

# **Choosing the best area of Winnipeg to live with children**

**Sultanova Liliya**

**Ufa, 2020**

## **1. Introduction**

### **1.1 Background**

Immigration to another country or even moving to a neighboring city is always difficult especially for families with children. Every parent wants to provide their children with a safe and comfortable environment for their growth.

Moving to another city it is stress for everybody and especially for children, as the same as transition from one school to another. Therefore, parents need to think about advances in which area they want to buy a house and choose school their child will go to and rent housing in the same area for the first time, so that they do not change the district and school later.

Another reason to choose school one time and not to change it later is that some Winnipeg schools study some subjects combined. For example, 7-8 first learn a subject together for grade 7 then for 8. If student moves from one school to another, it may happen, that a student will study physics for grade 7 twice, but for the eighth he will not study never.

So, I analyzed the criminal situation in Winnipeg and chose the safest areas in City. Then I have explored how close the elementary and secondary schools in the selected areas are to each other so that the distance and time to go to school does not increase after during the transition from eighth to ninth grade. I also explored the most common places in the selected areas of city to get an idea of the activities that are possible in them.

I chose the Winnipeg (the capital and largest city of the province of Manitoba in Canada), because this city is most interesting to me as a city of most appropriate for immigration in the future and I am curious to investigate the features of its areas.

### **1.2 Problem**

In my project, I analyzed which parts of the city of Winnipeg are best suited for living with children, in terms of safety, proximity to schools, and availability of various activities.

### 1.3 Interest

This project is intended for families with children who want to move to live in Winnipeg. It can also be useful for realtors offering to buy a house or apartment for families with school-age children.

## 2. Data acquisition and cleaning

### 2.1 Data sources

The most difficult task was to find information about each school, such as the number of students and the average score by given of feedbacks. Since there was no General information, I had to open each school's website and search for data there. For this purpose I used websites of a school division in Winnipeg: <https://www.lrsd.net/School/Community-of-Schools/Pages/default.aspx> (the Louis Riel School Division) and <https://www.pembinatrails.ca/School/SchoolList/Pages/default.aspx> (Pembina Trails School Division). Feedbacks for each school I found on the site: <https://www.google.ru/maps/place/>, to be clear, I indicated the number of those who evaluated this school.

To analyze the criminal situation of each district, I used the site: <https://www.winnipeg.ca/police/AnnualReports/annualreports.stm>.

I took Census data for 1996, 2001, 2006 and 2011, 2016 by district: <https://www.winnipeg.ca/census/sitemap/sitemap.asp>.

I used City of Winnipeg Currently Active Official Property Addresses and Associated Coordinates to get the coordinates of each school: <https://data.winnipeg.ca/City-Planning/Addresses/cam2-ii3u>.

I used Foursquare API to get the most common venues of given Borough of Winnipeg: <https://developer.foursquare.com/>.

### 2.2 Data cleaning

In the first part, I looked at the criminal situation in areas of the city of Winnipeg, data from 2015 to 2018 were presented on the police website. After seeing that the number of residents in each district is different (for example, in St.Vital there are twice as many as in Assiniboine South), it became clear that they can't be compared. So, I calculated the number of crimes per 10,000 residents.

In the second part, I looked at schools in two school division. I hoped to take the data on the rating and reviews in the Foursquare API, unfortunately, this information about Winnipeg schools is not available on the given website. So, I found this data

in [google.ru/maps/place](https://www.google.ru/maps/place). Since there was no General information on the number of students in each school, I had been searched for them on the site of each school.

I used the Winnipeg address database to determine the geographical coordinates of schools. The address of the school on the website and in the database differed in format. I brought them to the same species so I could work with them. Louis\_Riel\_School\_Division there was no information on the General site about which classes are taught in each school in the area. I found this information on the website of each school.

In the third part, I used the clustering method to find schools with similar characteristics also I used the Forsquare API to analyze areas for possible sports and cultural activities. I took a list of all the objects found related to sports and culture.

### 3. Methodology and Results section

#### 3.1 Choosing the safest areas of Winnipeg.

Every parent wants to create a safe environment for their child, so let's choose the three safest areas in Winnipeg.

Unfortunately, data for 2019 is not available yet, and data earlier than 2015 in the reports are provided for the entire city, so the review period is from 2015 to 2018.

