# Predicting optimal location for Multi-sport arena – Toronto, Canada

James Nevil

Feb 25, 2020

## 1. Introduction

In this project we will try to find an optimal location for an Indoor multi-sports arena. Specifically, this report will be targeted to stakeholders interested in opening a Recreational multi-sport arena in Toronto, Canada.

Since there is a population boom in Toronto and recreational play is usually preferred by people of all ages, we will try to detect locations that do not have many indoor sports arenas. We are also particularly interested in areas with less gyms in vicinity. We would also prefer locations as close to populous centers like Downtown Toronto as possible, if first two conditions are met.

## 2. Data

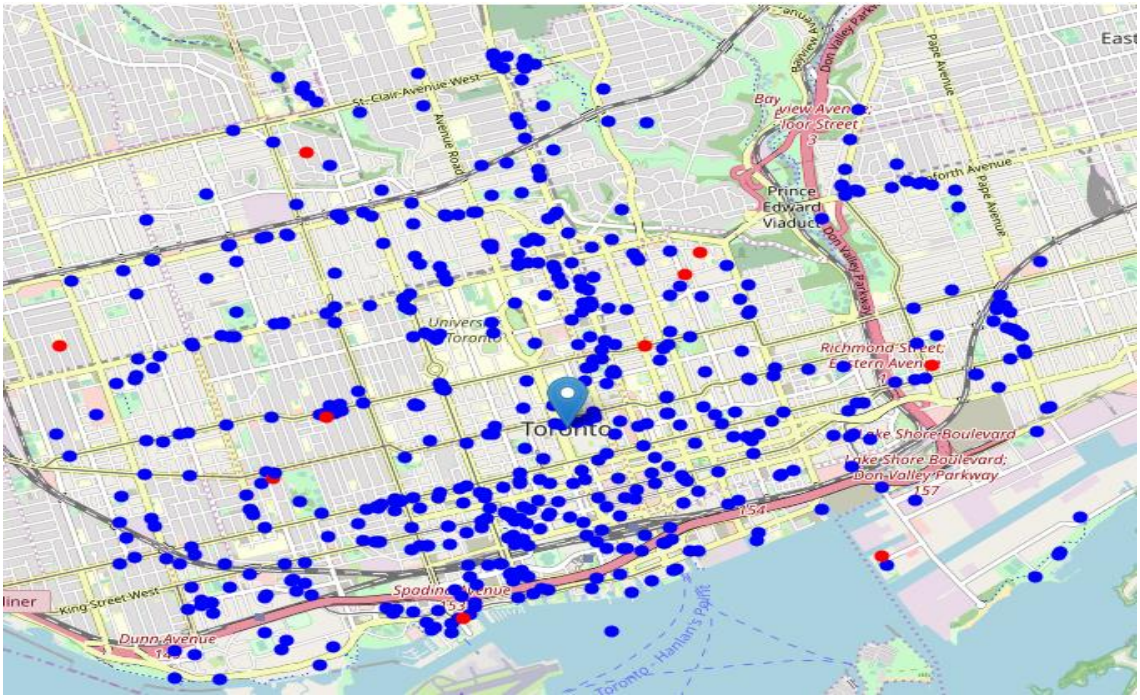
Based on definition of our problem, factors that will influence our decision are:

- number of existing sports arenas in the neighborhood (any type of arena).
- number of and distance to gyms in the neighborhood, if any.
- distance of neighborhood from city center.

Following data sources will be needed to extract/generate the required information:

- number of sports arena and their type/location in every neighborhood will be obtained using Foursquare API
- Coordinate of Toronto center will be obtained using Geopy geocoding of Toronto location (Downtown)

