

CAPSTONE PROJECT - BATTLE OF NEIGHBOURHOODS

*Identification of less Tennis facilitated
Neighborhoods in North York, Toronto*

Introduction/Business Problem

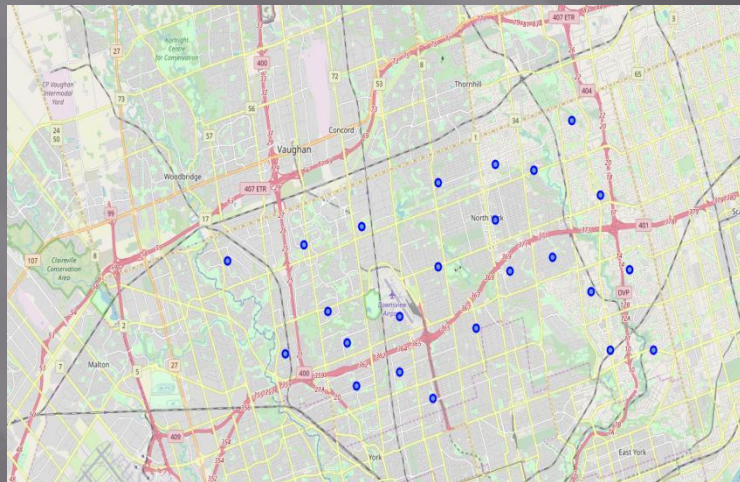
- ▣ **Background :** In preparation for the Summer olympics 2024, Sports Authority of Canada in collaboration with each local government and individual sports Federations assessing the current facilities and planning to improve the sports facilities all over Canada within two years.
- ▣ **Business Problem:** Identification of the locations for new Tennis facilities in Toronto neighborhoods.
- ▣ **Stakeholders:** Sports Authority, Tennis Federation, local governments and students to prove the necessity of a new Tennis facility in areas of low facility.

Data Acquisition and Preparation

- ▣ **Data Sources** : Wikipedia, Toronto neighborhood geospatial data and Foursquare
- ▣ **Data cleaning** : Removed unwanted data, missing data were replaced with appropriate values. Duplicates removed
- ▣ **Data preparation and Feature selection**: Considered North York Borough with 24 neighborhoods, education institutions and Tennis facilities, Geo-coordinates, distance from neighborhood center, data encoding of venues, normalization of distances

Methodology

- ▣ Exploratory Data Analysis: various statistical tables, graphs, geographical maps and machine learning techniques.
- ▣ Toronto map with north york neighborhoods-



Methodology

▣ North York neighborhoods-

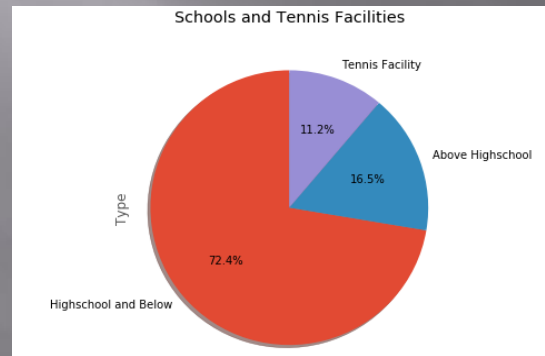
| Neighborhood | Above Highschool sum | Highschool and Below sum | Tennis Facility sum | AHS_dist mean | HSB_dist mean | TF_dist mean |
|---|-------------------------|-----------------------------|------------------------|------------------|------------------|-----------------|
| Bathurst Manor, Downsview North, Wilson Heights | 0 | 4 | 0 | 0.000000 | 636.500000 | 0.000000 |
| Bayview Village | 0 | 4 | 0 | 0.000000 | 780.000000 | 0.000000 |
| Bedford Park, Lawrence Manor East | 0 | 8 | 0 | 0.000000 | 691.250000 | 0.000000 |
| CFB Toronto, Downsview East | 1 | 2 | 0 | 237.000000 | 309.000000 | 0.000000 |
| Don Mills North | 2 | 9 | 4 | 33.866667 | 365.066667 | 141.733333 |
| Downsview Central | 1 | 6 | 0 | 126.714286 | 661.142857 | 0.000000 |
| Downsview Northwest | 1 | 6 | 0 | 107.857143 | 620.857143 | 0.000000 |
| Downsview West | 1 | 5 | 0 | 114.333333 | 728.500000 | 0.000000 |
| Downsview, North Park, Upwood Park | 0 | 5 | 0 | 0.000000 | 843.600000 | 0.000000 |
| Emery, Humberlea | 0 | 1 | 0 | 0.000000 | 680.000000 | 0.000000 |
| Fairview, Henry Farm, Oriole | 3 | 9 | 1 | 184.769231 | 594.000000 | 33.538462 |
| Flemingdon Park, Don Mills South | 0 | 2 | 0 | 0.000000 | 821.500000 | 0.000000 |
| Glencairn | 2 | 7 | 2 | 151.545455 | 646.181818 | 133.000000 |
| Hillcrest Village | 1 | 6 | 2 | 50.666667 | 465.111111 | 97.333333 |
| Humber Summit | 0 | 3 | 0 | 0.000000 | 727.666667 | 0.000000 |
| Lawrence Heights, Lawrence Manor | 1 | 5 | 1 | 158.142857 | 469.714286 | 170.285714 |
| Newtonbrook, Willowdale | 3 | 4 | 0 | 345.285714 | 476.285714 | 0.000000 |
| Northwood Park, York University | 0 | 1 | 2 | 0.000000 | 258.666667 | 624.000000 |
| Parkwoods | 1 | 8 | 1 | 109.700000 | 592.600000 | 92.400000 |
| Silver Hills, York Mills | 0 | 4 | 1 | 0.000000 | 300.200000 | 150.400000 |
| Victoria Village | 0 | 4 | 0 | 0.000000 | 779.250000 | 0.000000 |
| Willowdale South | 8 | 9 | 3 | 215.400000 | 344.300000 | 122.400000 |
| Willowdale West | 1 | 5 | 0 | 36.833333 | 785.166667 | 0.000000 |
| York Mills West | 2 | 6 | 2 | 170.300000 | 559.500000 | 155.000000 |