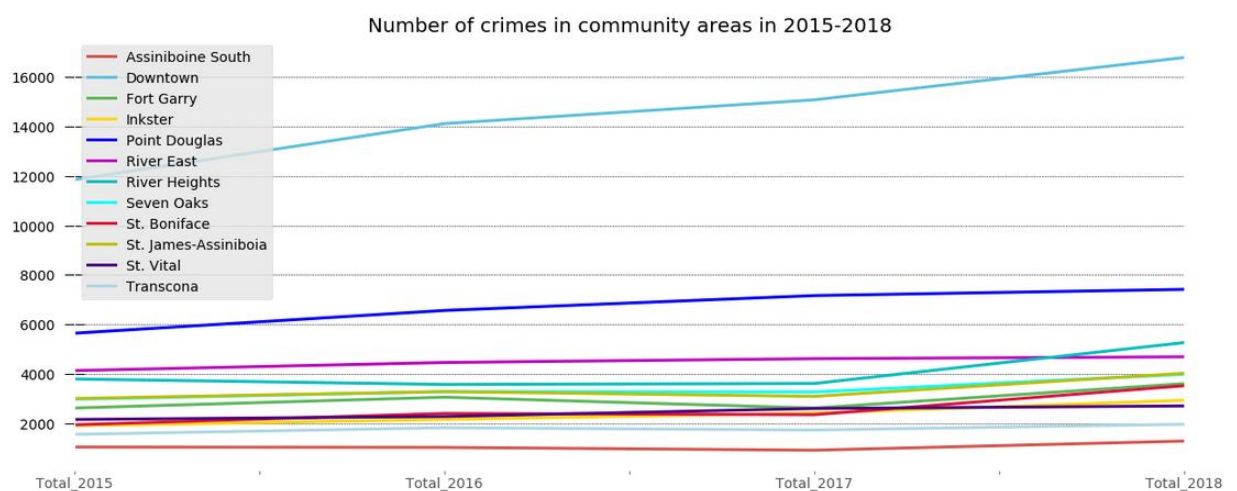


Figure 1. Number of crimes in community areas in 2015-2018

The graph shows that the most crimes were committed in the Downtown, the least crimes were committed in Assiniboine South. We also see an increase in the number of crimes by year in all areas of Winnipeg.

I looked at the number of criminal acts by type for each district of the city for 2018. The result is shown on the histogram.



Figure 2. Criminal activity in 2018

I took Census data for 1996, 2001, 2006 and 2011, 2016 by district and combined them into one table. For clarity, I have built a graph based on the data obtained.

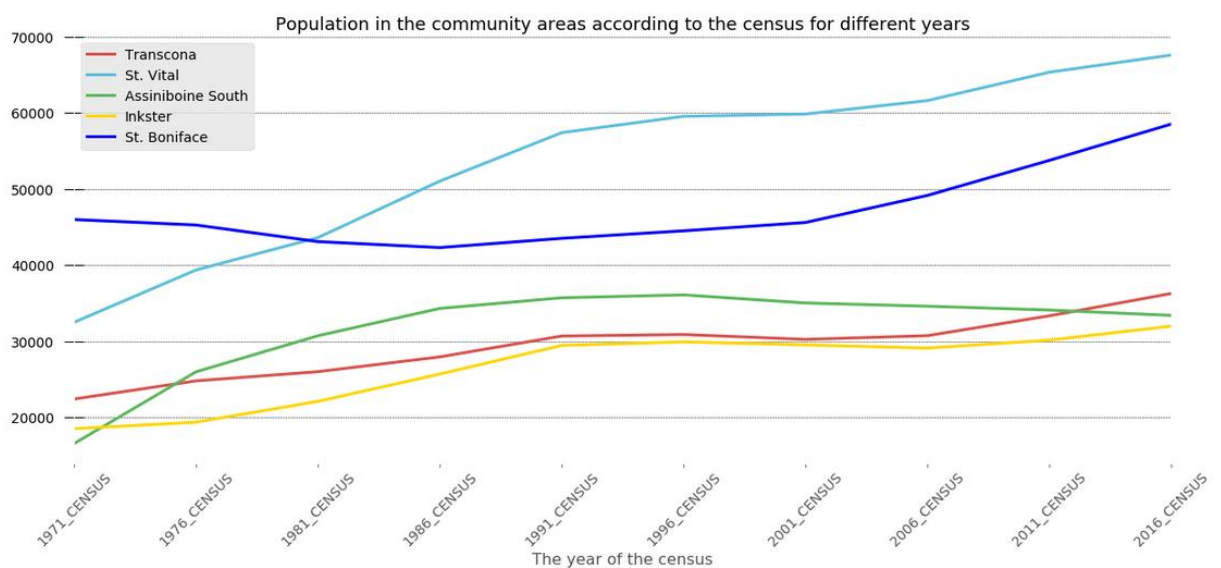


Figure 3. Population in the community areas according to the census for different years

From the graph we see that the St.Vital number of population is two times more than in Assiniboine South, therefore these data cannot be considered to compare.