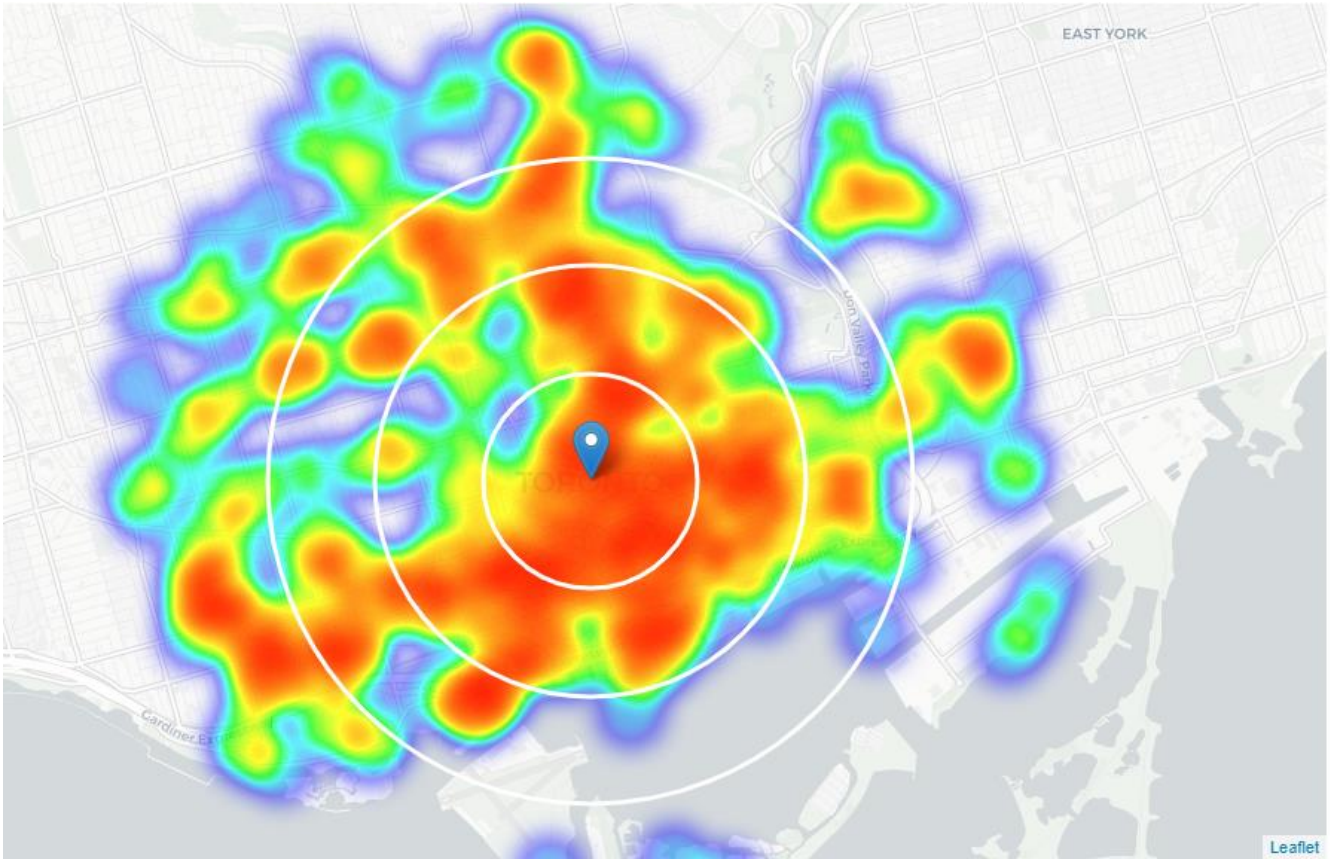
Below map displays the gathered data. Red dot refers to existing indoor sports arenas/courts while blue dot refers to the gyms and other sports arenas /fields in the location.



