New York City Taxi and Limousine Commission

Hardik Gupta (71620027), Reema Malhotra (71620057)

Executive Summary

Taxi transportation is a vital element in the transportation network of New York City (TEAM, 2014). In order to strike the right balance between demand and supply of taxis at different locations, at different times of the day, it is important for any taxi management or government authorities to plan and forecast the demand in order to provide better service to the commuters while maximising the profits.

This report highlights the results of the analysis for NYC Taxi and Limousine Commission TLC trip data and demonstrates how any business can accomplish the goals of balancing the demand while maximising profits using historical trip data such as date-time and locations of pick-up and drop-off, trip distances, trip durations, fare details, and others

The results are based on data of over 9 million green taxi trip records from January-June 2016. The analysis shows how using unsupervised learning algorithm such as K-means clustering, the city of NYC can be divided into 15 high pickup zones. Further using time series techniques such as Regression based models, ARIMA and Holt-Winters, how business can benefit by forecasting the revenue (overall) and dynamic demand for pickups in these zones over different time periods of the day. Based on these forecasted demands and average revenue for a ride spread across different zones and time period, the report highlights how using Linear Programming technique one can effectively optimize the use of existing fleet for better profits and effective service. The analysis also presents patterns of daily commute between different regions which can promote carpooling, enabling business to cater to larger demands and avoiding congestion on the road.

Business Problem

In order to improve the transportation network and optimize the availability of the taxis for New York City, following key business objectives are addressed

- Forecast the revenue from taxi services based on the day of the week for the first week of July 2016
- Forecasting demand for taxi, based on the day of the week, zone and time period of the day for the first week of July 2016
- Optimize the existing taxi availability at different location, different time period based on the demand and profits per ride
- Identify potential locations to promote pooled car services.

Methodology and Data

1. **Data -** Green taxi trip records from January-June 2016 (see references for data link)

2. **Data Cleaning and Pre-processing in R**

   a. **Raw Data Cleaning**

   The raw dataset contained over 9 million records. However, the data contained lot of discrepancies such as negative fare amount, zero trip distance, zero latitude etc. In order to handle such discrepancies and keep authentic data, an index is calculated – Fare_amount/Trip_distance and only those records are retained where the index value is between 2 and 10. Further the latitude and longitude values are rounded off to 3rd decimal places as this separates two data points only by a distance of 110m (Zhang, n.d.). Many derived variables are also created such as TripDuration, weekday, slot, zone etc.

   b. **Time Slot**

   Number of pickups varies across different hours of the day as well as across different regions. In order to accurately capture the demand, the entire data set is divided into 6 slots – each of 4 hours duration.

   Slot 1 – 12 to 4am, Slot 2 – 4 to 8am, Slot 3 – 8am to 12pm, Slot 4 – 12 to 4pm, Slot 5 – 4 to 8pm, and Slot 6 – 8pm to 12am

   c. **Creation of Zones**

   Using K-means clustering, the city is divided into 15 zones based on their pickup latitudes and longitudes. K-means is an unsupervised learning algorithm used for identifying homogenous groups ("clusters") of records. Since K-means inherently uses Euclidean distance method and because the dataset contained latitude and longitude in polar coordinates, the two dimensional dataset is first converted to three dimensional Euclidean feature space. Calinski-Harabasz index (CH index) is used to calculate the best number of clusters for the given data points. The index is ratio of between-cluster error to within-cluster error. Higher the CH index, the better is that cluster number for the dataset (refer appendix)

3. Time-Series Analysis in XLMiner and R

Time series is a technique used to predict the future values of a parameter. Different time series models are built using R and XLMiner to forecast revenue and demand using techniques such as Regression based, ARIMA and Holt-Winters

4. Linear Programming (Linear Optimization) using XLMiner

Linear programming is an optimization technique used to meet an objective given certain constraints, whose requirements can be solved as linear equations. Here LP is used to optimize the current demand (number of pickups) spread across different zones and time slots, and given the existing fleet maximise the profit generated by meeting these demands

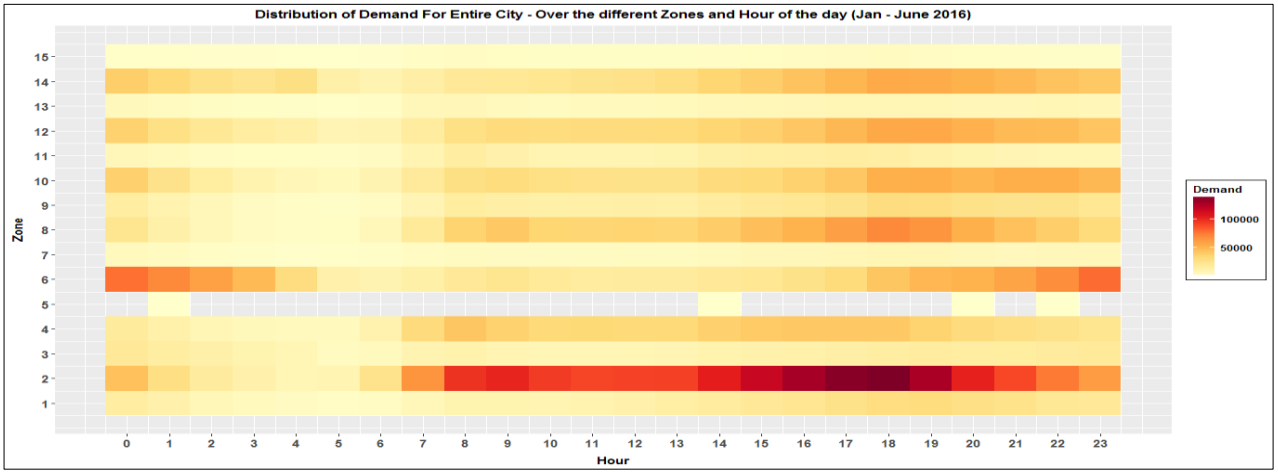
5. Tableau

Tableau 10 to identify different hotspots and areas where carpooling can be achieved