Find the number of crimes committed per capita in each district. Since we know the total number of residents by district only for 2016, let's consider the total crime rate per 10000 people for 2016.

Formula for calculating the total crime rate per 10000 people:

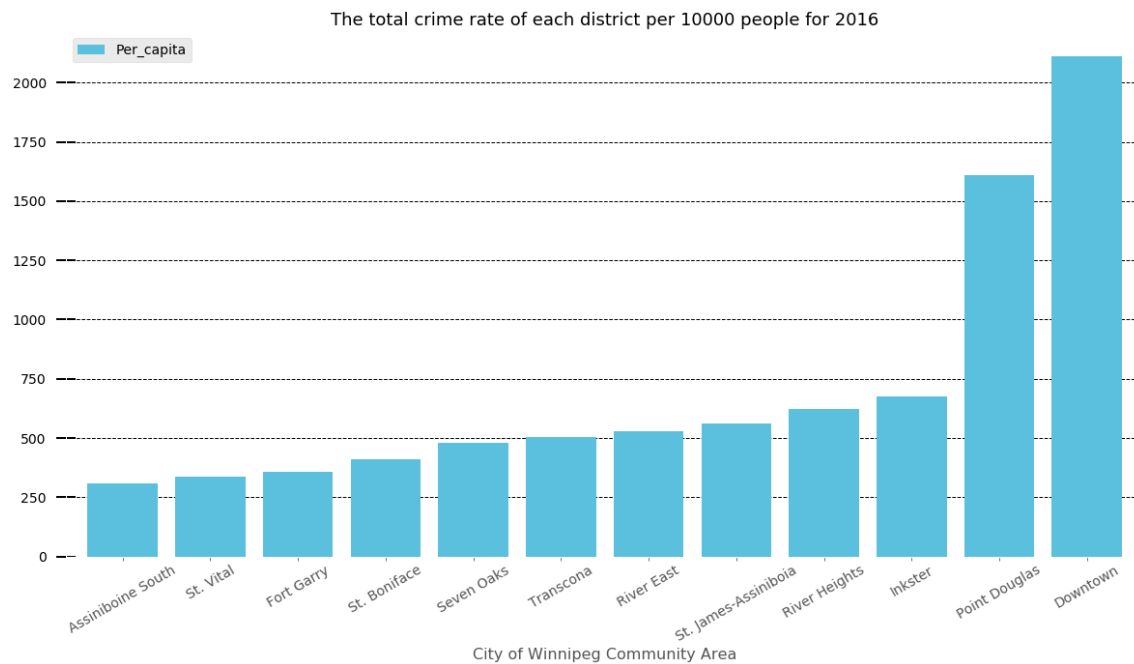
$$K = P * 10000 / N$$

where K is the crime rate

N - the number of facts or the number of persons who have committed crimes

P – population (total or at the appropriate age).