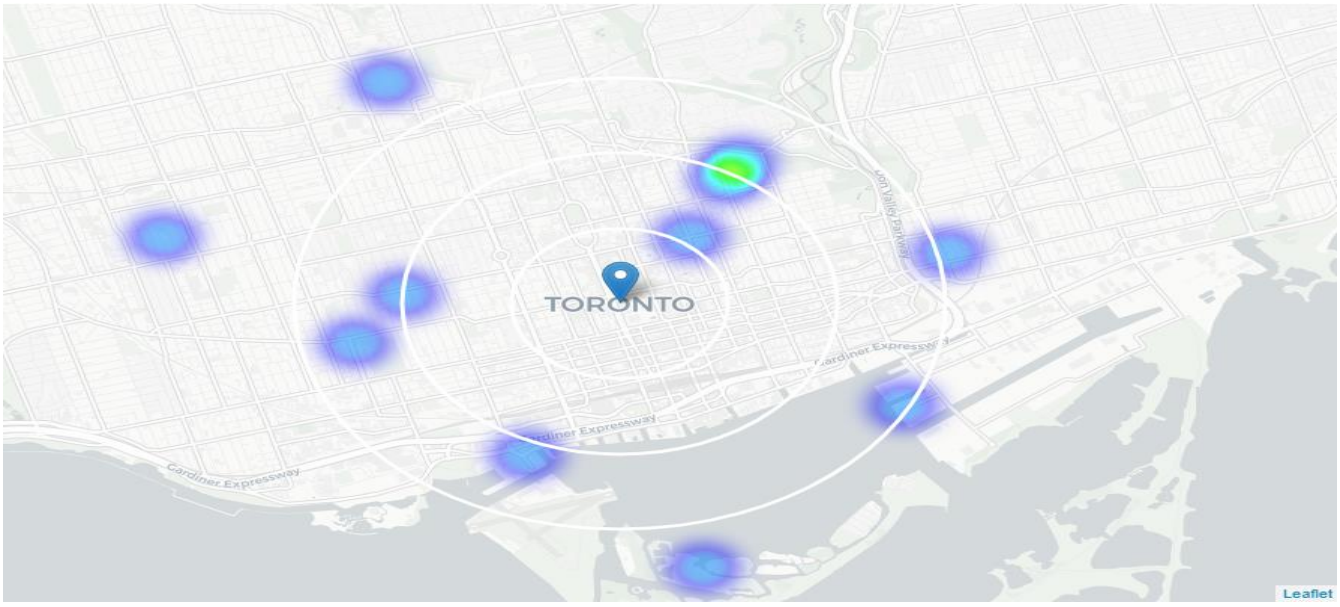
## 3. Methodology and Analysis

In this project we will direct our efforts on detecting areas of Toronto that have low arena/gym density, particularly those with low number of indoor multi-court arena. We will limit our analysis to area ~6km around city center.

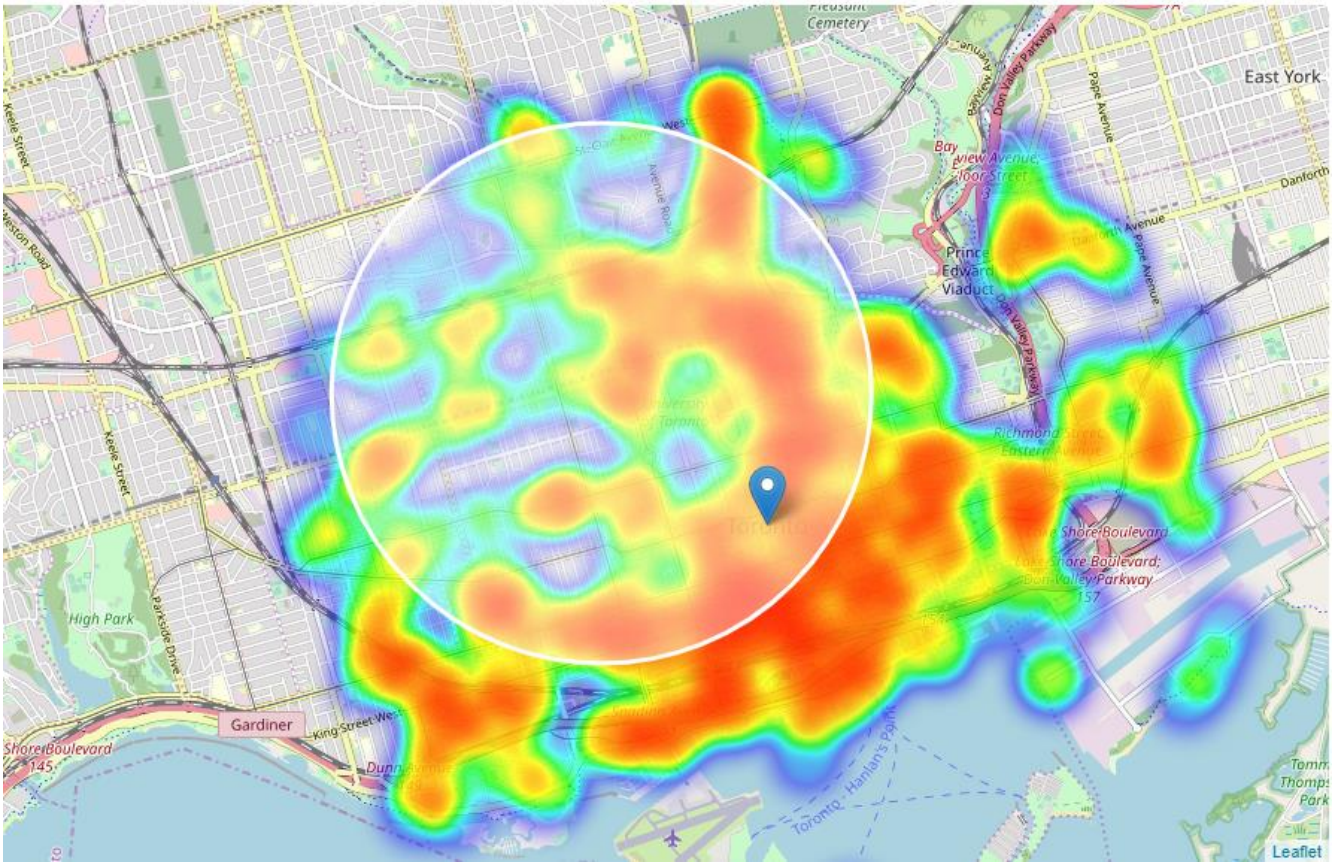
In first step we have collected the required data: location and type (category) of every arena within 6km from Toronto center (Downtown Toronto). We can visualize the same using the heatmaps as shown below




We have also identified indoor courts (according to Foursquare categorization).