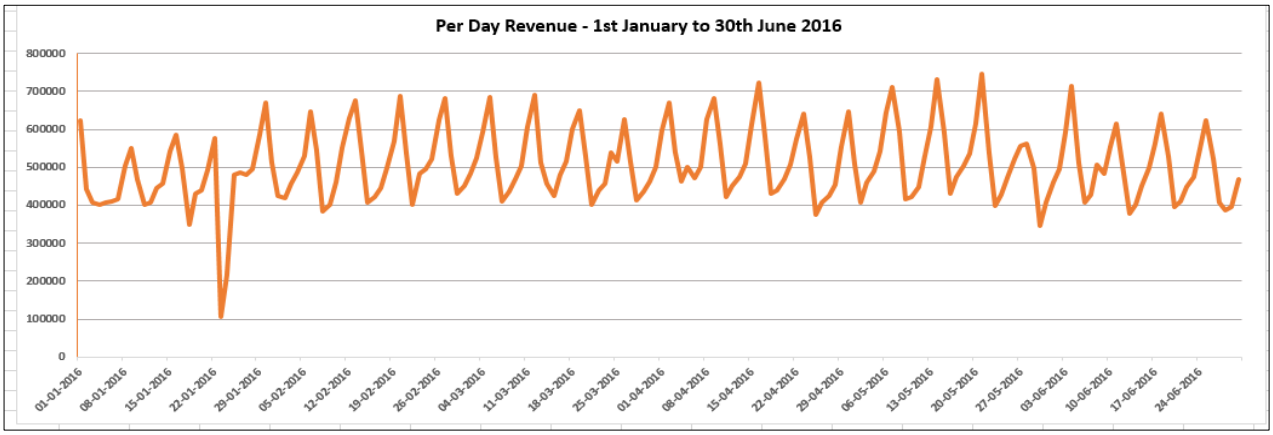
Analysis and Results

1. Distribution of demand over different zones and time period

Variation in demand across different parts of the city, day of the week (Sunday, Monday etc.) and hour of the day provided the necessary motivation to divide the city into 15 zones based on their location of pickups and further divide the time in 6 slots. A heat map shows this distribution of demand for all days across different zones and slots (Note: Zone 5 contained only 4 records and hence is not used for any analysis)



2. Forecasting the Revenue from the taxi services



A regression based model with 7 dummy variables - for capturing week of the day seasonality (Monday, Tuesday etc.) and an increasing trend is built having the final equation as

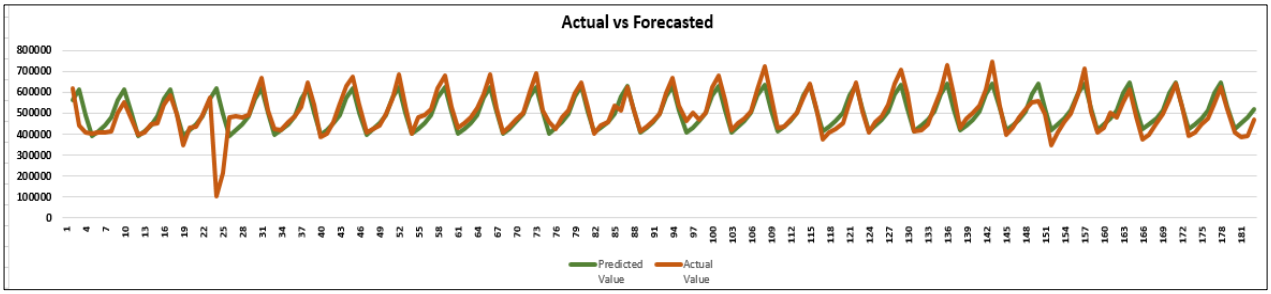
$$Y = B_0 + B_1t + B_2d_1 + B_3d_2 + ... + B_7d_6$$

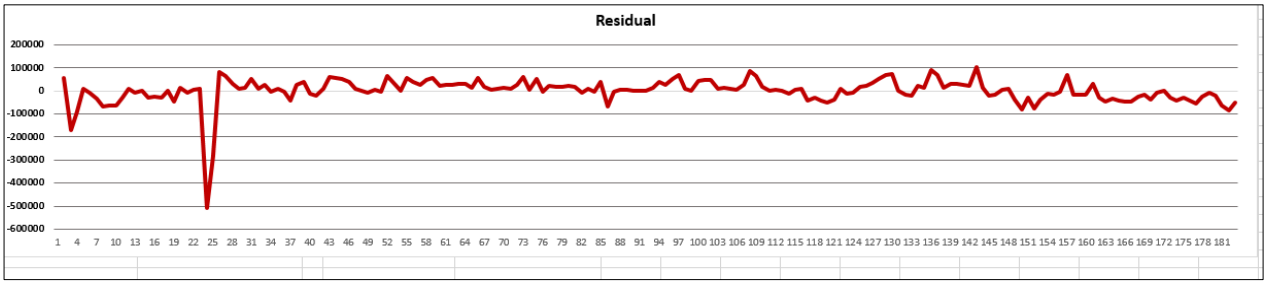
Summary Statistics of the Model

	Regression Model
Training RMSE	58914.07264
Validation RMSE	53725.63561
Training MAPE	9.24%
Validation MAPE	10.59%

