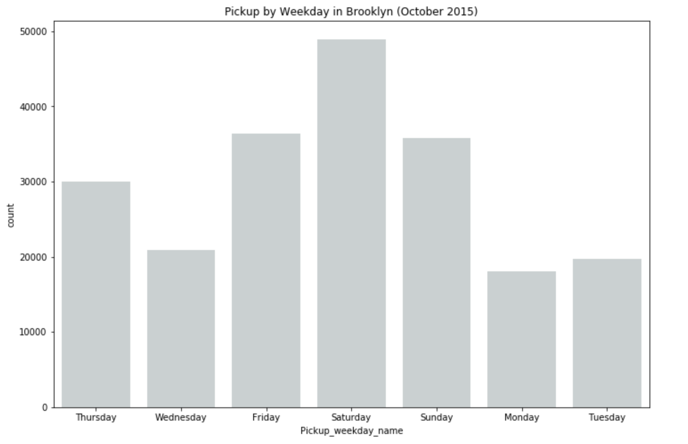
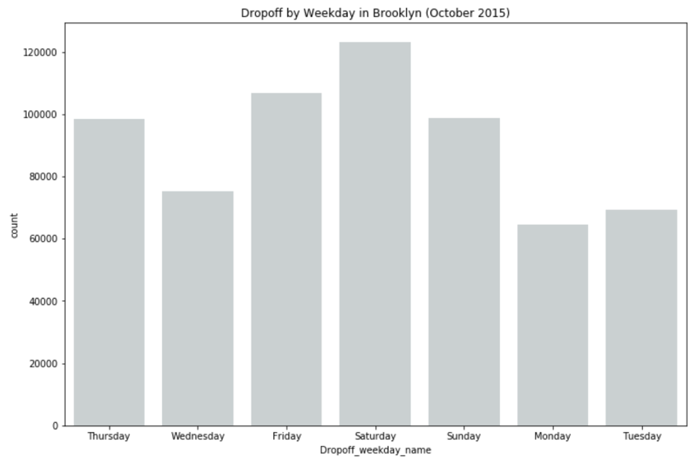


* Activity level of different zoning districts =
* Commercial districts have the highest level of taxi pick-up & drop-off activity.
* More people hail more taxis when they come from & go to commercial districts than they go or come from other districts.
* People are more likely to hail taxis when leaving from the commercial districts such as shops, offices, theaters, restaurants than when they come.

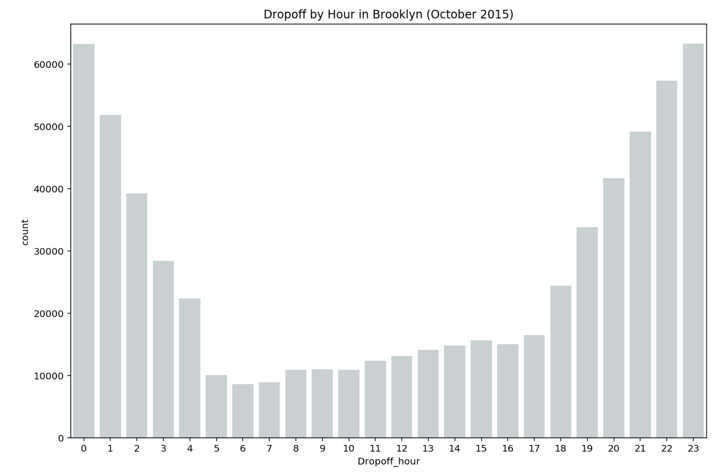
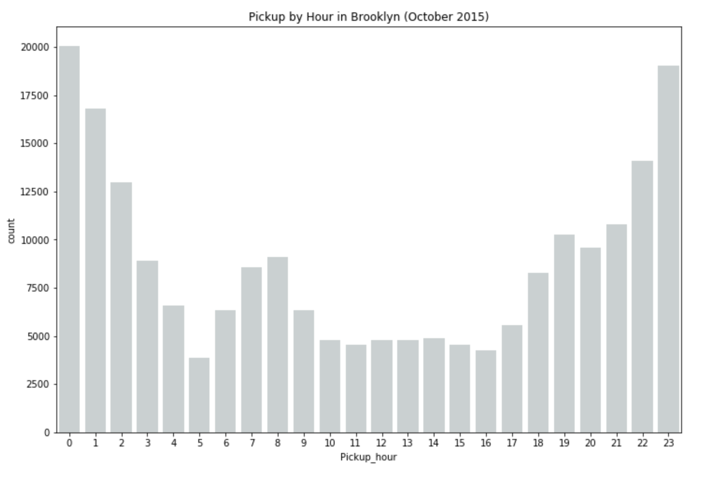


