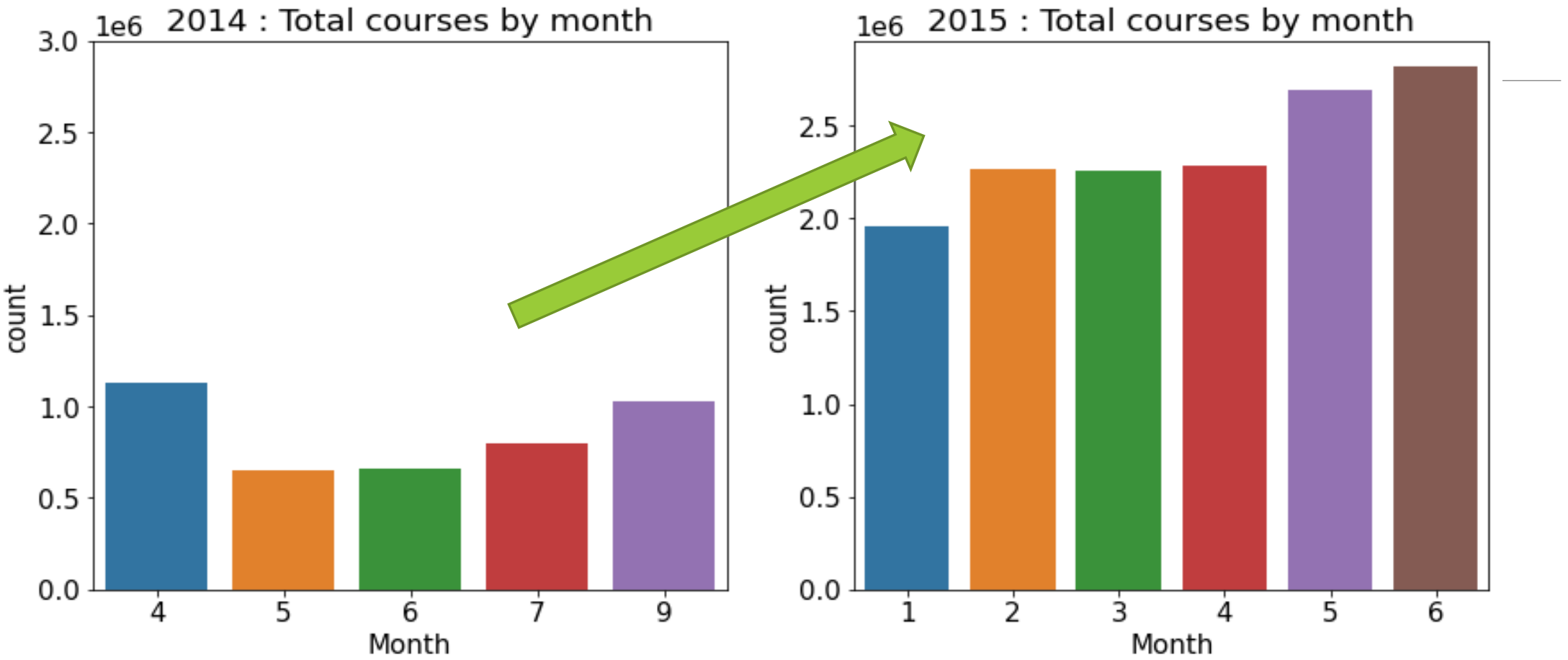
Uber pick-up

TRANG NGHIEM

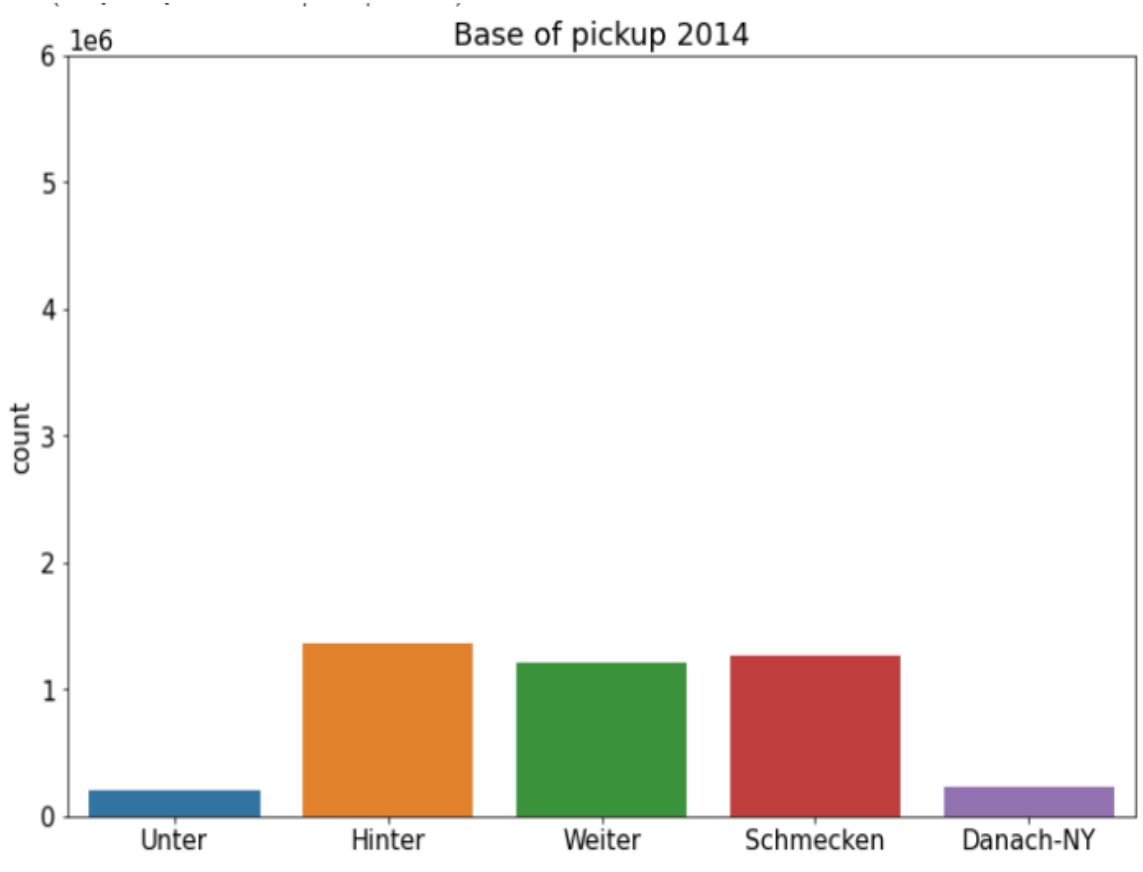
WHERE & WHEN
to find clients quickly?