	Regression Model
Overall RMSE	58088.8919
Overall MAPE	9.06%

Performance chart for Regression Model





The above charts show that a regression model fits well on the data, with overall MAPE as 9%

Forecast of Revenue using Regression Model (1st-7th July 2016)

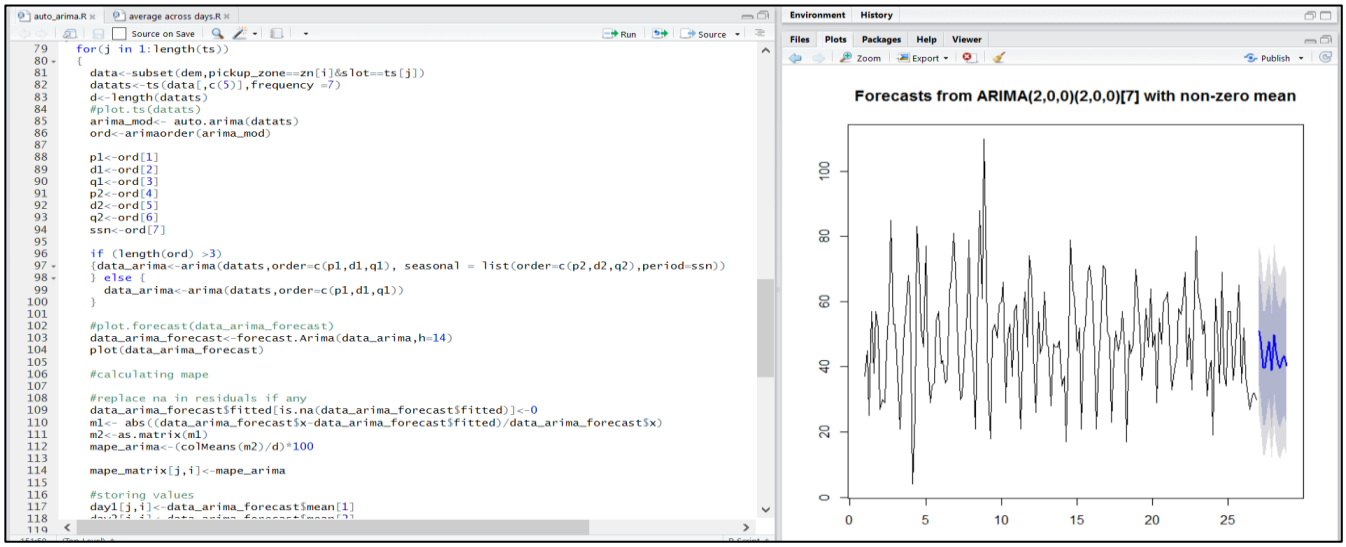
Predicted Value	95% Confidence Intervals		95% Prediction Intervals	
	Lower	Upper	Lower	Upper
601700.32	573887.32	629513.3078	481191.28	722209.35
650202.41	622389.42	678015.3997	529693.37	770711.4419
532500.68	504687.69	560313.6743	411991.65	653009.7166
428468.3	400655.31	456281.2917	307959.27	548977.3339
452903.06	425090.06	480716.0474	332394.02	573412.0896
481085.33	453272.34	508898.3236	360576.3	601594.3658
520030.12	492217.13	547843.1101	399521.08	640539.1523

3. Forecasting the taxi demand over different zones and different slots

In order to accurately forecast the demand for each slot and each zone, 84 (14 zones, 6 time slots) time series models are created. This is achieved by writing an algorithm in R which uses a function called "auto.arima". This function outputs the parameters of ARIMA (p, d, q values which stand for auto-correlation, seasonality and partial auto-correlation respectively) that are appropriate for a given time-series. MAPE for each model is recorded to monitor accuracy. Steps to build the model can be summarised as

- Subset the data to a particular zone and time slot
- Using the function 'auto.arima' in R, get the p,d,q values
- Build forecasting model to forecast for next 7 days, passing the values as obtained from step 2
- Evaluate the MAPE

An example of a model, coded in R and using ARIMA model for forecasting



```
for(j in 1:length(ts))
{
  data<-subset(dem,pickup_zone==zn[i]&slot==ts[j])
  datats<-ts(data[,c(5)],frequency =7)
  d<-length(datats)
  #plot.ts(datats)
  arima_mod<- auto.arima(datats)
  ord<-arimaorder(arima_mod)

  p1<-ord[1]
  d1<-ord[2]
  q1<-ord[3]
  p2<-ord[4]
  d2<-ord[5]
  q2<-ord[6]
  ssn<-ord[7]

  if (length(ord) >3)
  {data_arima<-arima(datats,order=c(p1,d1,q1), seasonal = list(order=c(p2,d2,q2),period=ssn))
  } else {
    data_arima<-arima(datats,order=c(p1,d1,q1))
  }

  #plot.forecast(data_arima_forecast)
  data_arima_forecast<-forecast.Arima(data_arima,h=14)
  plot(data_arima_forecast)

  #calculating mape

  #replace na in residuals if any
  data_arima_forecast$fitted[is.na(data_arima_forecast$fitted)]<-0
  m1<- abs((data_arima_forecast$x-data_arima_forecast$fitted)/data_arima_forecast$x)
  m2<-as.matrix(m1)
  mape_arima<-(colMeans(m2)/d)*100

  mape_matrix[j,i]<-mape_arima

  #storing values
  day1[j,i]<-data_arima_forecast$mean[1]
```