* In the mid of the day, there is a trend of increasing pick-up and drop-off trips in the commercial area, compared with the same time period in the residential districts.

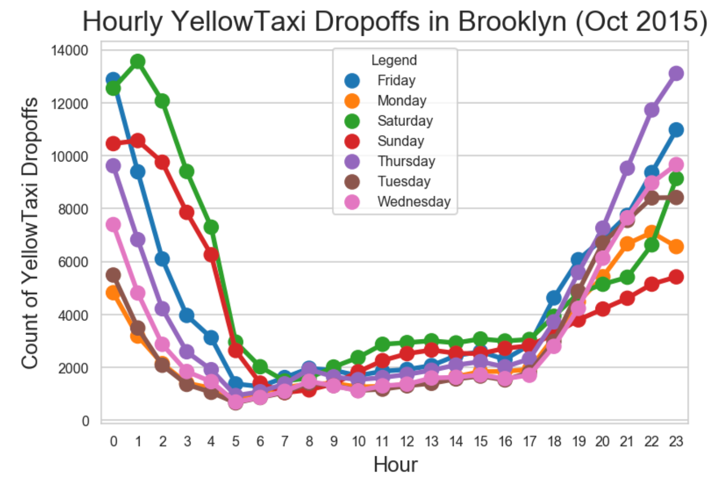
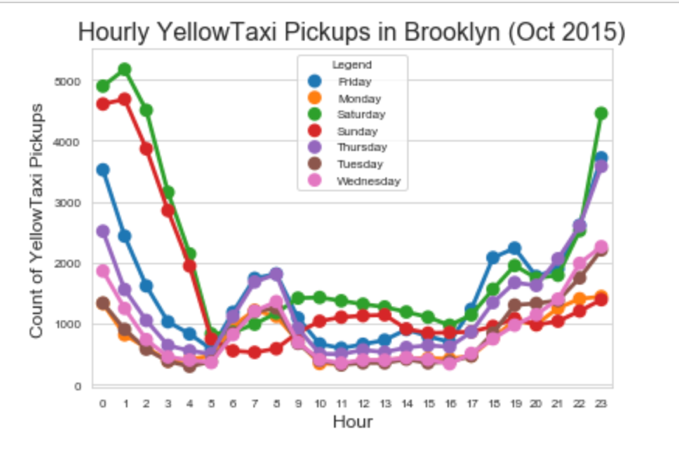
**Overall Descriptive Analysis**



