# Capstone Final Project

Introduction:

Every business or service is required to focus assets to the areas that it will provide the most benefit. In the transportation industry, this means focusing vehicles and drivers where they will be needed most. Different locations will have peak demand at different times of day, and different times of the week. There are also certain areas that will only have a high demand due to special events, such as sporting events, concerts, or other similar high demand times. To provide the greatest service, as well as gain the most income from providing the service, it is necessary to have the drivers and vehicles in those locations when the demand hits. This provides more opportunities for drivers to attract passengers, and shorter wait times for passengers which leads to higher satisfaction.

Data:

I will be using the neighborhood and borough information for New York City, NY, as well as Foursquare location data to analyze venues and likely high demand areas and times of day or week. The Foursquare data along with the NYC data will help me divide information into small regions and corridors that are likely going to need transportation services. There will be locations that are mostly serviced during the day, such as shopping centers, banks, pharmacies, while others will require peak services at other times. Areas with a high number of bars or clubs will require higher supply of vehicles late into the evening and early morning, while some venues like sports arenas will regularly require little to no supply outside of special events. The location data can then be used to plan high supply corridors, and to stage vehicles in preparation for high demand or special events.

Methodology:

I began with the neighborhood data for New York City, NY. This information breaks down NYC into small chunks that can be used to categorize regions based on venues that are present in the area. I then used the Foursquare API to pull information about different venues that are located in all the neighborhoods in NYC. I then checked how many unique categories existed in the data. I took those venues and grouped them into more general venue categories. It would have weighted the results to leave restaurants broken out into dozens of different categories. So I reduced the number of categories from 432 down to 16. This allowed for a simpler analysis of the neighborhoods and the pertinent venues. From there I clustered the neighborhoods using KMeans. This allowed me to put neighborhoods with similar breakdown of venues into categories. I then took those clusters and mapped all the NYC neighborhoods in order to visualize the data breakdown.

Results:

I categorized NYC neighborhoods into 4 clusters. I then did an analysis of transportation needs based on the most prevalent venues in those clusters.

Cluster 1:

|  | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
| --- | --- | --- | --- | --- | --- | --- |
| 0 | Wakefield | Dining | Shopping | Services | Sports/Arenas | Religion |
| 1 | Co-op City | Shopping | Dining | Transportation | Sports/Arenas | Outdoors |
| 5 | Kingsbridge | Dining | Shopping | Bars/Clubs | Outdoors | Gallery/Exhibit |
| 6 | Marble Hill | Dining | Shopping | Fitness | Services | Sports/Arenas |
| 8 | Norwood | Dining | Shopping | Outdoors | Services | Sports/Arenas |
| ... | ... | ... | ... | ... | ... | ... |
| 290 | Middle Village | Dining | Shopping | Outdoors | Services | Bars/Clubs |
| 295 | Highland Park | Dining | Shopping | Outdoors | Fitness | Transportation |
| 298 | Allerton | Dining | Shopping | Services | Outdoors | Fitness |
| 300 | Erasmus | Shopping | Dining | Fitness | Outdoors | Performing arts |
| 301 | Hudson Yards | Dining | Shopping | Fitness | Bars/Clubs | Services |

Cluster 2:

|  | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
| --- | --- | --- | --- | --- | --- | --- |
| 3 | Fieldston | Outdoors | Transportation | Sports/Arenas | Religion | Gallery/Exhibit |
| 4 | Riverdale | Outdoors | Services | Transportation | Sports/Arenas | Fitness |
| 18 | West Farms | Dining | Transportation | Outdoors | Sports/Arenas | Services |
| 27 | Clason Point | Outdoors | Transportation | Shopping | Fitness | Dining |
| 35 | Spuyten Duyvil | Shopping | Fitness | Dining | Outdoors | Services |
| 76 | Mill Island | Services | Fitness | Sports/Arenas | Religion | Outdoors |
| 78 | Coney Island | Dining | Sports/Arenas | Entertainment | Outdoors | Gallery/Exhibit |
| 85 | Sea Gate | Services | Outdoors | Transportation | Sports/Arenas | Religion |
| 91 | Bergen Beach | Services | Sports/Arenas | Outdoors | Shopping | Dining |
| 127 | Battery Park City | Dining | Shopping | Outdoors | Gallery/Exhibit | Bars/Clubs |
| 164 | South Jamaica | Dining | Transportation | Shopping | Services | Sports/Arenas |
| 176 | Edgemere | Dining | Transportation | Outdoors | Shopping | Hotels etc |
| 177 | Arverne | Outdoors | Dining | Transportation | Shopping | Sports/Arenas |
| 183 | Jamaica Estates | Services | Dining | Transportation | Sports/Arenas | Religion |
| 198 | New Brighton | Transportation | Outdoors | Shopping | Dining | Entertainment |
| 202 | Grymes Hill | Services | Fitness | Dining | Transportation | Sports/Arenas |
| 204 | South Beach | Outdoors | Dining | Shopping | Transportation | Sports/Arenas |
| 207 | Port Ivory | Transportation | Sports/Arenas | Religion | Outdoors | Gallery/Exhibit |
| 210 | Travis | Dining | Hotels etc | Services | Fitness | Bars/Clubs |
| 212 | Oakwood | Services | Bars/Clubs | Transportation | Sports/Arenas | Religion |
| 224 | Park Hill | Shopping | Transportation | Fitness | Dining | Hotels etc |
| 226 | Graniteville | Dining | Transportation | Shopping | Sports/Arenas | Religion |
| 238 | Butler Manor | Sports/Arenas | Fitness | Shopping | Transportation | Religion |
| 245 | Bloomfield | Shopping | Outdoors | Transportation | Entertainment | Social |
| 255 | Emerson Hill | Services | Sports/Arenas | Religion | Outdoors | Gallery/Exhibit |
| 256 | Randall Manor | Transportation | Outdoors | Dining | Sports/Arenas | Religion |
| 265 | Pomonok | Outdoors | Dining | Shopping | Bars/Clubs | Transportation |
| 275 | Stuyvesant Town | Transportation | Outdoors | Services | Dining | Bars/Clubs |
| 286 | Sandy Ground | Services | Shopping | Transportation | Sports/Arenas | Religion |
| 288 | Roxbury | Outdoors | Dining | Sports/Arenas | Bars/Clubs | Religion |
| 294 | Malba | Fitness | Transportation | Sports/Arenas | Religion | Outdoors |
| 302 | Hammels | Outdoors | Dining | Transportation | Services | Shopping |
| 304 | Queensbridge | Hotels etc | Dining | Outdoors | Sports/Arenas | Performing arts |
| 305 | Fox Hills | Dining | Bars/Clubs | Transportation | Sports/Arenas | Religion |

Cluster 3:

|  | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
| --- | --- | --- | --- | --- | --- | --- |
| 172 | Breezy Point | Outdoors | Gallery/Exhibit | Sports/Arenas | Religion | Performing arts |
| 179 | Neponsit | Outdoors | Sports/Arenas | Religion | Gallery/Exhibit | Performing arts |
| 192 | Somerville | Outdoors | Sports/Arenas | Religion | Gallery/Exhibit | Performing arts |
| 203 | Todt Hill | Outdoors | Sports/Arenas | Religion | Gallery/Exhibit | Performing arts |
| 303 | Bayswater | Outdoors | Sports/Arenas | Religion | Gallery/Exhibit | Performing arts |

Cluster 4:

|  | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
| --- | --- | --- | --- | --- | --- | --- |
| 2 | Eastchester | Dining | Transportation | Shopping | Entertainment | Sports/Arenas |
| 7 | Woodlawn | Dining | Outdoors | Bars/Clubs | Shopping | Transportation |
| 9 | Williamsbridge | Dining | Bars/Clubs | Sports/Arenas | Religion | Outdoors |
| 11 | Pelham Parkway | Dining | Services | Performing arts | Shopping | Fitness |
| 13 | Bedford Park | Dining | Transportation | Shopping | Bars/Clubs | Sports/Arenas |
| ... | ... | ... | ... | ... | ... | ... |
| 292 | Lighthouse Hill | Dining | Outdoors | Gallery/Exhibit | Services | Sports/Arenas |
| 293 | Richmond Valley | Dining | Services | Transportation | Sports/Arenas | Religion |
| 296 | Madison | Dining | Services | Shopping | Fitness | Sports/Arenas |
| 297 | Bronxdale | Dining | Shopping | Outdoors | Performing arts | Services |
| 299 | Kingsbridge Heights | Dining | Shopping | Transportation | Outdoors | Sports/Arenas |

Discussion:

These clusters allow for some planning for when the highest likely demand will occur, and what other factors may play into surges in demand.

Cluster 1

Dining and shopping dominates these neighborhoods. They are likely to need the majority of transportation during regular business hours and early evening. The arenas and performing arts venues may need additional resources flexed to them for major events.

Cluster 2

There exists a large portion of public and other forms of transportation already in place. This will reduce the amount of additional services required. The high number of outdoors venues will make transportation needs very weather dependent. There are some arenas and performing arts venues that will need to be flexed for when they hold a major event.

Cluster 3

This group of neighborhoods will need very minimal transportation on a regular basis. It is dominated by outdoors venues, arenas, galleries, and religious venues. This means that outside of special events, there will be almost no demand for transportation on a regular basis.

Cluster 4

Dining dominates these neighborhoods, but there are an above average number of bars and clubs. So this means that there will be a higher than average demand for transportation in the evening and into the early morning.

Conclusion:

Grouping neighborhoods into categories allows for prediction and analysis of what the likely transportation needs will be. This will allow for pre-staging of drivers and vehicles before the surge of demand, which will allow for higher customer satisfaction as well as more customers serviced, which leads to greater profits.