MAPE Values for each of 84 models

SLOT\ZONE	1	2	3	4	6	7	8	9	10	11	12	13	14	15
MAPE VALUES (IN %)														
1	0.14114261	0.15950239	1.25596877	0.63768569	0.2326816	0.13342041	0.19707333	0.34749325	0.21438476	0.2371495	0.185669	0.33693033	0.12488857	0.29138679
2	0.19596892	0.144354	0.27063756	0.15069797	1.26197728	0.12038663	0.30009296	0.14977716	0.29988795	0.12908235	0.61568242	0.14896662	0.96672893	0.29333628
3	0.0958267	0.10587114	0.0793361	0.11349573	0.14253121	0.12957748	0.16475756	0.12243625	0.10378128	0.1419143	0.09683574	0.08325175	0.08684841	0.31693376
4	0.08647565	0.11249059	0.10977201	0.09901715	0.25640226	0.07931359	0.22289758	0.15149542	0.15954877	0.08909945	0.11137614	0.08682931	0.12822765	0.11036918
5	0.28008624	0.58332642	0.45364391	0.38328	0.081527	0.54967199	0.86780201	0.74889111	3.52940403	0.23987097	0.22444618	0.18479365	0.31431391	0.13963988
6	1.56421508	5.99979212	0.09829495	0.08173248	0.09209866	0.07161172	1.81761795	0.09773548	2.08915048	0.10168407	3.89341877	0.15718009	0.92178782	0.15879386

The maximum error is 6% while the least is 0.08%. Overall ARIMA models fit well on the data for all 84 models

Forecast of Demand using ARIMA Model across all Zone sand Slots (Shown only for 1st July 2016)

	day1	day1	day1	day1	day1	day1	day1	day1	day1	day1	day1	day1	day1	day1
1	365.917267	883.524762	613.451646	154.977095	2250.61944	109.920509	424.366794	427.484364	739.252946	168.818597	787.142499	54.0598306	996.221499	36.0387776
2	81.5156643	514.734746	161.721391	221.508228	449.417491	36.057715	154.720965	80.6358789	224.561359	91.6593935	260.947991	62.787433	312.647596	21.2909355
3	247.779808	1954.37208	195.676454	775.065847	417.953278	73.9095457	777.734636	333.381673	592.658214	234.808196	609.837968	121.629395	484.187307	78.5223444
4	359.434234	2175.60436	182.172729	780.006664	355.900105	100.41821	891.771512	274.217658	709.646555	254.698375	690.673429	149.082919	603.461052	60.4002093
5	615.308677	2686.47876	269.134005	803.02448	701.192756	165.934068	1346.59377	587.995252	965.437621	279.340366	1093.09442	187.655185	1132.61735	82.2842814
6	564.613491	1798.8206	431.862108	695.293553	1533.1004	173.293012	985.826252	659.847339	1108.30551	227.623076	1274.60625	190.299789	1146.28511	50.9292533

4. Optimizing the taxi availability at different location

Objective for the business is to maximise the revenue generated from taxi services given the constraints of demand and supply. Demand here is the forecasted number of rides. In order to quantify constraints of supply two sets of information is required – Number of cabs available during a time slot and relation between number of rides and number of cabs. Since this information is not present, following assumptions are made

- It is assumed that the highest number of rides which NYC can cater to (across all zones for a slot) is the maximum number of rides it had catered to in the past for a given day and slot (see appendix). This is a high-level assumption which is used to develop constraints for Supply.
- Average revenue per unit ride for each zone-time slot, each day of the week is calculated (by processing historical data, see appendix for example).
- The objective for optimisation problem is to maximise revenue given the constraints of demand and supply. An additional assumption is - NYC would ensure that it caters to a minimum number of rides for each time-slot and zone. This threshold should come from business, however for this analysis the threshold is considered to be 10.

Optimization Formulation (shown for 1st July 2016)

Day -1 01/07/2016- WEDNESDAY

Forecast of Demands- number of pick ups

	zone1	zone2	zone3	zone4	zone6	zone7	zone8	zone9	zone10	zone11	zone12	zone13	zone14	zone15	Sum
slot1	365	880	610	150	2250	105	420	425	735	165	785	50	995	35	7970
slot2	80	510	160	220	445	35	150	80	220	90	260	60	310	20	2640
slot3	245	1950	195	775	415	70	775	330	590	230	605	120	480	75	6855
slot4	355	2175	180	780	355	100	890	270	705	250	690	145	600	60	7555
slot5	615	2685	265	800	700	165	1345	585	965	275	1090	185	1130	80	10885
slot6	560	1795	430	695	1530	170	985	655	1105	225	1270	190	1145	50	10805