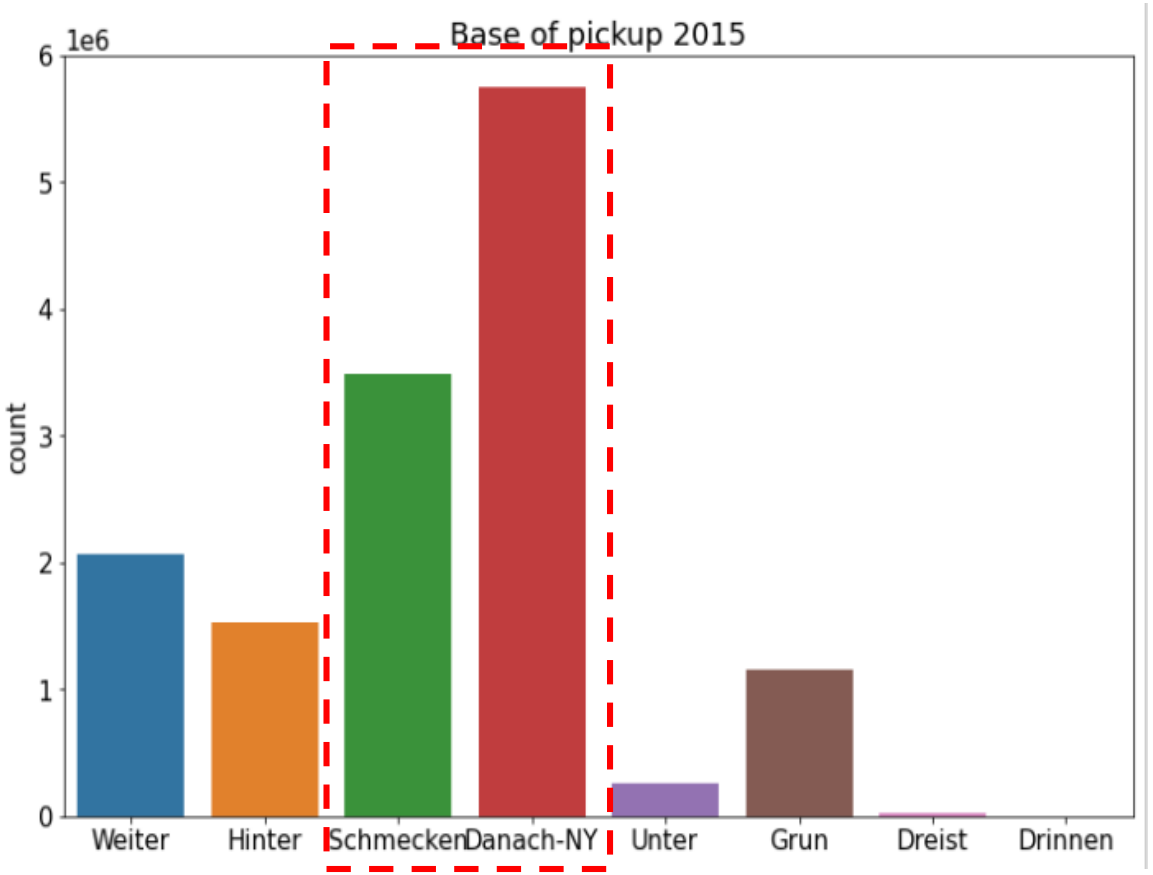
Rides in 2014 vs. 2015



Total rides increase progressively from 2014 to middle 2015

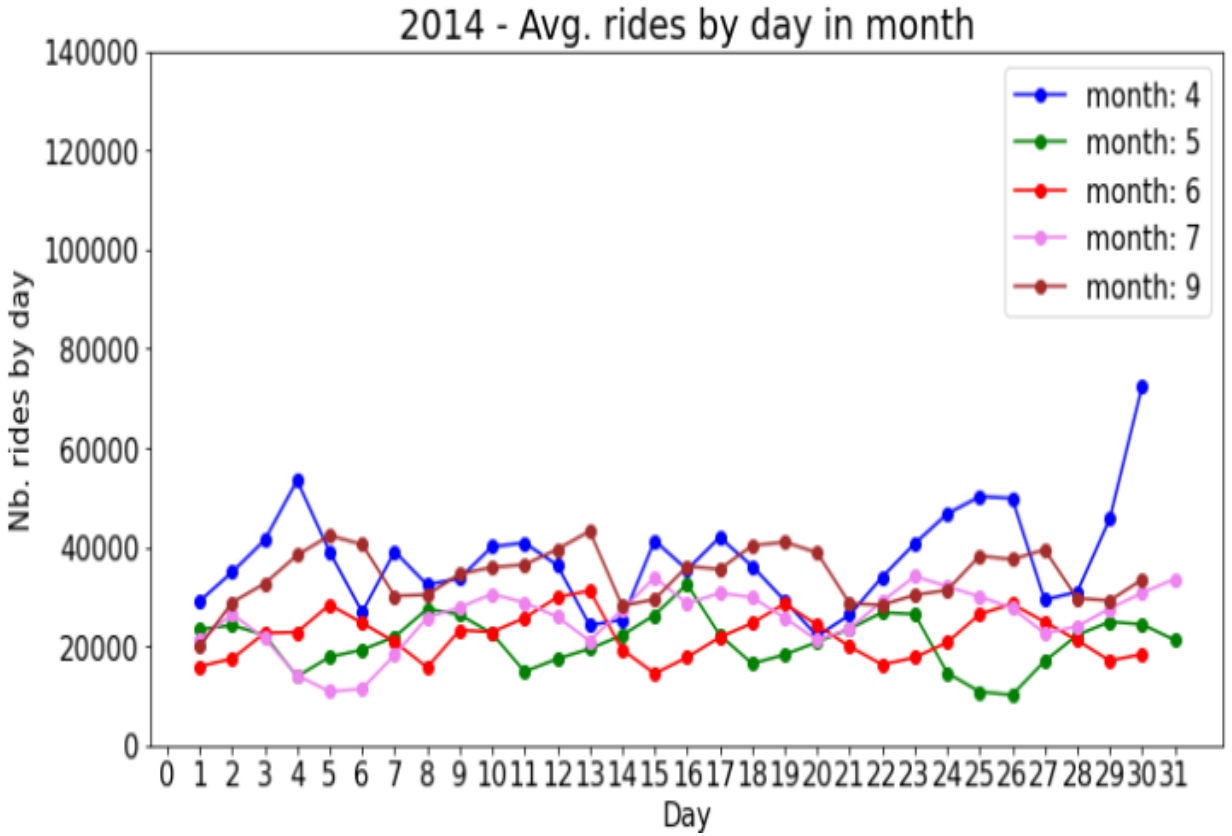
Pick-up Bases 2014 vs. 2015

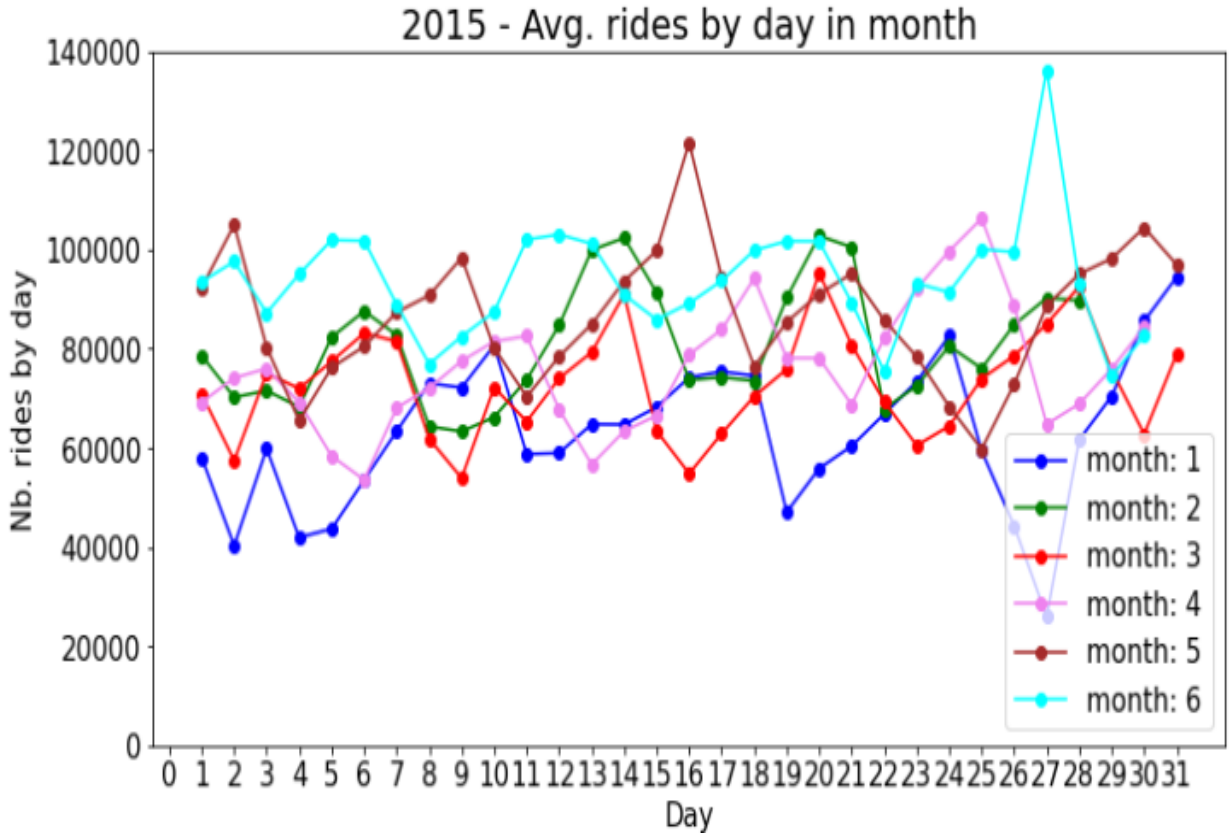




- Extend from 4 bases in 2014 to 8 bases in 2015
- 2015 : 2 important bases are Schmecker & Danach NY