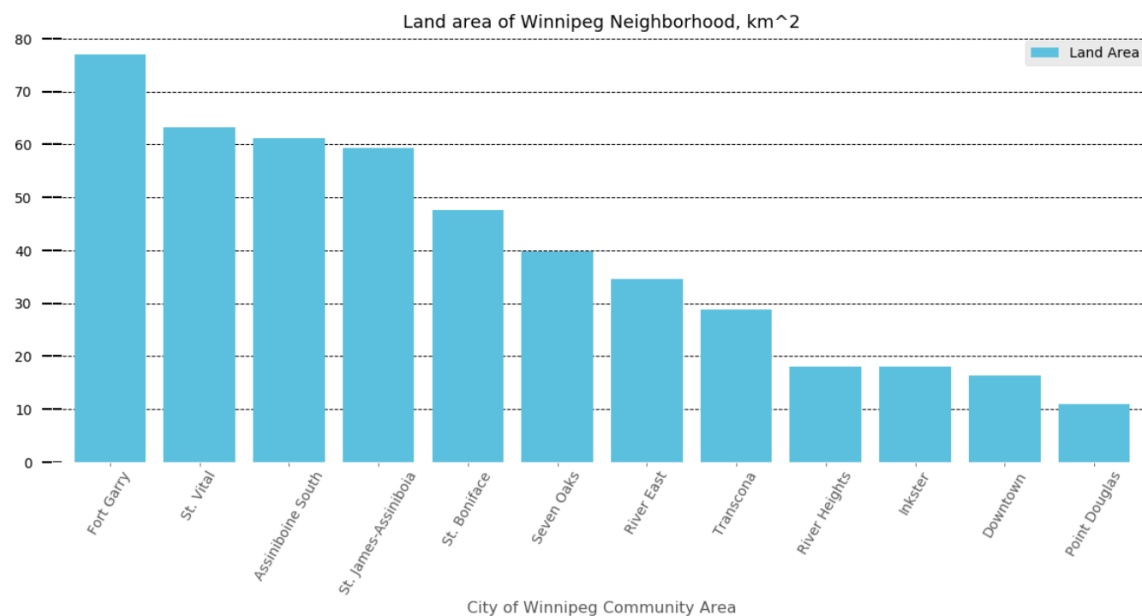
When the calculation is made for the entire population, the coefficient actually reflects only how the population is affected by crime.



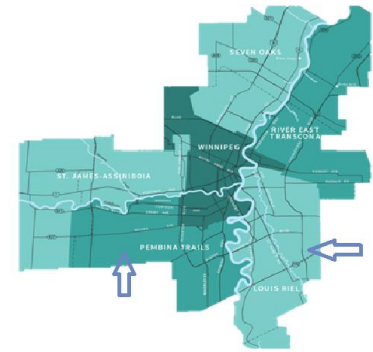
Therefore, we will select the following areas for further wards:

Assiniboine South, St. Vital, Fort Garry.

Note that these wards have the largest area:



St.Vital and St.Boniface are part of the **Louis Riel School Division**, the main part of the Ford Garry and Assiniboine South are part of the **Pembina Trails School Division**, so we will continue to review the schools of only these school divisions.



### 3.2 Determining the distance between schools.

I mapped all Winnipeg schools that belong to the Louis Riel and Pembina Trails School Division. The map shows the number of students in each school (I found the radius of the circle to indicate each school by dividing the number of students by 100 and rounding up to an integer in the larger direction). For example, if there are 51 students in the school, then it will be marked on the map with a circle of radius 1, if the number of students is 1062, then the radius is 11.

The color of the circle indicates which classes are in school, if the school has:

○ 9 - 12 grade

○ 4 - 8 grade

○ a kindergarten and primary

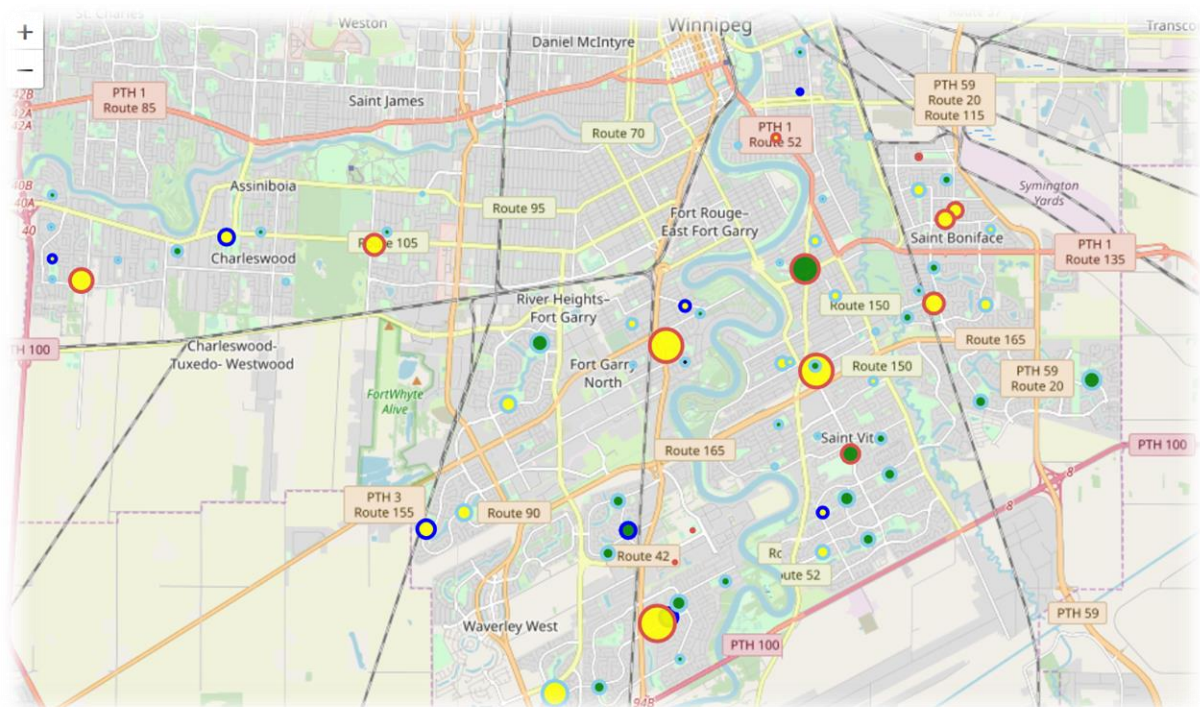
Filling the circle indicates what reviews about the school:

● if  $0 < \text{mark} < 2$

● if  $2 \leq \text{mark} < 4$

● if  $\text{mark} \geq 4$

● if no reviews





I made a request for grades 9-12 and saw there are only 11 schools. In order not to change the place of residence when a child grows up (especially if there are several children), it is necessary to look for housing closer to these schools:

	Program_English	Program_French	Latitude	Longitude	Enrollment	Average_mark	count_review	SK	Kindergarten	K_8	H_9_12	Count_!
count	11.000000	11.000000	11.000000	11.000000	11.000000	11.000000	11.000000	11.0	11.000000	11.0	11.0	
mean	0.818182	0.363636	49.846108	-97.136738	737.363636	3.336364	18.727273	1.0	0.090909	0.0	1.0	
std	0.404520	0.504525	0.020889	0.072938	362.751781	1.149150	10.257591	0.0	0.301511	0.0	0.0	
min	0.000000	0.000000	49.801995	-97.309285	55.000000	0.000000	0.000000	1.0	0.000000	0.0	1.0	
25%	1.000000	0.000000	49.837529	-97.149216	595.000000	3.450000	13.500000	1.0	0.000000	0.0	1.0	
50%	1.000000	0.000000	49.850943	-97.112557	754.000000	3.600000	19.000000	1.0	0.000000	0.0	1.0	
75%	1.000000	1.000000	49.859447	-97.088967	967.500000	3.750000	23.500000	1.0	0.000000	0.0	1.0	
max	1.000000	1.000000	49.875913	-97.071911	1235.000000	4.300000	37.000000	1.0	1.000000	0.0	1.0	

I found the distance between the primary and secondary schools, for this purpose I translated the coordinates into a spherical system:

$$\cos(d) = \sin(\varphi_A) \cdot \sin(\varphi_B) + \cos(\varphi_A) \cdot \cos(\varphi_B) \cdot \cos(\lambda_A - \lambda_B)$$

where:

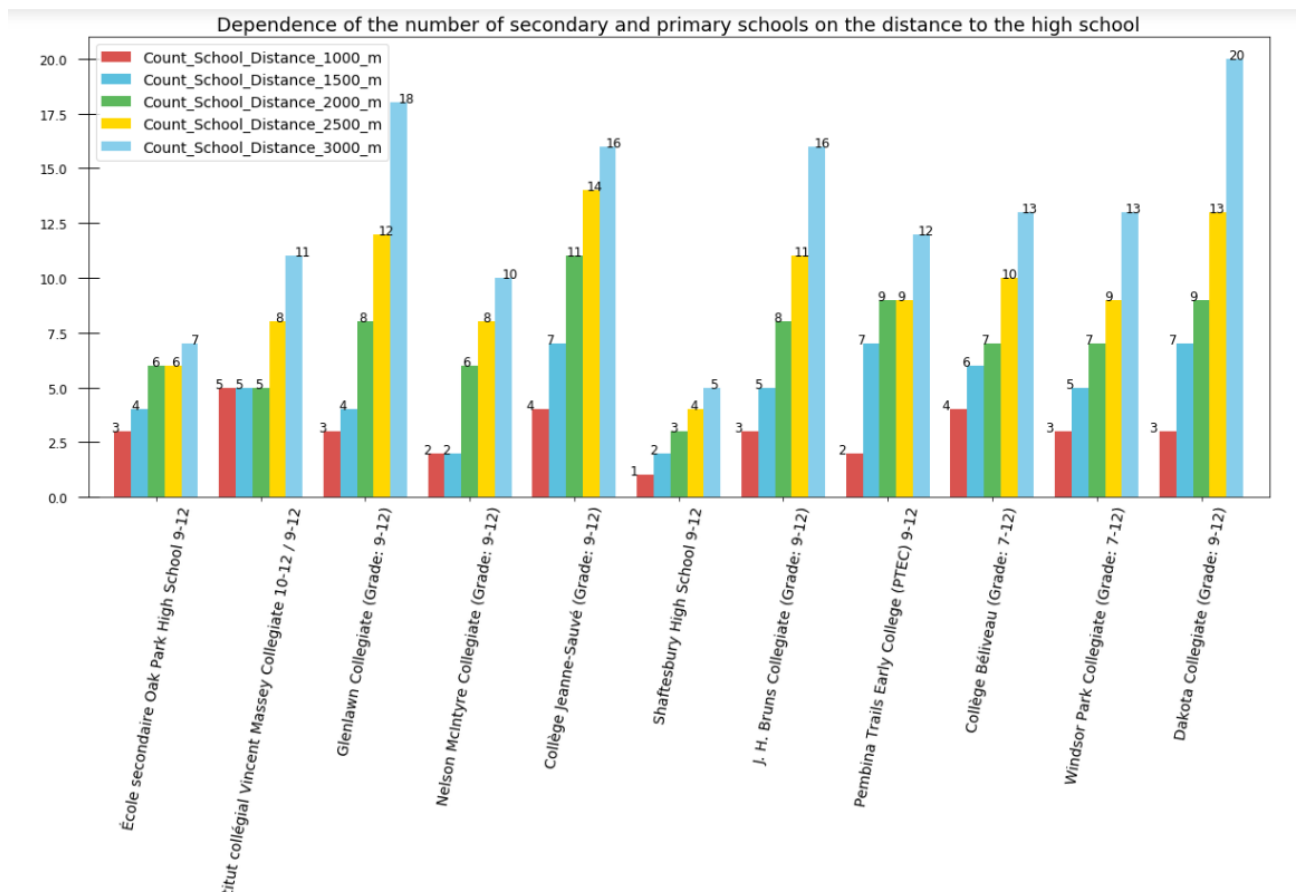
$\varphi_A, \varphi_B$  are the latitudes

$\lambda_A, \lambda_B$  are the longitudes

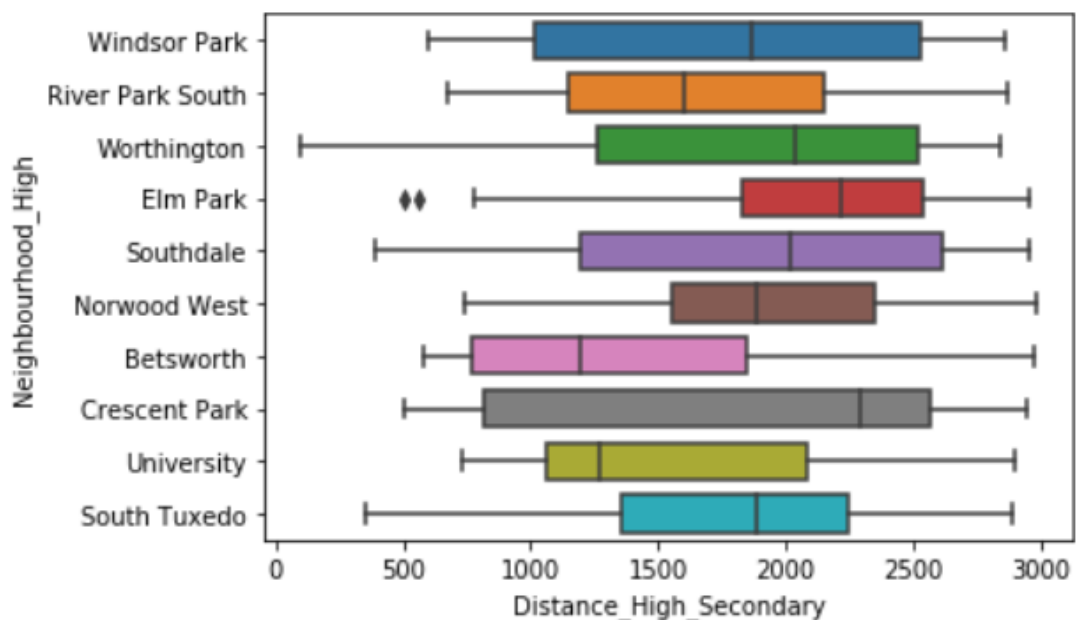
d is the distance between points measured in radians by the length of the arc of a large circle of the globe