Revenue Per Ride

	zone1	zone2	zone3	zone4	zone6	zone7	zone8	zone9	zone10	zone11	zone12	zone13	zone14	zone15
slot1	9	9	10	9	11	11	11	10	11	9	9	11	8	10
slot2	12	10	14	12	13	14	15	14	14	12	11	15	10	12
slot3	13	10	14	11	13	13	15	14	13	11	11	15	11	11
slot4	12	11	15	12	13	16	14	12	12	12	11	16	10	14
slot5	11	10	13	12	11	14	12	11	11	11	10	15	10	10
slot6	9	10	11	10	11	11	11	10	11	10	9	12	9	8

Optimized Demand * Revenue per Ride

464371

Optimized Demand (Decision variables)

	zone1	zone2	zone3	zone4	zone6	zone7	zone8	zone9	zone10	zone11	zone12	zone13	zone14	zone15	Sum		
slot1	10	10	10	10	2250	105	420	10	376	10	10	10	10	10	3251	<=	3251
slot2	80	510	160	220	445	35	150	80	220	90	260	60	310	20	2640	<=	2992
slot3	245	1950	195	775	415	70	775	330	590	230	605	120	480	75	6855	<=	9667
slot4	355	2175	180	780	355	100	890	270	705	250	690	145	600	60	7555	<=	11538
slot5	615	2685	265	800	700	165	1345	585	965	275	1090	185	1130	80	10885	<=	12861
slot6	560	1795	430	695	1530	170	985	655	1105	225	1270	190	855	10	10475	<=	10475

After optimization and analysing the results for 7 days, the following table shows where NYC would not be able to meet the demand (which day, slot and zone)

Demand not met

Day-slot	zone1	zone2	zone3	zone4	zone5	zone6	zone7	zone8	zone9	zone10	zone11	zone12	zone13	zone14	zone15	Non-zero demand (Sum)
Day1-slot1	-355	-870	-600	-140	0	0	0	-415	-359	-155	-775	-40	-985	-25		-4719
Day1-slot6	0	0	0	0	0	0	0	0	0	0	0	0	-290	-40		-330
Day2-slot1	0	-508	0	0	0	0	0	0	0	0	-397	0	-596	-8		-1509
Day5-slot3	0	-1430	0	0	0	0	0	0	0	0	0	0	0	0		-1430
Day6-slot6	0	0	0	0	0	0	0	0	0	0	-63	0	-1076	-36		-1175
Day7-slot1	0	0	0	0	0	0	0	0	0	0	0	0	-503	0		-503

It is observed from above that for slots 1 and 6 across zones 1,2,3,4,9,10,11,12,13 and 14, NYC is likely to not meet its demand of rides. In order to address this, the first step is to estimate the numbers of cabs that would be required to meet these number of rides. The data is at a ride level and not cab level, which is why the number of cabs which would be required to fulfil the above demand cannot be estimated. Once that is estimated, NYC can choose to have more cabs available for this time slot.

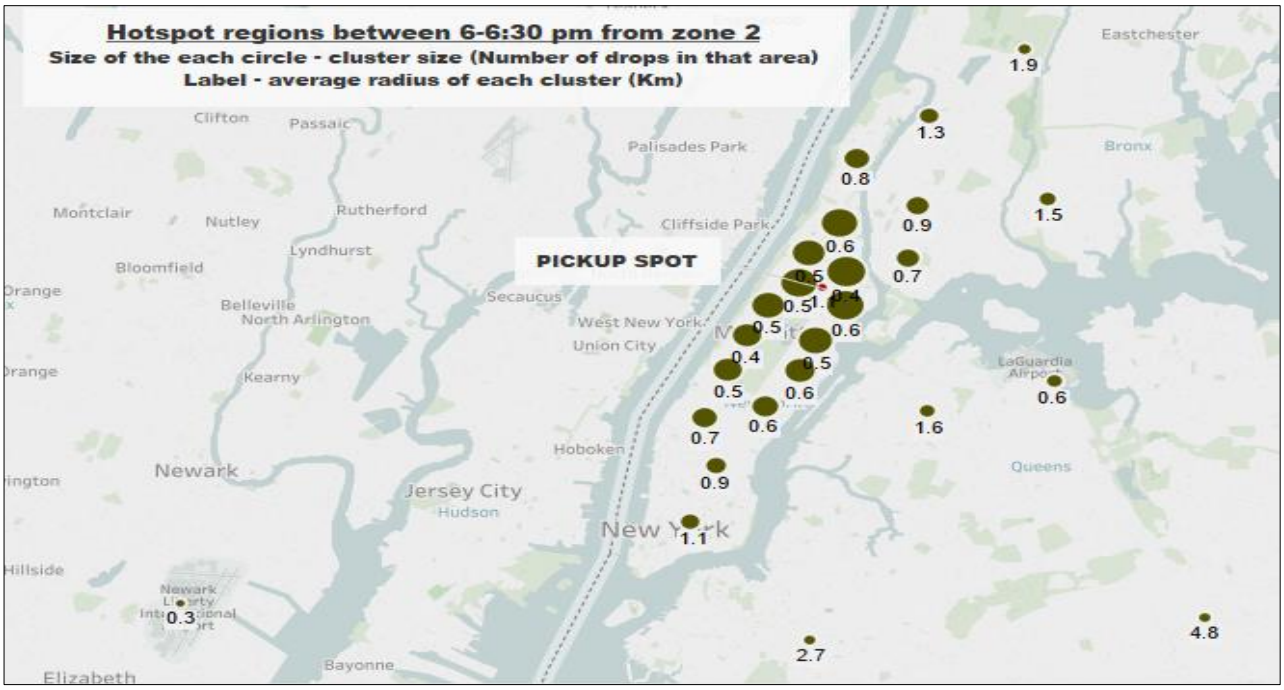
Our analysis do not suggest buying new cars because as per the data, NYC is capable of catering to more number of rides than what is predicted. Inspecting slot 1, it is observed that there are less number of cabs available as compared to rest of the slots like slot 5 and slot 6. It may be because slot 1 is an odd hour (12-4 am). A proposed solution could be that NYC could look for drivers who are available to work on a shift that includes slot 1, this can be an economical solution to above problem.