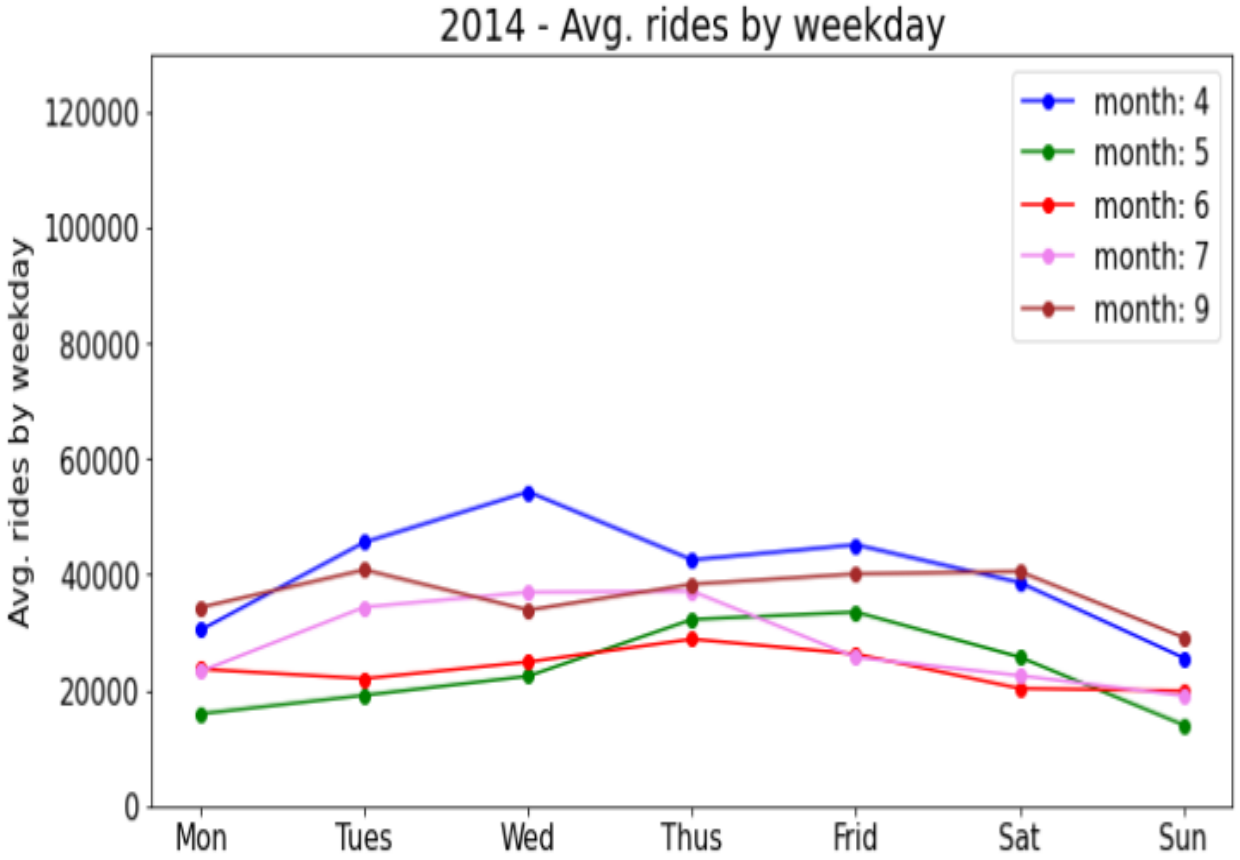
Average rides by days in month

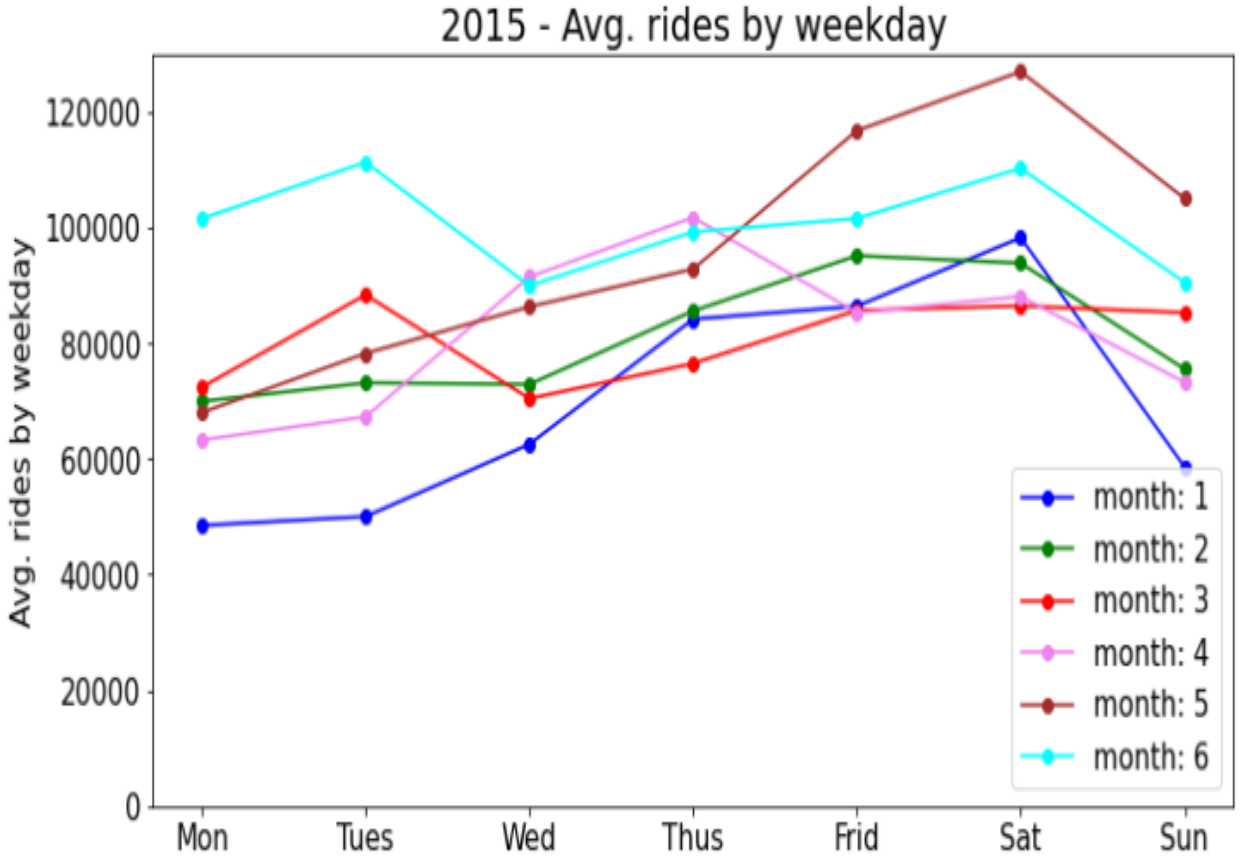




No explicit influence of days in month on the number of rides

Avg. rides by weekday