Second step in our analysis will be calculation and exploration of 'arena density' across different areas of Toronto - we will use heatmaps to identify a few promising areas close to center with low number of arena/gyms in general (and no indoor sports arena in vicinity) and focus our attention on those areas. Based on this we will now focus our analysis on areas north, north-west from Downtown Toronto - we will move the center of our area of interest and reduce its size to have a radius of 2.5km. This places our location candidates mostly in boroughs West Toronto




In third and final step we will focus on most promising areas and within those create clusters of locations that meet some basic requirements established in discussion with stakeholders: we will take into consideration locations with no more than two gyms/arenas in radius of 500 meters, and we want locations without indoor arena in radius of 2000 meters. Based on the calculation below are the locations(geo-coordinates)

| | Latitude | Longitude | X | Y | Arenas nearby | Distance to Indoor arenas |
|---|---|---|---|---|---|---|
| 0 | 43.666416 | -79.384776 | -5.308459e+06 | 1.050741e+07 | 14 | 979.669451 |
| 1 | 43.667037 | -79.384871 | -5.308359e+06 | 1.050741e+07 | 14 | 1062.614203 |
| 2 | 43.662941 | -79.384997 | -5.309009e+06 | 1.050749e+07 | 14 | 701.246731 |
| 3 | 43.663562 | -79.385091 | -5.308909e+06 | 1.050749e+07 | 14 | 742.390523 |
| 4 | 43.664183 | -79.385186 | -5.308809e+06 | 1.050749e+07 | 13 | 794.065741 |
| 5 | 43.664804 | -79.385281 | -5.308709e+06 | 1.050749e+07 | 13 | 854.363572 |
| 6 | 43.665425 | -79.385375 | -5.308609e+06 | 1.050749e+07 | 13 | 921.593091 |
| 7 | 43.666046 | -79.385470 | -5.308509e+06 | 1.050749e+07 | 15 | 994.349303 |
| 8 | 43.666667 | -79.385565 | -5.308409e+06 | 1.050749e+07 | 15 | 1071.506999 |
| 9 | 43.667288 | -79.385659 | -5.308309e+06 | 1.050749e+07 | 18 | 1152.182260 |

On filtering these location to ones of our interest - no more than two gyms/arenas in radius of 500 meters, and we want locations without indoor arena in radius of 2000 meters we find we have 39 locations matching these conditions as shown in map below