5. Potential locations to promote pooled car services

Pooled car services can be effective when the pickup and drop locations are either at a minimum distance or is in the same or minimum deviated direction and the start time of travel is within an acceptable time (around 20 mins). At the start we clustered different locations into zones based on pickups. We consider these zones as the starting point. In order to identify areas which shows potential for carpooling, we performed following steps

1. Identify a particular zone with highest number of pickups – Zone 2 (for this example). Next for the selected zone, identify the hour which has maximum number of pickups (starting point) – Hour 18 (refer section 'Distribution of demand over different zones and time period' for this)
2. Divide each hour into 30 mins slot. For better accuracy this slot can be further decreased to 20mins and hence will get 3 slots per hour
3. Perform K-Means clustering using Dropoff Latitudes and Longitudes. This will cluster regions which are closer to each other. We ran K-means and identified 25 drop off regions

4. Identify the centroid, dropoff cluster size, median distance of each point from their respective centroids (radius of the cluster). Also using R, identify the address of each centroid point
5. Plot the data using Tableau

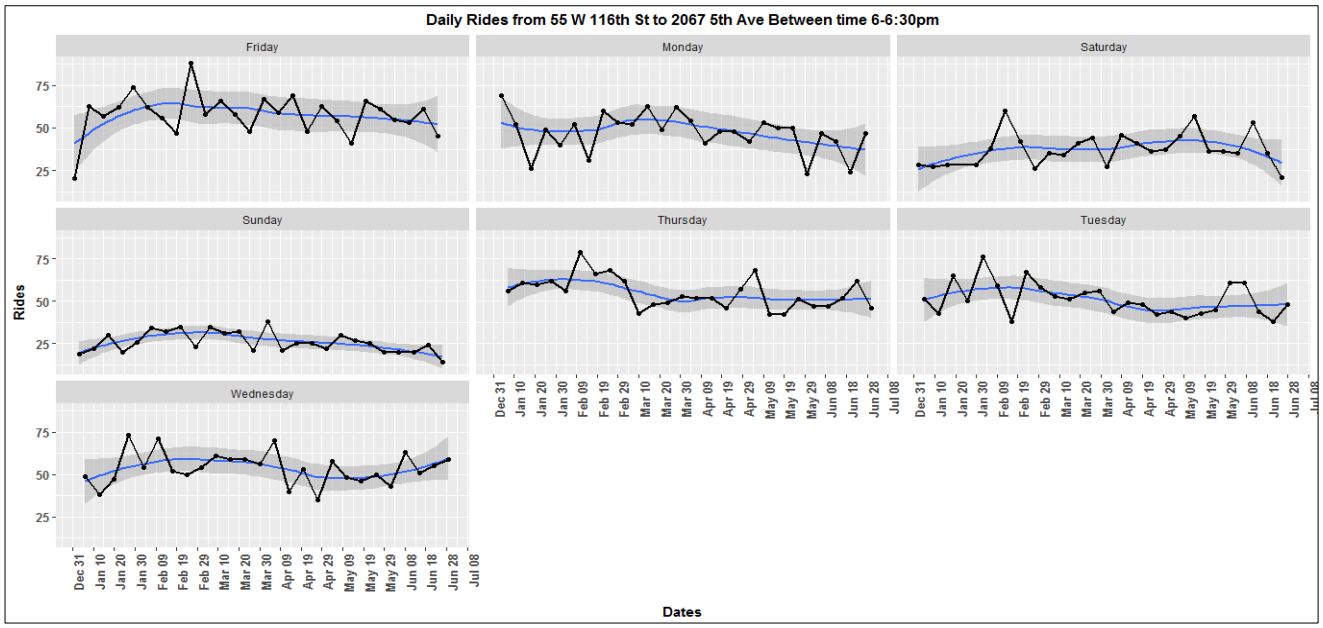


Top 15 Dropoff regions

CENTER OF PICKUP		
latitude	longitude	textAddress
40.80194303	-73.94907667	55 W 116th St, New York, NY 10026

CENTER OF DROP LOCATIONS				
latitude	longitude	No. of Drops	Average distance (km)	Locations
40.80767077	-73.9408099	8485	0.4373	2067 5th Ave, New York, NY 10035
40.79597882	-73.94111048	6943	0.5796	226 E 113th St, New York, NY 10029
40.80374092	-73.95707691	6925	0.4991	313 W 114th St, New York, NY 10026
40.82438872	-73.9430389	6598	0.6229	267 Edgecombe Ave, New York, NY 10031
40.78367992	-73.95136331	6375	0.5289	181 E 93rd St, New York, NY 10128
40.81386912	-73.95364712	5940	0.5282	449 W 128th St, New York, NY 10027
40.79571039	-73.96836059	5289	0.4747	160 W 100th St, New York, NY 10025
40.77288081	-73.95665544	4719	0.5515	244 E 78th St, New York, NY 10075
40.78547834	-73.97587913	3523	0.4422	168 W 83rd St, New York, NY 10024
40.77321853	-73.98288749	2417	0.515	132 W 65th St, New York, NY 10023
40.76080076	-73.9695856	2302	0.6464	696 Lexington Ave, New York, NY 10022
40.75667722	-73.99110187	1740	0.7458	327 W 41st St, New York, NY 10036
40.84669912	-73.93677283	1518	0.7509	61 Wadsworth Ave, New York, NY 10033
40.81237519	-73.91848186	1403	0.7375	447 E 144th St, Bronx, NY 10454
40.83025514	-73.91505904	1252	0.9174	280 E 166th St, Bronx, NY 10456
40.740108	-73.98712227	1218	0.8929	48 E 23rd St, New York, NY 10010