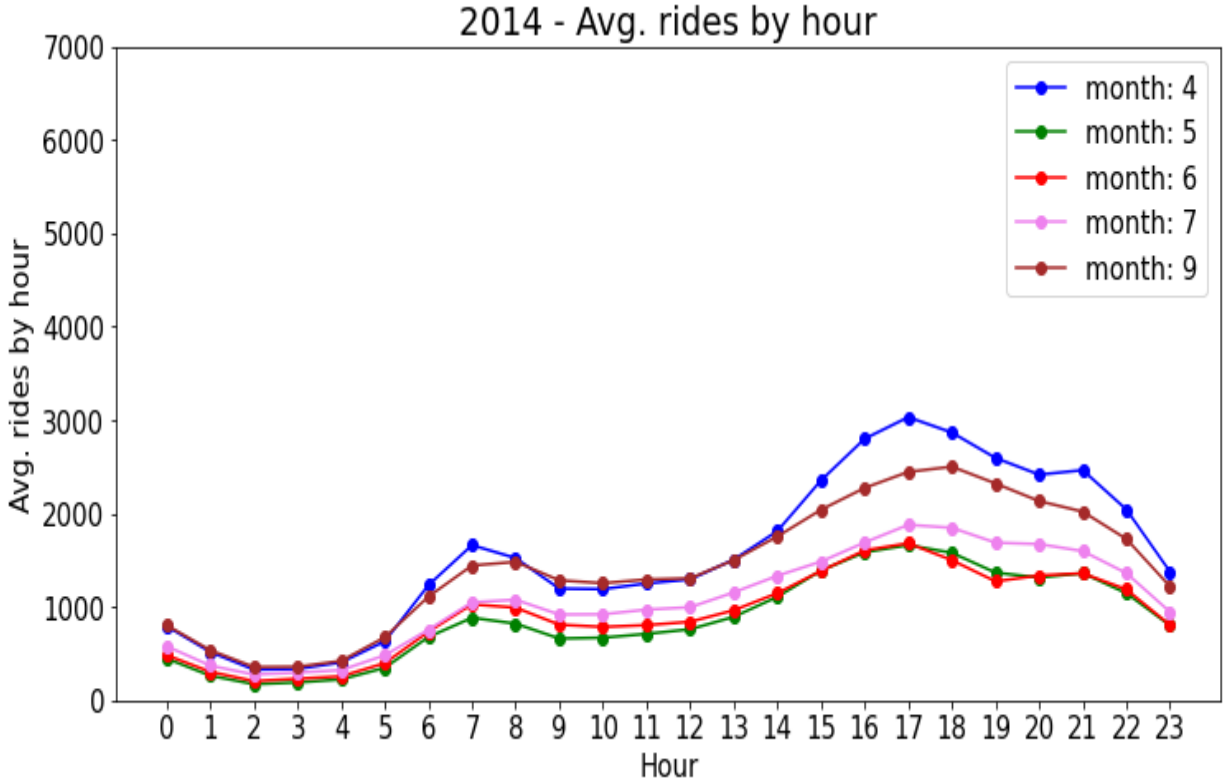


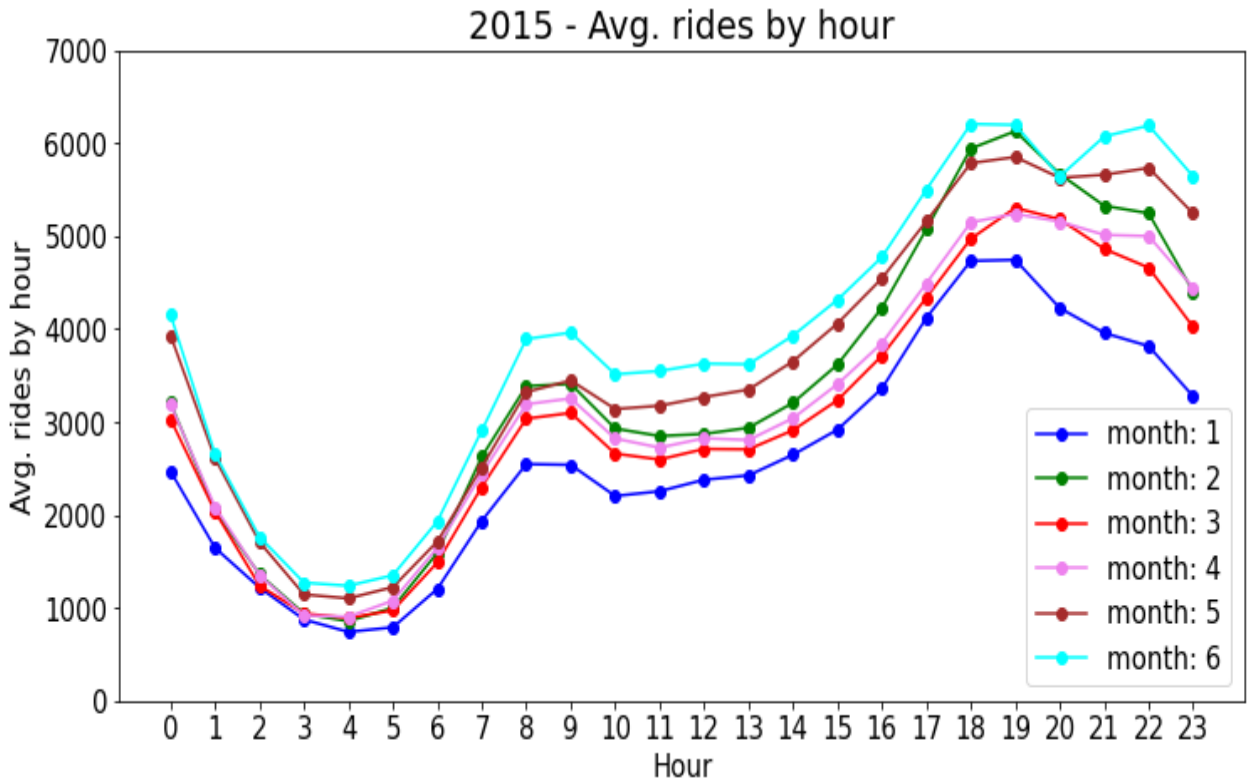


More rides on : Sat., Frid.

Less rides on : Sun., Mon.

Avg. rides by hour in weekday



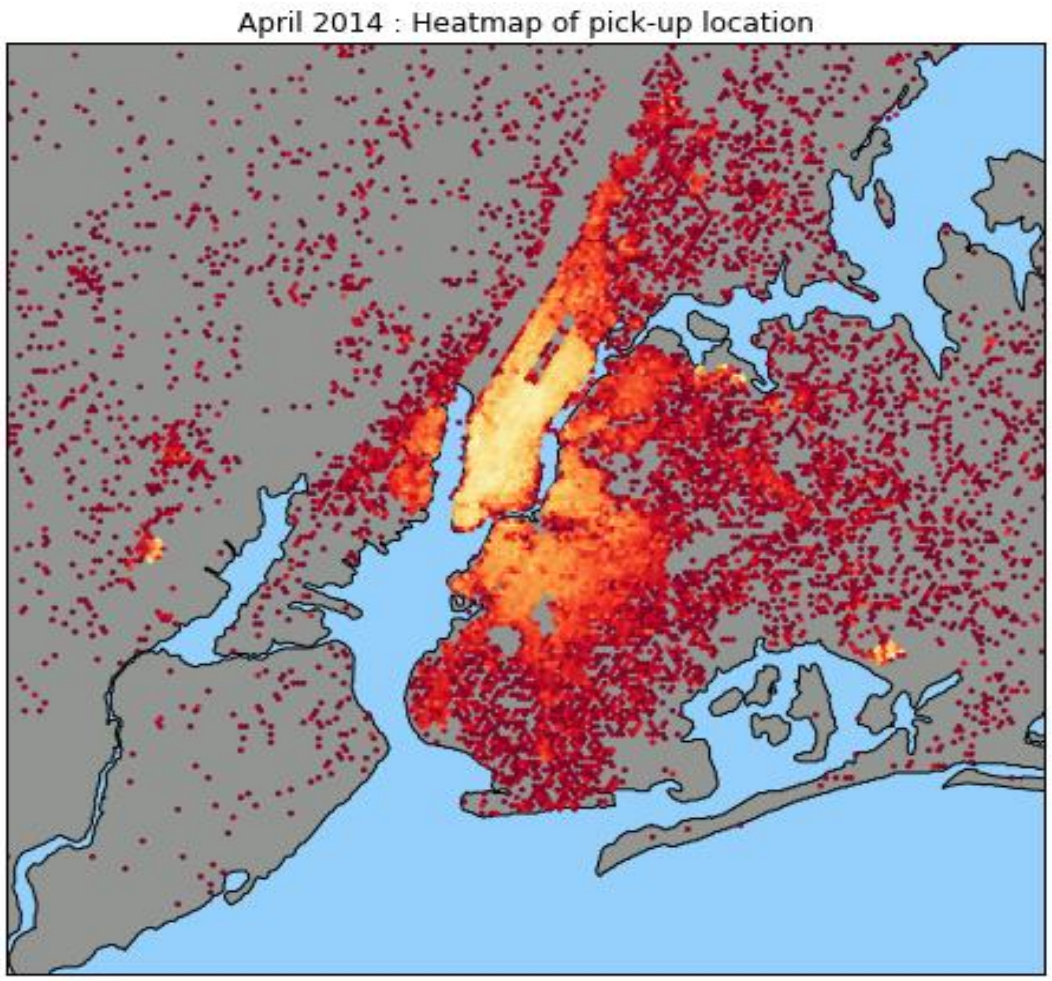


More rides at :