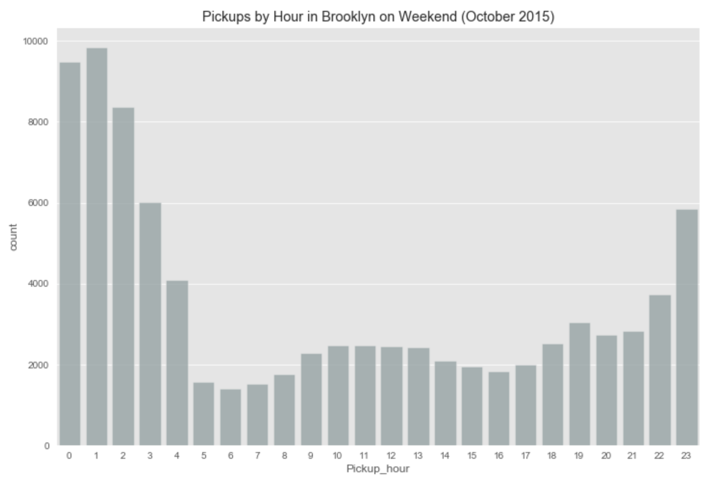
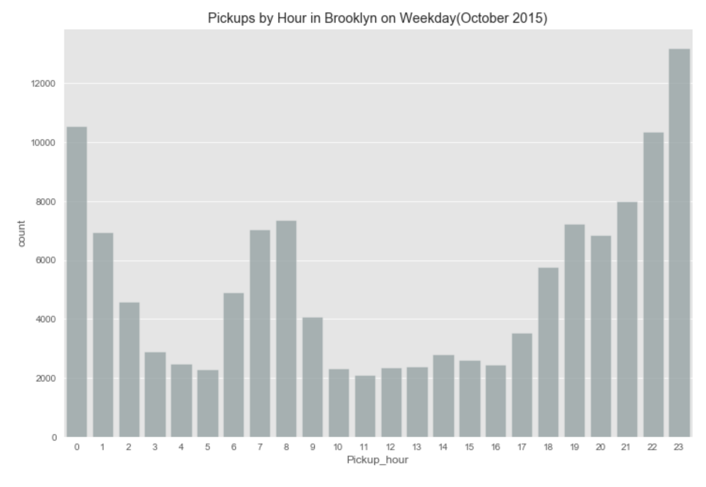
* Highest volume of taxi trips (pickup&dropoff) on Saturday.

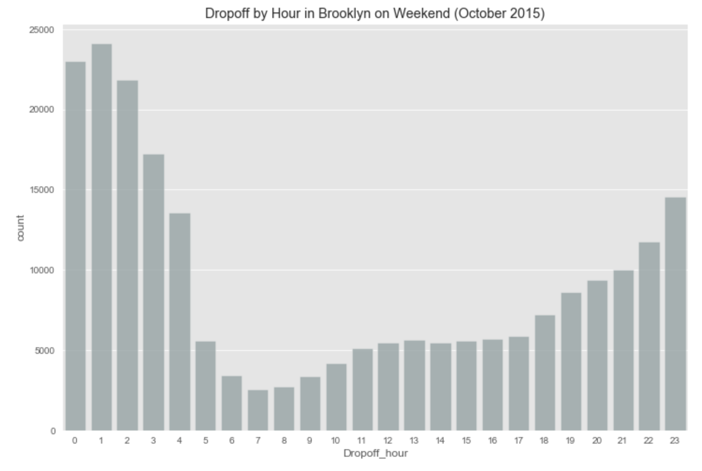
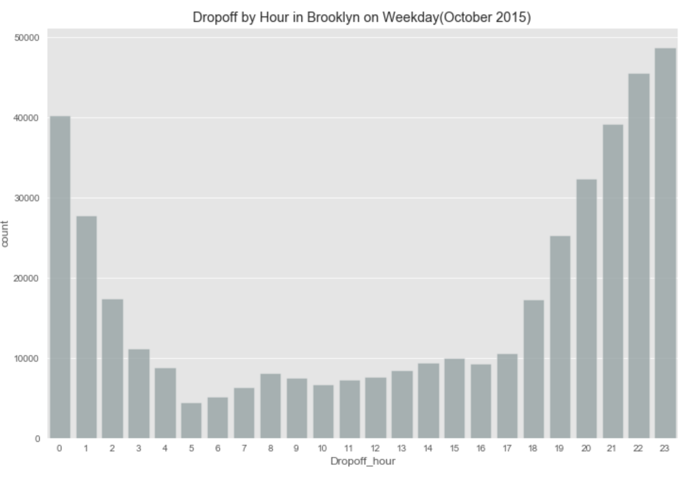


* Much more drop-off trips to BK from 10pm to 12am.
* More pick-up trips from BK from 11pm to 1am.
* Much more drop-off trips than pick-up trips in BK.



* Different Taxi travel patterns on weekdays and weekend.





* Weekday peak hours (pickup & dropoff) in BK: 10 pm to 12 am.
* Weekend peak hours (pickup & dropoff) in BK: 12 am to 2 am.
* Much more drop-off than pick-up one all-day long.