The above data suggests there is definitely potential to carpool starting from zone2 and starting time is between 6-6:30 pm. The following time series graph also suggest that there is a steady to and fro (except for weekends) between the starting point (55 W 116th St) and drop location 2067 5th Ave (drop off region with maximum number of drop offs)



A similar analysis can be performed for any other zone and time period

Conclusion

The analysis conducted can help NYC analyse and deploy the correct number of cabs in different parts of the city with an objective to maximise the revenue. Even with certain assumptions, NYC can maximise the profits by targeting the correct number of pickups in different parts of the city at different hours with the existing fleets. Identification of carpooling areas can suggest introduction of new services between the high demand-drop zones, leading to efficient utilization of existing fleet. All the above analysis can be further improved with increased number of zones, better precision and accuracy can be achieved with more time slots and dynamic models capturing special days can be built to cater to all sorts of challenges in future.

References

TEAM, R. R. (2014, December 16). *case-study-new-york-city-taxis-596.* Retrieved from Datasmart City Solutions: http://datasmart.ash.harvard.edu/news/article/case-study-new-york-city-taxis-596

Zhang, S. (n.d.). *How precise is one degree of longitude or latitude?* Retrieved from Gizmodo: http://gizmodo.com/how-precise-is-one-degree-of-longitude-or-latitude-1631241162

Dataset - http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml

Code and Data Files

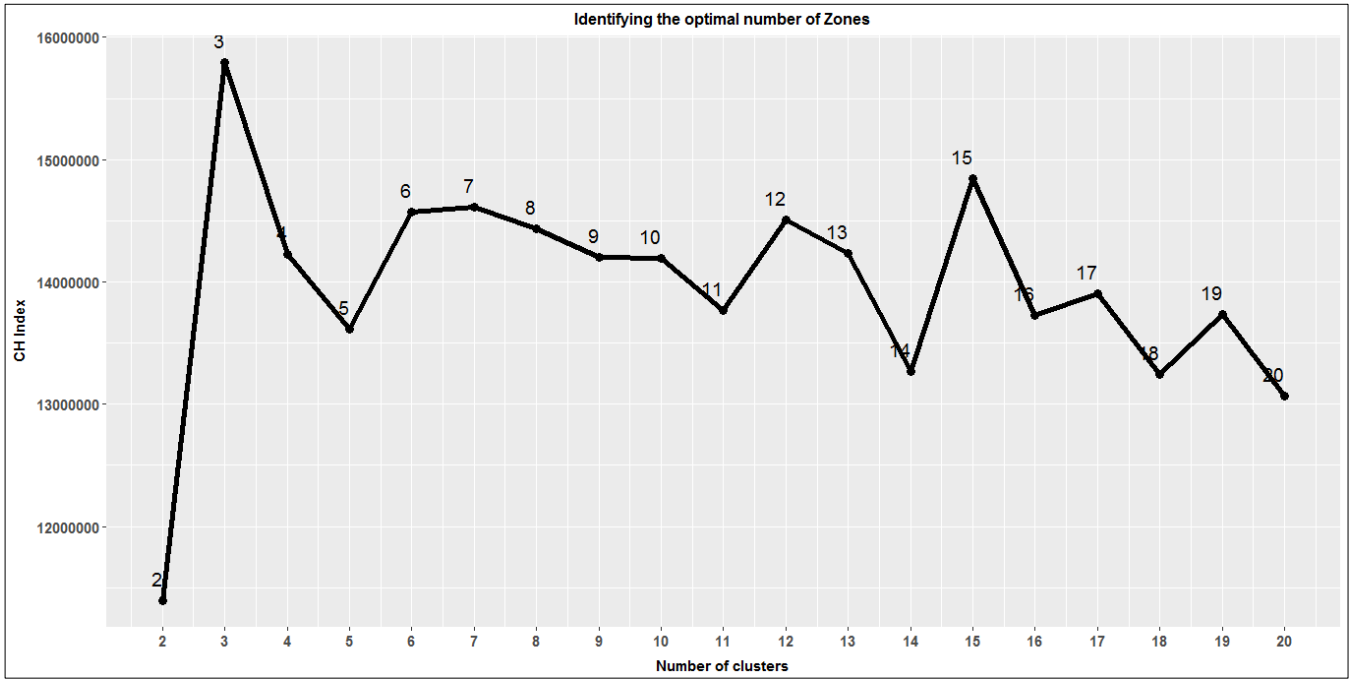
- Folder 'R Codes' – Contains all the R codes used for cleaning, pre-processing the dataset, demand forecasting, and identification of hotspot for carpooling. The necessary data files that are generated after each processing is kept in folder 'data files'. (One raw file which is of size 1.9 GB is not uploaded. The same can be send upon request. This file contains the aggregated 82L trip records obtained after cleaning and pre-processing the original >90L records
- Folder 'DemandFoecasting' – Results of 84 time series models that are built to forecast demand for 7 days starting from 1st July 2016
- Folder 'RevenueForecasting' – Results of Regression based model used to forecast revenue for 7 days starting from 1st July 2016
- Folder 'Optimization' – Results of performing LP optimization to target maximum profits, calculated for each of the 7 days starting 1st July 2016. Also contains the last sheet contains information of all those areas where demand is not met
- Folder 'Supporting files' – Contains few charts of hotspots, and tableau workbook for the same