$L = R \cdot d$ , took the earth's radius  $R = 6378137m$ .

	Description_High	Description_Secondary	Distance_High_Secondary	Enrollment_High	Enrollment_Secondary	Average_mark_High	Average_mark_Secondary
9	Dakota Collegiate (Grade: 9-12)	Victor Mager School (Grade: K-8)	101	1235	375	3.4	4.8
32	Shaftesbury High School 9-12	Laidlaw School K-8	349	754	242	3.8	4.1
15	J. H. Bruns Collegiate (Grade: 9-12)	Niakwa Place School (Grade: K-8)	395	800	253	3.6	4.1
25	Institut collégial Vincent Massey Collegiate ...	École Crane K-4	501	1209	259	3.5	0.0
12	Glenlawn Collegiate (Grade: 9-12)	Windsor School (Grade: K-8)	509	1059	167	4.3	4.3



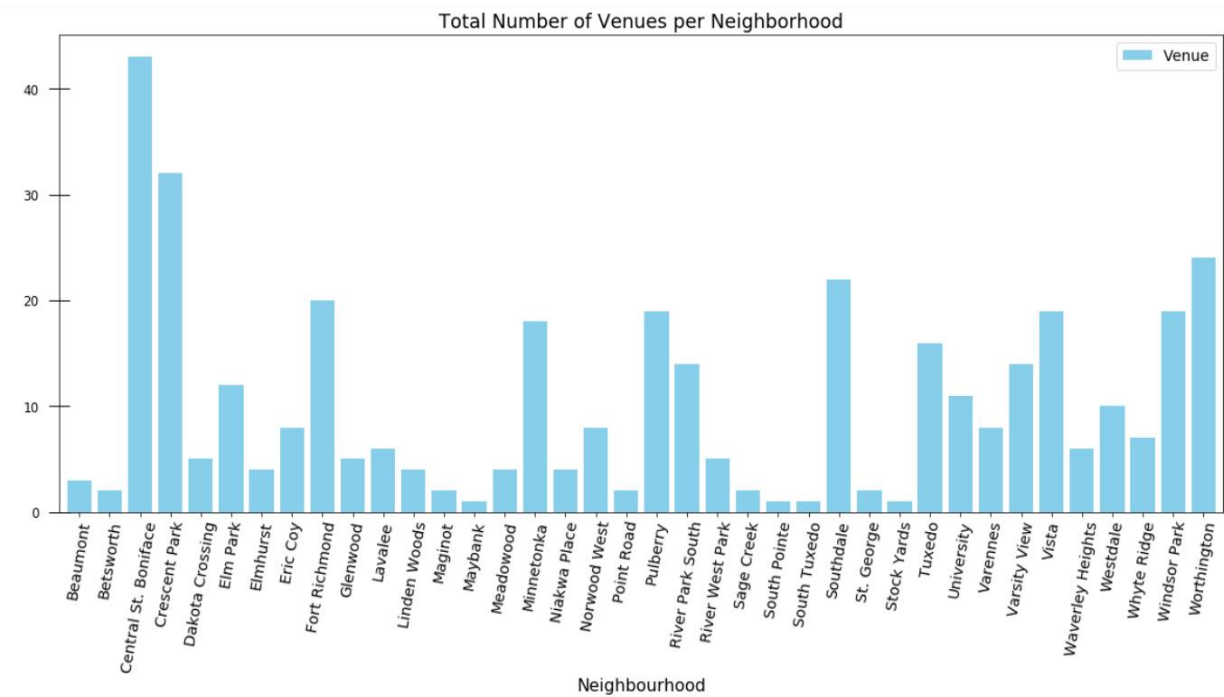
In this plot show the distribution of distance High and Secondary school in a way that facilitates comparisons between neighborhoods.