▣ Venue summary -

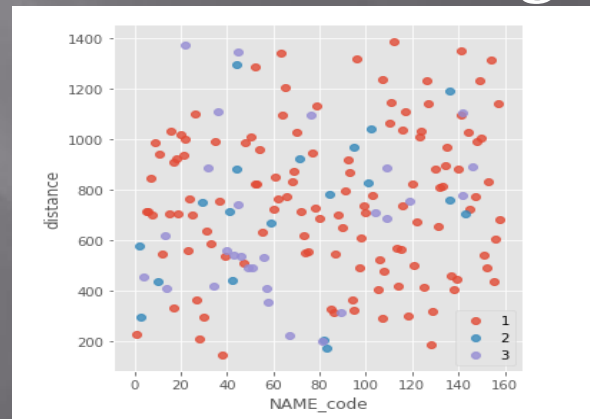
| | Type |
|----------------------|------|
| Highschool and Below | 123 |
| Above Highschool | 28 |
| Tennis Facility | 19 |

Methodology

▣ Education Institutions and Tennis Facility -

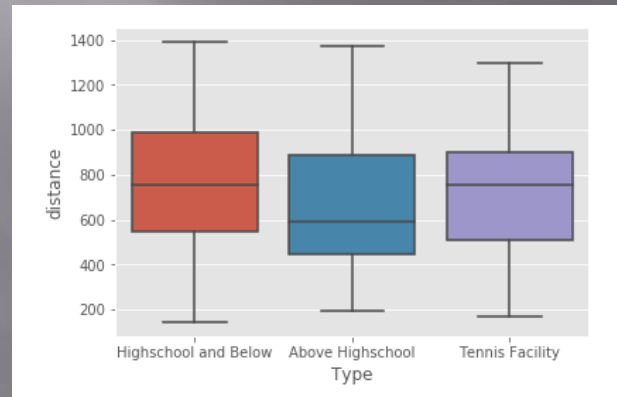


▣ Venue distance from neighborhood center -

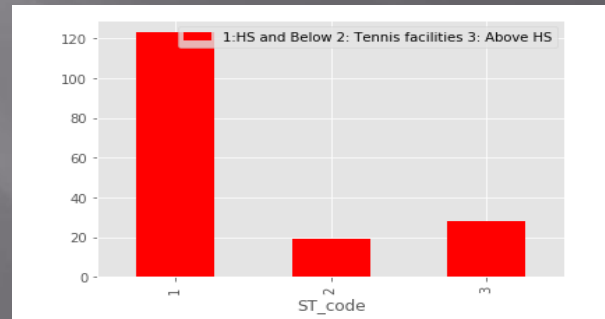


Methodology

- ▣ Box plot showing distance for categories -

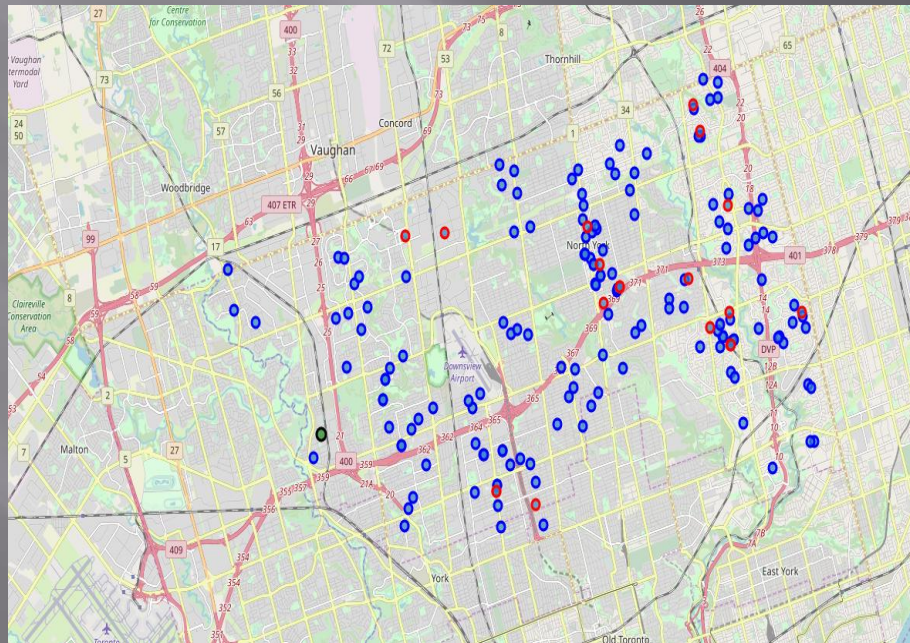


- ▣ Histogram for number of various category venues -



Methodology

- ▣ Map with Education institutions and Tennis facilities in north york neighborhoods:



Methodology

- Machine Learning – K means cluster analysis-
- Neighbourhood cluster with good facility



| Cluster Labels | | | Neighborhood | Above Highschool sum | Highschool and Below sum | Tennis Facility sum | AHS_dist mean | HSB_dist mean | TF_dist mean |
|----------------|---|---|------------------|-------------------------|-----------------------------|------------------------|------------------|------------------|-----------------|
| 21 | 2 | 2 | Willowdale South | 8 | 9 | 3 | 215.4 | 344.3 | 122.4 |

- Neighbourhood cluster with poor facility

| Cluster Labels | | | Neighborhood | Above Highschool sum | Highschool and Below sum | Tennis Facility sum | AHS_dist mean | HSB_dist mean | TF_dist mean |
|----------------|---|---|------------------------------------|-------------------------|-----------------------------|------------------------|------------------|------------------|-----------------|