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

* Clustering = 50; Pink points represent originations while blue points represent destinations.

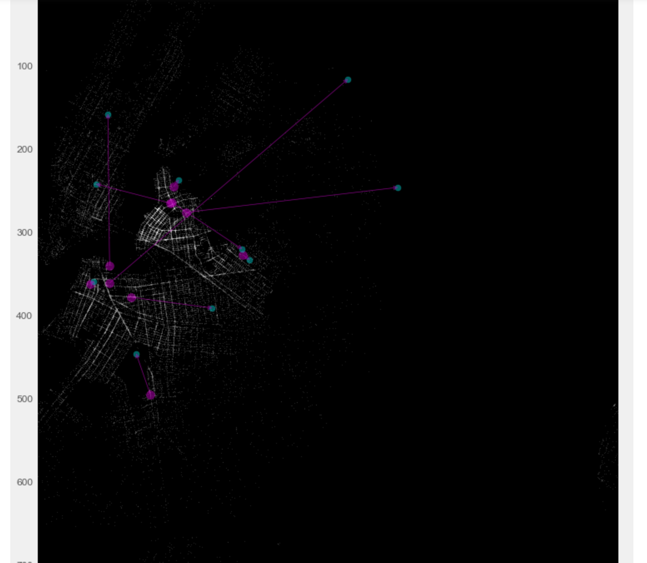


Figure 4 weekend typical pick-up trips Figure 5: weekend typical drop-off trips

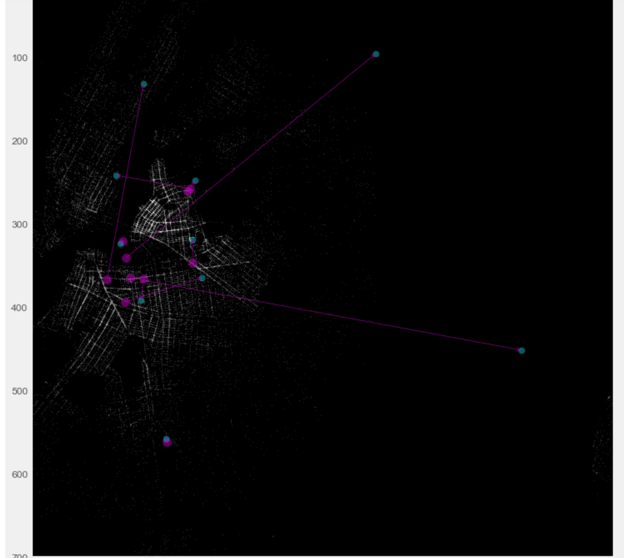


Figure 6 weekday typical pick-up trips Figure 7 weekday typical drop-off trips

* Typical pick-up trip maps on weekdays and on weekend
  + Commons: the originations of most typical pick-up trips are concentrating in west BK around Williamsburg and Brooklyn Heights.
  + Difference: more people take taxis from BK to the airport on weekdays.
* Typical drop-off trip maps on weekdays and on weekend
  + Common: the destinations of most typical drop-off trips are concentrating in northwest BK around Bedford-stuyvesant, Red Hook, Williamsburg and Brooklyn Heights.
  + Difference: more people take taxis from MN to BK on weekdays.