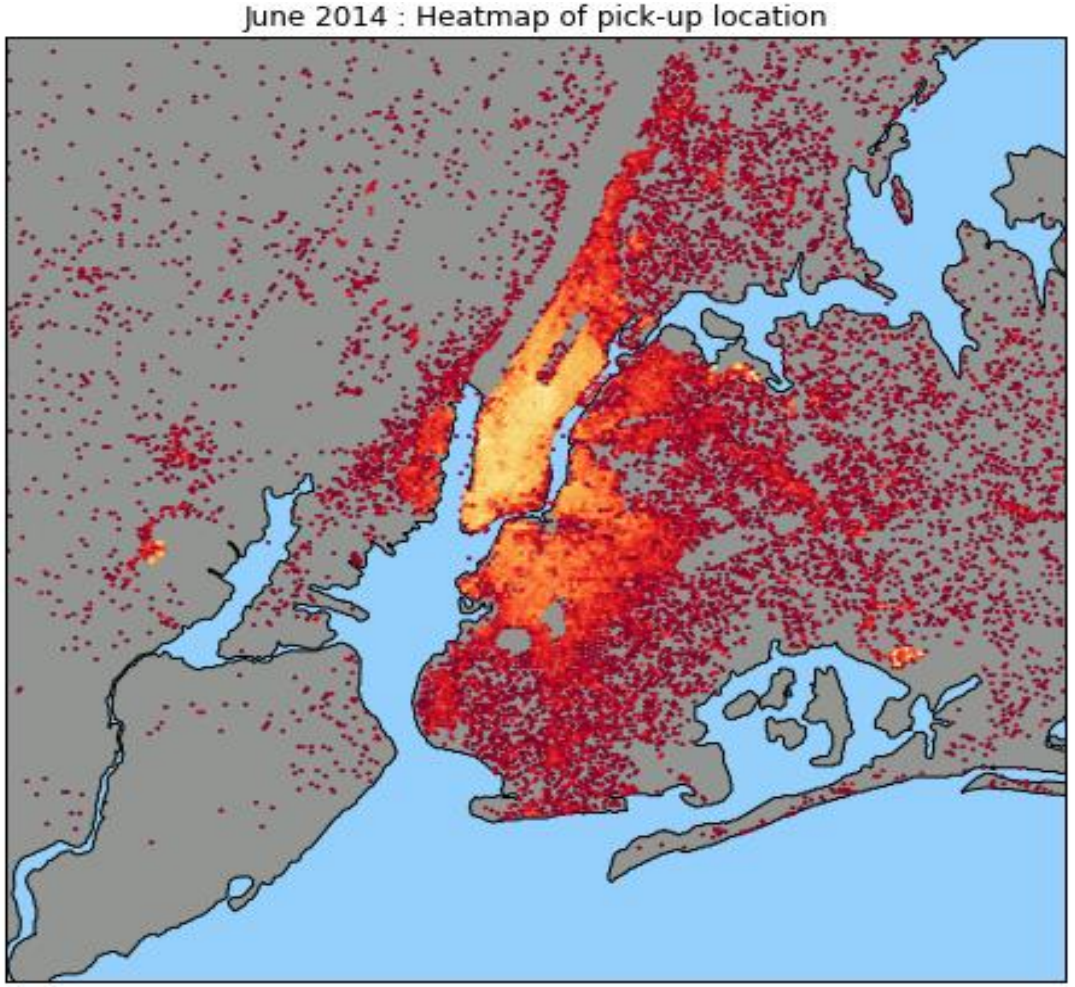
- 16 – 19PM (suggest : leave office, come home)
- 19 PM – 0AM (suggest : go-out)