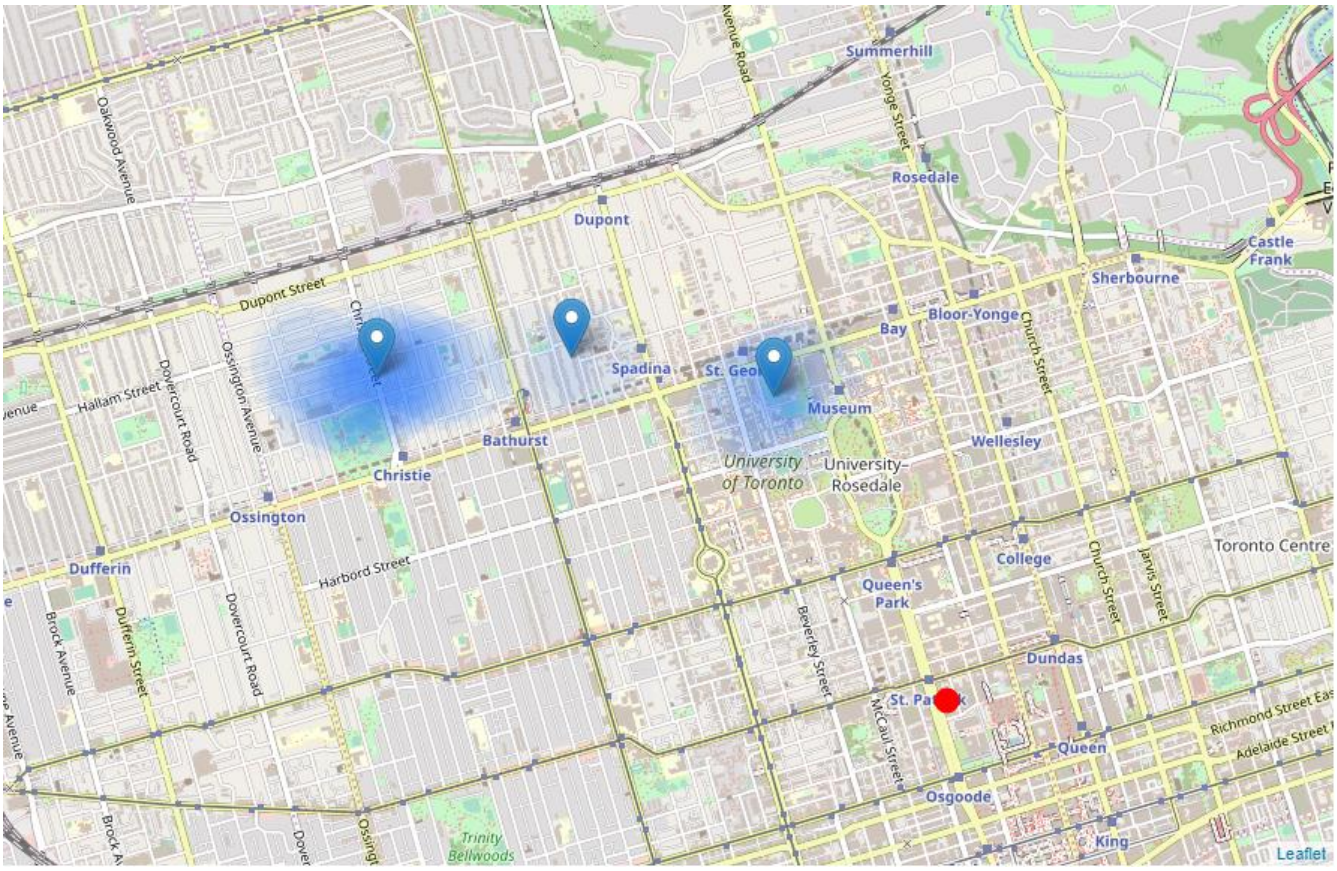

We will create clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

We can reverse geocode and obtain the location which can be presented to stakeholders.

```
============================================================
Addresses of centers of areas recommended for further analysis
============================================================

Varsity Centre, Devonshire Place, Bloor Street Culture Corridor, University–Rosedale, Old Toronto, Toronto, Golden Horseshoe, Ontario, M5S 1V6 => 2.3km from DownTown
151, Christie Street, Christie Pits, University–Rosedale, Old Toronto, Toronto, Golden Horseshoe, Ontario, M6G 3B9 => 4.4km from DownTown
366, Brunswick Avenue, The Annex, University–Rosedale, Old Toronto, Toronto, Golden Horseshoe, Ontario, M5S 1X5 => 3.4km from DownTown
```

So, to conclude our analysis we have created 3 addresses representing centers of zones containing locations with low number of gyms or arena and no indoor sports courts nearby, all zones being fairly close to city center (all less than 5km from Downtown).




## 4. Results and Discussion

Our analysis shows that although there is a great number of arenas or gyms in Toronto (~525 in our initial area of interest which was 6x6km around Downtown Toronto), there are pockets of low gym/arena density relatively close to city center. Highest concentration of gym/arena was detected south and east from Downtown Toronto, so we focused our attention to areas north, north-west and west, corresponding to boroughs West Toronto.

After directing our attention to this narrower area of interest we first created a dense grid of location candidates (spaced 100m apart); those locations were then filtered so that those with more than two gym/arena in radius of 500m and those with an Indoor sports arena closer than 2000m were removed.

Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 3 zones containing largest number of potential new sports arena locations based on number of and distance to existing venues. This, of course, does not imply that those zones are optimal locations for a new arena! Purpose of this analysis was to only provide info on areas close to Toronto center but not crowded with existing gyms or sports arena. It is entirely possible that there is a very good reason for small number of gyms or arenas in any of those areas, reasons which would make them unsuitable for a new sports arena regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition, but also other factors considered and all other relevant conditions met.

## 5. Conclusion

Purpose of this project was to identify Toronto areas close to center with low number of gyms and sports arenas in order to aid stakeholders in narrowing down the search for optimal location for a new multi-sport recreational sports court. By calculating gyms and sports arena density distribution from Foursquare data we have first identified general boroughs that justify further analysis (West Toronto), and then generated extensive collection of locations which satisfy some basic requirements. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like projected population boom, real estate development etc...