Apendix

1. Summary of the dataset

```
 lpep_pickup_datetime Lpep_dropoff_datetime Pickup_longitude Pickup_latitude Dropoff_longitude Dropoff_latitude Passenger_count
Length:8235198       Length:8235198        Min.   :-82.16   Min.   :36.63   Min.   :-82.17    Min.   :17.85    Min.   :0.000
Class :character     Class :character      1st Qu.:-73.96   1st Qu.:40.69   1st Qu.:-73.97    1st Qu.:40.70    1st Qu.:1.000
Mode  :character     Mode  :character      Median :-73.95   Median :40.75   Median :-73.95    Median :40.75    Median :1.000
                                           Mean   :-73.94   Mean   :40.75   Mean   :-73.94    Mean   :40.75    Mean   :1.358
                                           3rd Qu.:-73.92   3rd Qu.:40.80   3rd Qu.:-73.92    3rd Qu.:40.79    3rd Qu.:1.000
                                           Max.   :-67.90   Max.   :43.17   Max.   :-49.32    Max.   :48.12    Max.   :9.000
 Trip_distance       Fare_amount      pickup_date        pickup_time        pickup_hour     dropoff_date       dropoff_time
Min.   :  0.0161   Min.   :   0.07   Length:8235198     Length:8235198     Min.   : 0.00   Length:8235198     Length:8235198
1st Qu.:  1.6737   1st Qu.:   6.00   Class :character   Class :character   1st Qu.: 9.00   Class :character   Class :character
Median :  2.8807   Median :   9.00   Mode  :character   Mode  :character   Median :15.00   Mode  :character   Mode  :character
Mean   :  3.9467   Mean   :  11.15                                         Mean   :13.71
3rd Qu.:  5.1338   3rd Qu.:  13.50                                         3rd Qu.:19.00
Max.   :450.8738   Max.   :3347.50                                         Max.   :23.00
 dropoff_hour        slot        TripDuration        weekday
Min.   : 0.00   Min.   :1.000   Min.   :   0.000   Length:8235198
1st Qu.: 9.00   1st Qu.:3.000   1st Qu.:   6.017   Class :character
Median :15.00   Median :4.000   Median :   9.833   Mode  :character
Mean   :13.67   Mean   :4.052   Mean   :  21.732
3rd Qu.:19.00   3rd Qu.:5.000   3rd Qu.:  16.117
Max.   :23.00   Max.   :6.000   Max.   :1439.983
```

2. CH Index to identify correct number of Zones



3. Maximum number of rides during a particular slot for the given day of the week

weekday	slot	Maxdemand
Friday	1	22317
Friday	2	9110
Friday	3	9435
Friday	4	9084
Friday	5	15956
Friday	6	15931

weekday	slot	Maxdemand
Saturday	1	11299
Saturday	2	3800
Saturday	3	7469
Saturday	4	11996
Saturday	5	16978
Saturday	6	16872

weekday	slot	Maxdemand
Sunday	1	13114
Sunday	2	4956
Sunday	3	5804
Sunday	4	9963
Sunday	5	13492
Sunday	6	11465

weekday	slot	Maxdemand
Monday	1	6605
Monday	2	3206
Monday	3	9211
Monday	4	8463
Monday	5	12064
Monday	6	8043

weekday	slot	Maxdemand
Tuesday	1	2635
Tuesday	2	3084
Tuesday	3	9275
Tuesday	4	9149
Tuesday	5	13292
Tuesday	6	9902

weekday	slot	Maxdemand
Wednesday	1	3251
Wednesday	2	2992
Wednesday	3	9667
Wednesday	4	11538
Wednesday	5	12861
Wednesday	6	10475

weekday	slot	Maxdemand
Thursday	1	3761
Thursday	2	3194
Thursday	3	9592
Thursday	4	9323
Thursday	5	15161
Thursday	6	12140

4. Average revenue per ride (for each day of the week, each slot and each zone)

An example for Friday (similar analysis is done for all other days of the week)

Average Revenue Per Ride for Friday

	Zone1	Zone2	Zone3	Zone4	Zone5	Zone6	Zone7	Zone8	Zone9	Zone10	Zone11	Zone12	Zone13	Zone14	Zone15
slot1	9	9	10	9	8	11	11	12	12	11	9	9	12	8	9
slot2	13	10	14	12	NA	14	15	14	14	14	12	11	14	10	13
slot3	13	11	14	11	NA	13	14	14	13	13	11	11	15	11	12
slot4	12	11	15	12	NA	14	15	14	13	13	12	11	16	11	14
slot5	11	11	14	12	NA	11	14	13	12	12	12	10	14	11	11
slot6	10	10	11	11	NA	11	11	12	11	11	11	9	13	10	10