Heatmap of pick-up in 2014

April



June

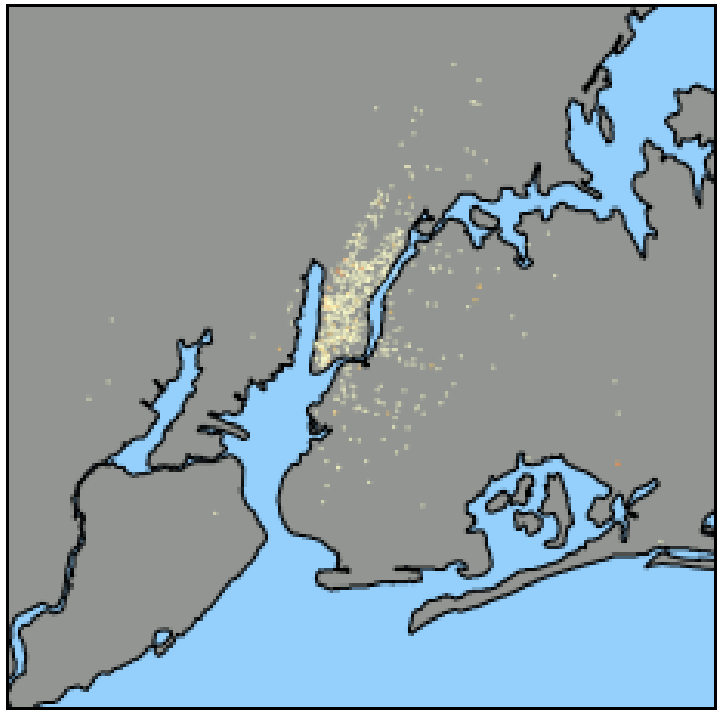


Almost rides are in the centers of New York

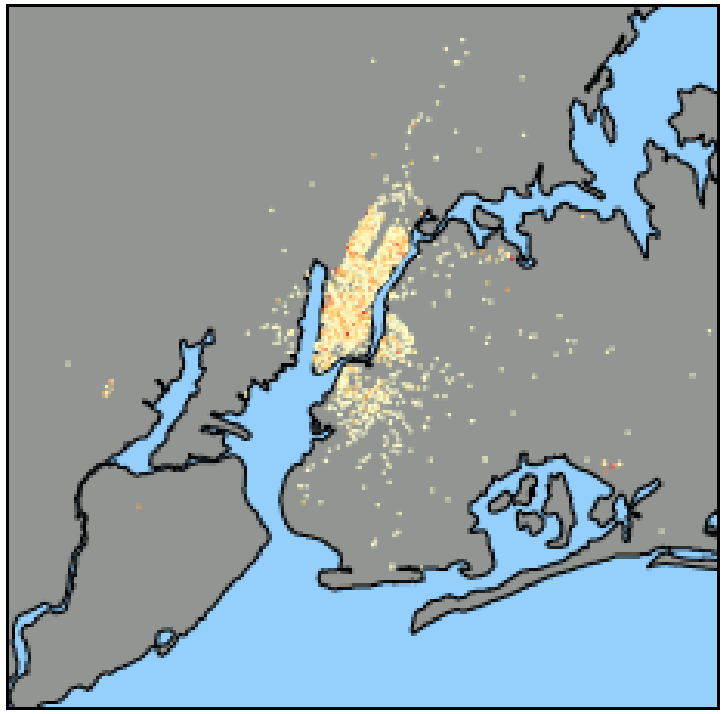
Heatmap of pick-up in 2014

Friday – April 2014

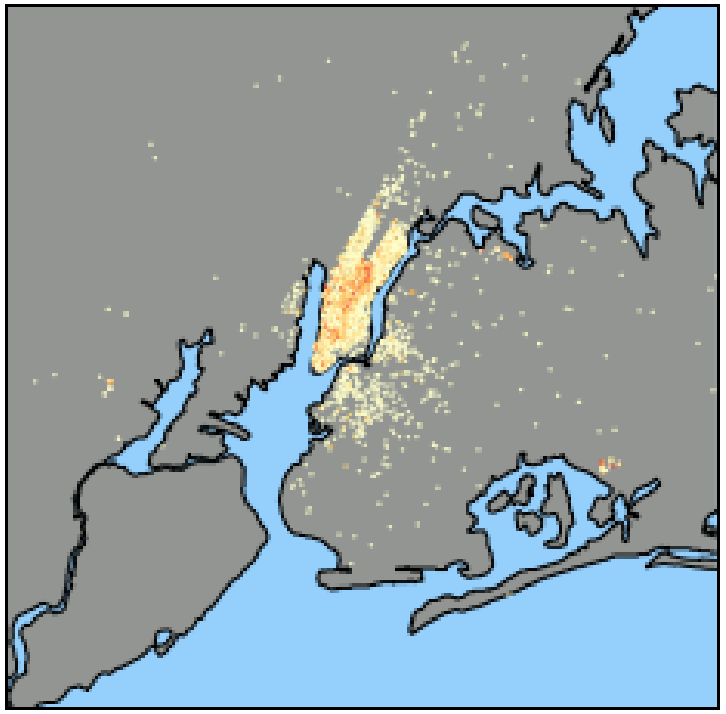
3 AM



7 AM



17 PM



CLUSTERING of pick-up region

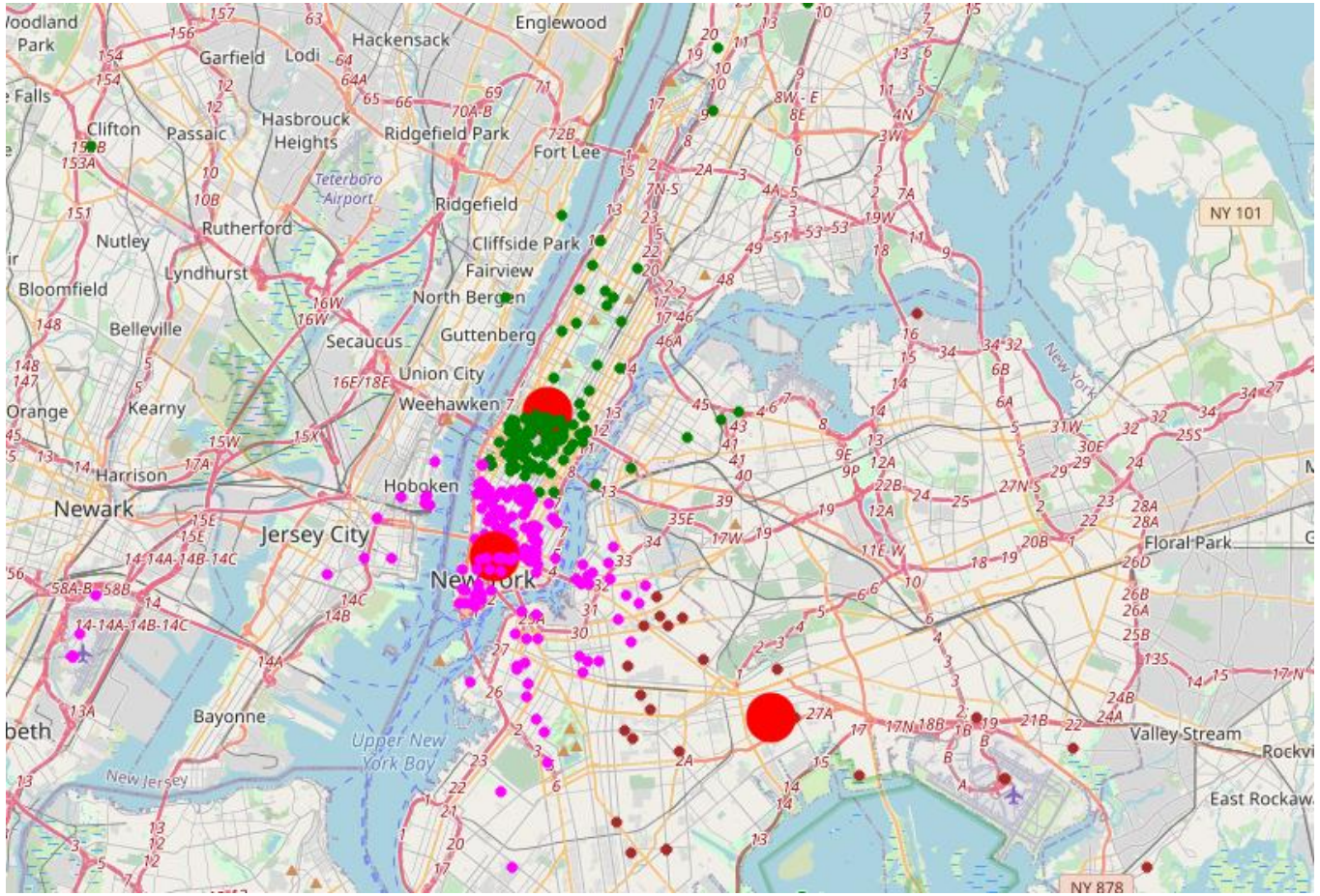
2 techniques

KMeans

DBSCAN

KMeans

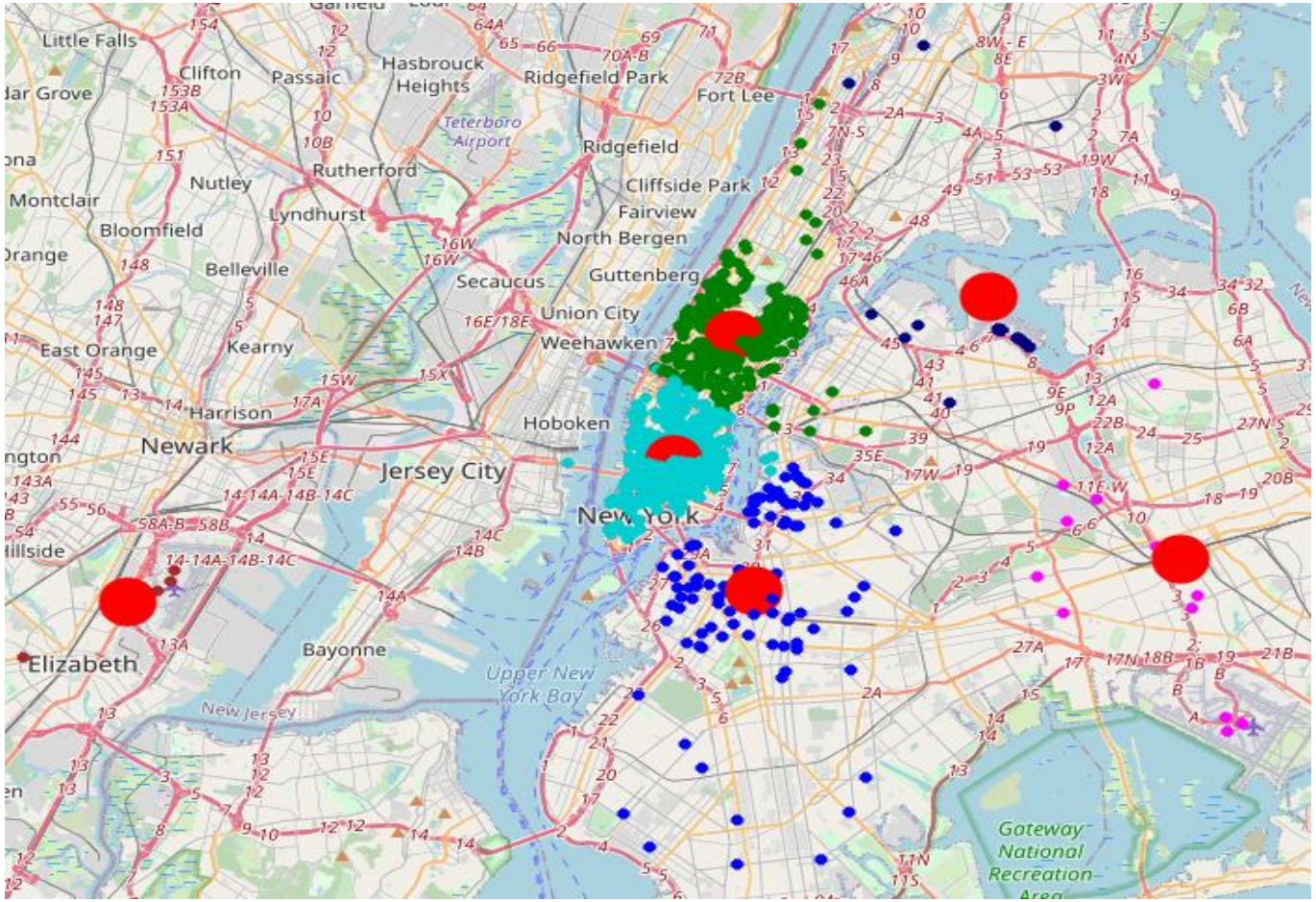