| 7 | 4 | 4 | Downsview West | 1 | 5 | 0 | 114.333333 | 728.500000 | 0.0 |
| 8 | 4 | 4 | Downsview, North Park, Upwood Park | 0 | 5 | 0 | 0.000000 | 843.600000 | 0.0 |
| 9 | 4 | 4 | Emery, Humberlea | 0 | 1 | 0 | 0.000000 | 680.000000 | 0.0 |
| 11 | 4 | 4 | Flemingdon Park, Don Mills South | 0 | 2 | 0 | 0.000000 | 821.500000 | 0.0 |
| 14 | 4 | 4 | Humber Summit | 0 | 3 | 0 | 0.000000 | 727.666667 | 0.0 |

Results

- ▣ Among 24 neighborhoods, there are total 128 high school or below institutions and 28 above high school education institutions. Number of Tennis facility is only 19 for the total 151 education institutions. More schools and facilities are distributed between 300m and 700m from neighborhood centers. **Neighbourhood cluster with good facility**
- ▣ Cluster analysis shows cluster5 with 5 neighborhoods and cluster2 with 7 neighborhoods are the two sets of neighbors without Tennis facility

Discussion and Conclusion

- ▣ Selected 11 neighborhoods for considering to new Tennis facilities.
- ▣ Three new facilities recommended are Bedford Park, Lawrence Manor East (8 institutions benefitted), Downsvie Central and Downsvie Northwest (7 institutions benefitted).

References

- ▣ <https://www.coursera.org>
- ▣ [**https://en.wikipedia.org**](https://en.wikipedia.org)
- ▣ <https://www.google.com>
- ▣ <https://stackoverflow.com>
- ▣ <https://www.markdownguide.org>
- ▣ <https://matplotlib.org/index.html>
- ▣ <https://pandas.pydata.org/pandas-docs/stable/reference/index.html>