**Popular trips**

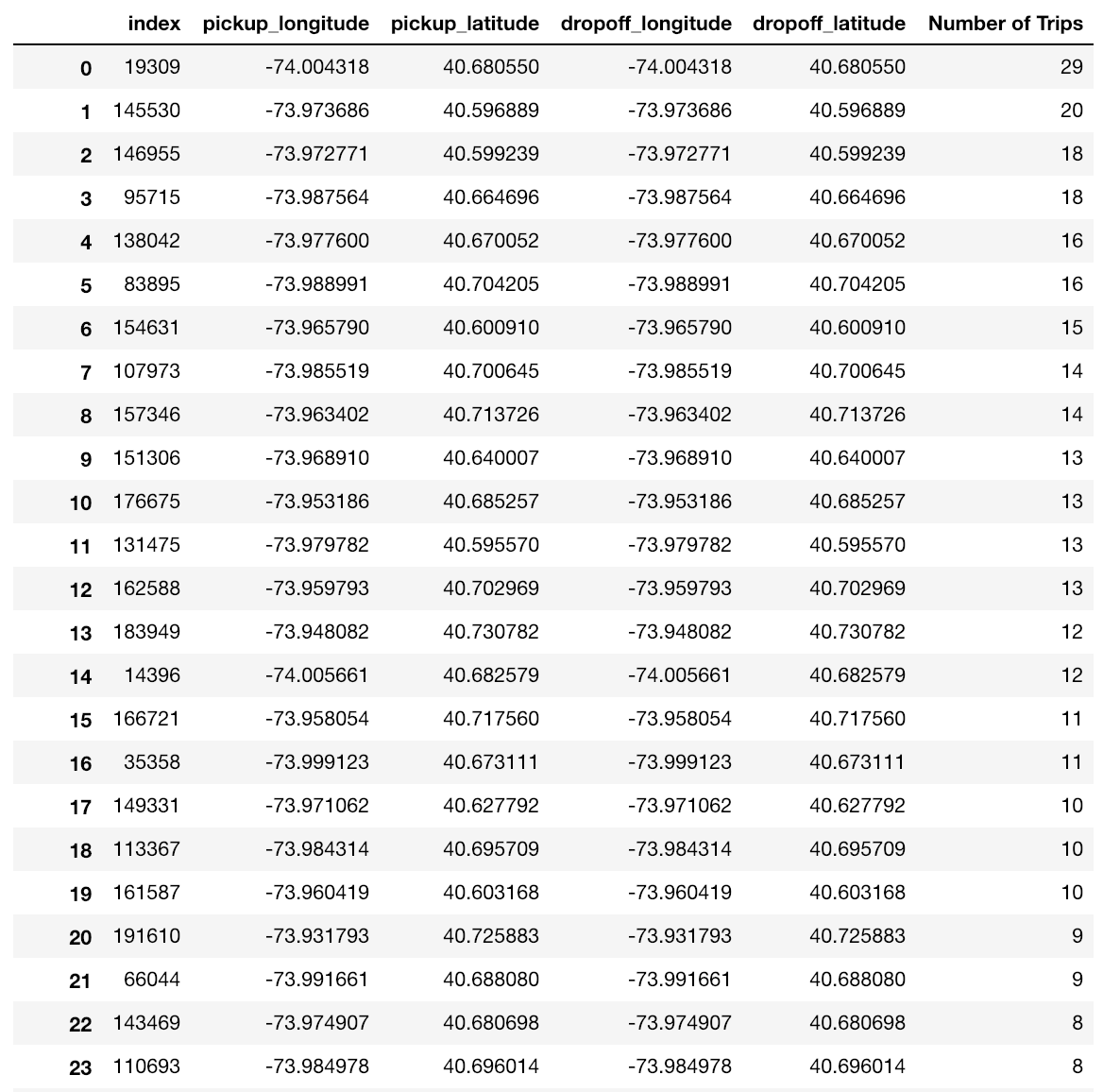
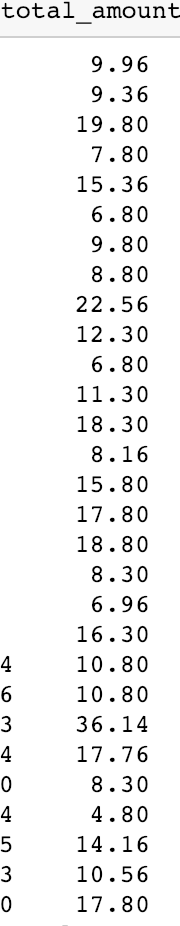


Figure 8 The most popular pick-up/drop-off trips in Brooklyn.

* By Grouping & Soring Trips Coordinates, the most popular pick-up and drop-off trips are those pickup coordinates = drop-off coordinates. However, the trip cost and trip distance of those identified popular trips are not zero, which may be caused 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>