Day : 10/04/2014 Hours : **0 – 3 AM**



3 clusters
with quite large zones

KMeans

Day : 10/04/2014 Hours : **7 – 9 AM**

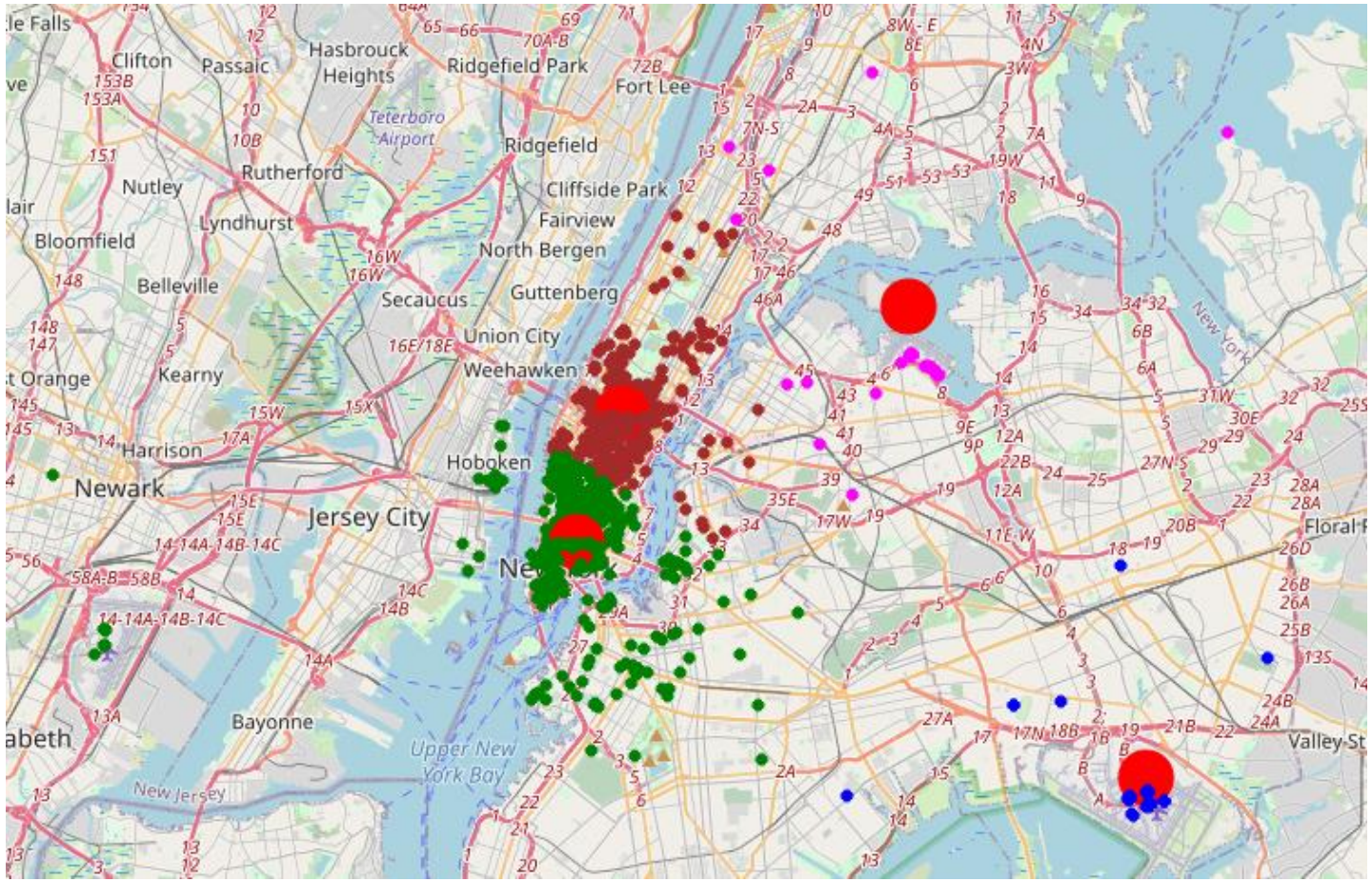


6 clusters

- 3 clusters with few drives (brown, pink, navy)
- 2 clusters with too large zones (blue, green)

KMeans

Day : 10/04/2014 Hours : **21 – 23 PM**

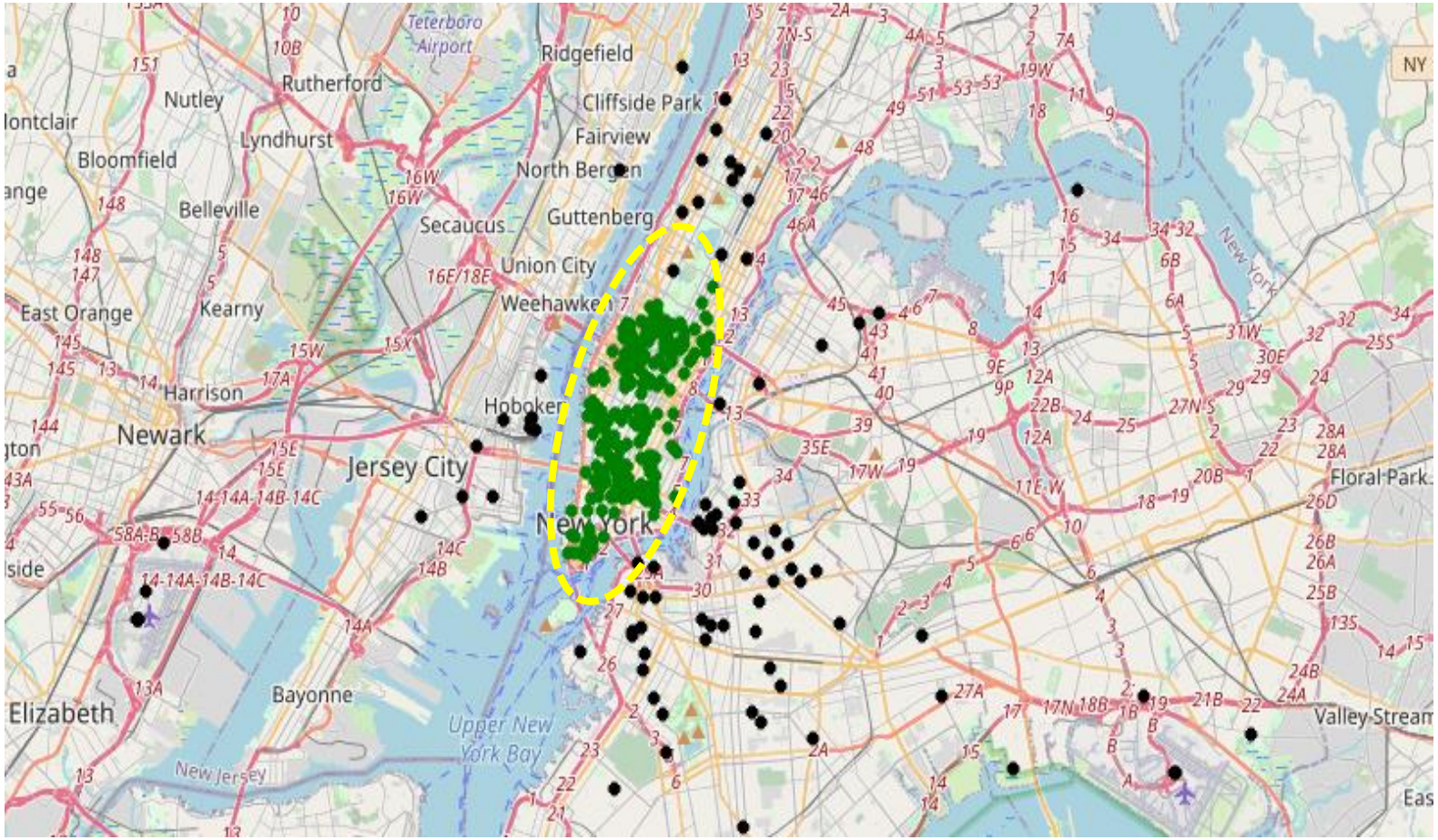


4 clusters

- 2 clusters with few drives (blue, pink)
- 1 cluster with too large zones (green)

DBSCAN

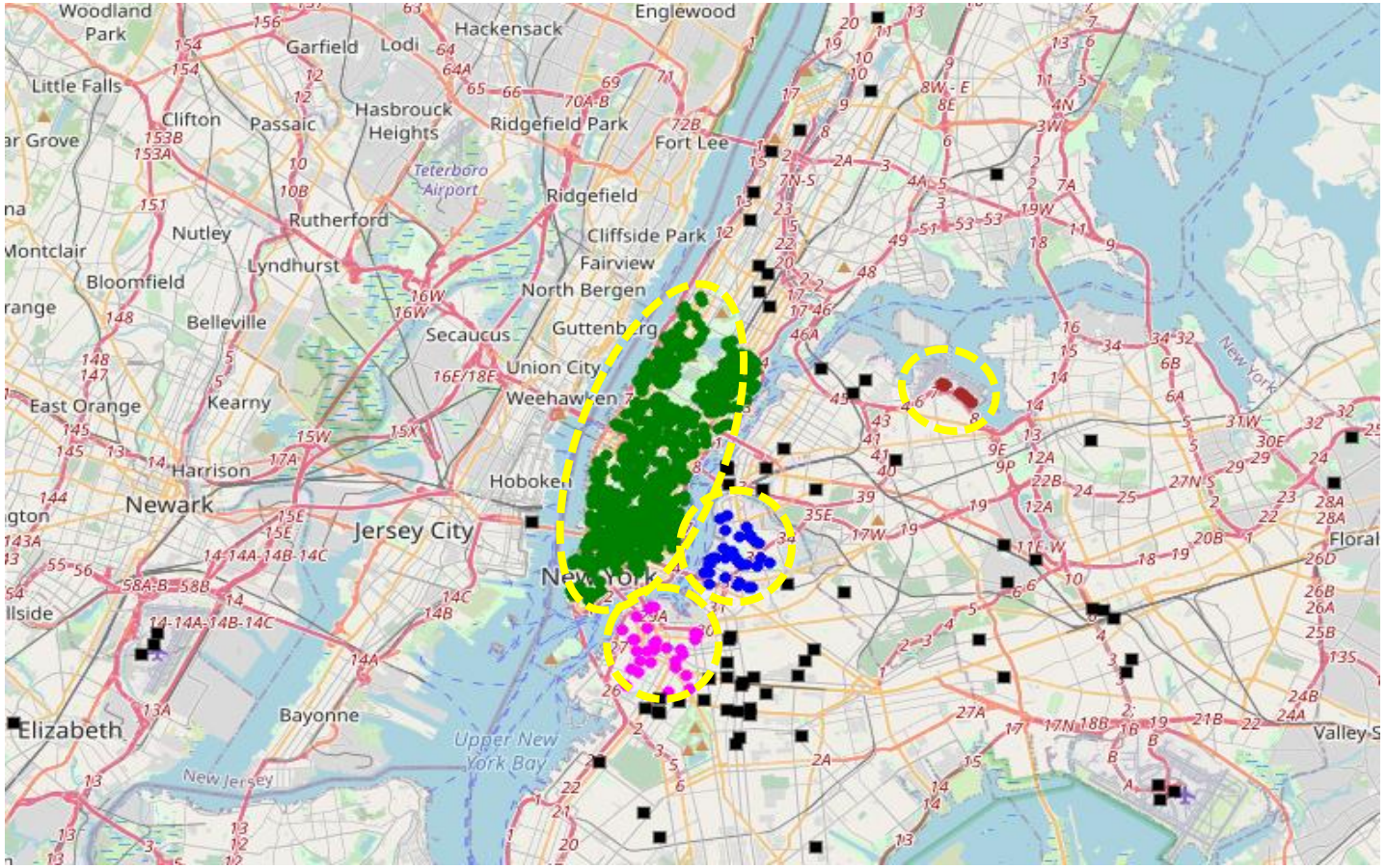
Day : 10/04/2014 Hours : **0 – 3 AM**



- 1 clusters (center NY)
- **Black : Outliers (far from center NY)**

DBSCAN

Day : 10/04/2014 Hours : **7 – 9 AM**

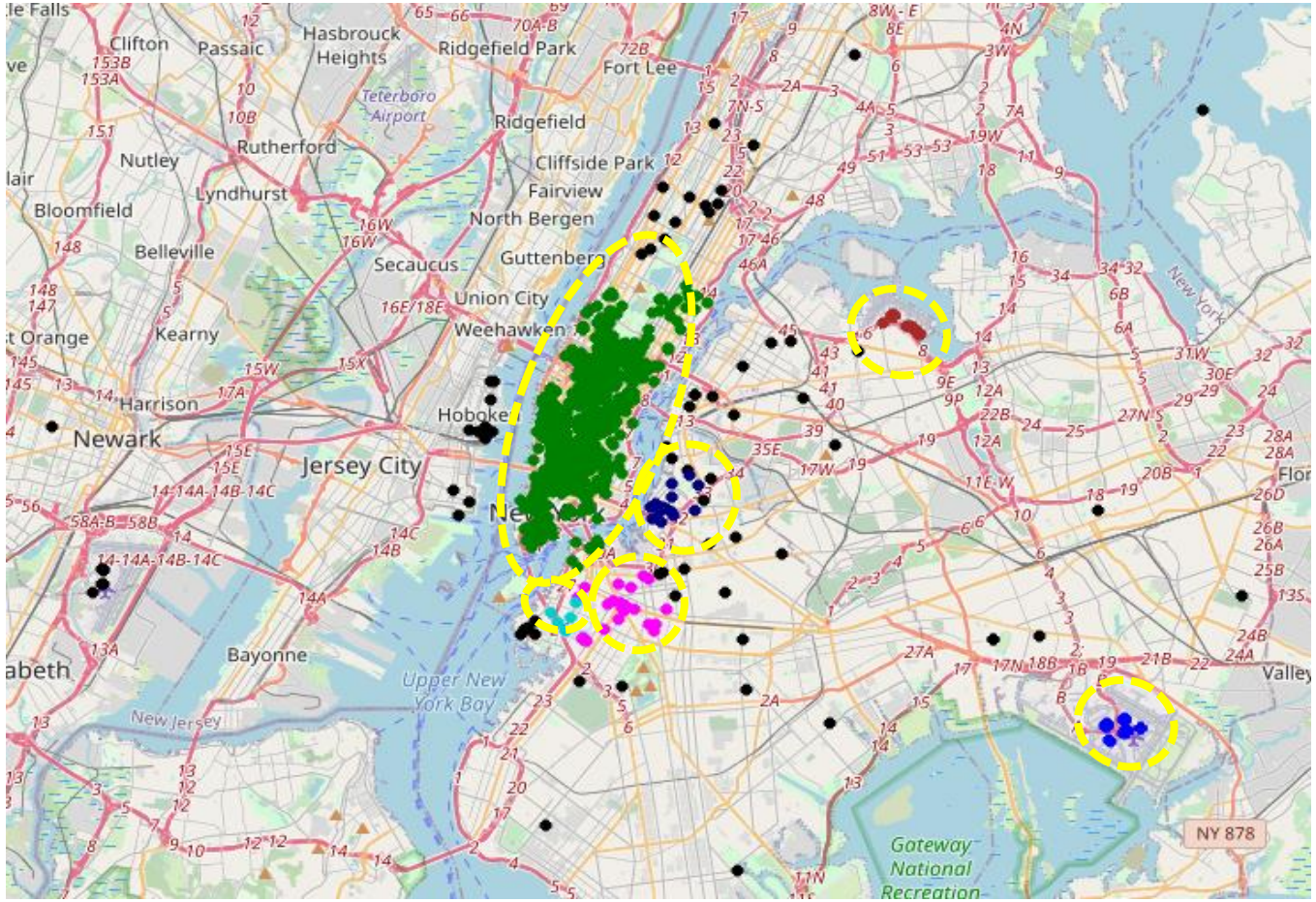


4 clusters

- Green : center NY
- Blue : near center NY
- Pink : near center NY
- Brown : aeroport
- **Black : Outliers (far from center NY)**

DBSCAN

Day : 10/04/2014 Hours : **21 – 23 PM**



6 clusters

- Green : center NY
- Navy : near center NY
- Pink : near center NY
- Cyan : near center NY
- Brown, Blue : airport
- **Black : Outliers (far from center NY)**

Conclusion

- Exploration data of pick-up's uber in 2014, 2015

  + Month, Day, Weekday, Hour

  + Density of rides

- Clustering pick-up zones by using KMeans and DBSCAN

  + KMeans : clusters quite well, but takes into account all points -> some clusters with little points (rides)

  + DBSCAN : clusters well, shows outliers points

-- > DBSCAN is better for clustering in this case.