### 3.3 Clustering method to find schools with similar characteristics.

I use Foursquare API to generate a list of venues per borough:



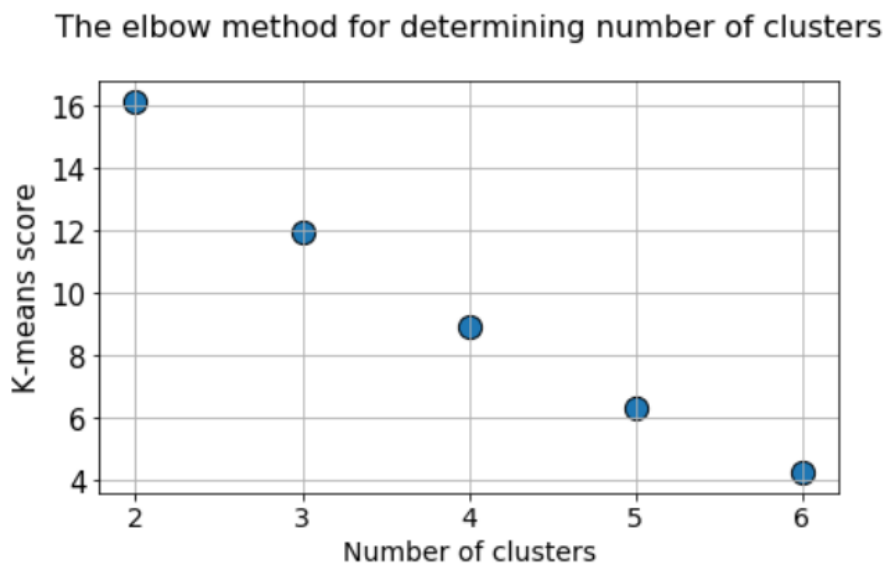
I looked at what places there are in each area for sports:

Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Beaumont	Park	Event Space	Baseball Field	Athletics & Sports	Stadium
Betsworth	Park	Event Space	Baseball Field	Athletics & Sports	Stadium
Central St. Boniface	Gym / Fitness Center	Bowling Alley	Park	Event Space	Baseball Field
Crescent Park	Athletics & Sports	Park	Event Space	Baseball Field	Stadium
Dakota Crossing	Park	Event Space	Baseball Field	Athletics & Sports	Stadium

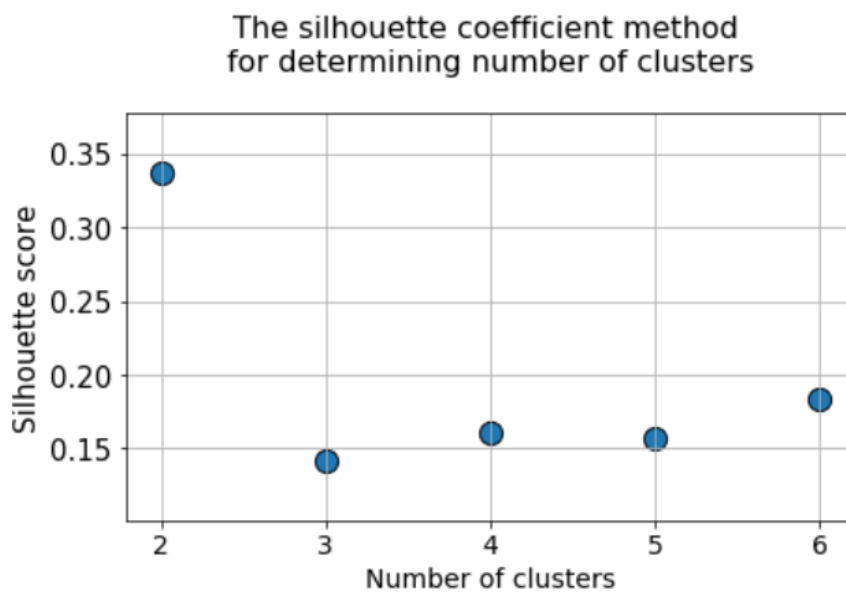
I used the clustering method to find schools with similar characteristics.

I selected the following data to use for clustering: Enrollment of school, program English or French, average mark, the number of secondary schools is closer than 1000 meters, area where there is soccer field or stadium, baseball field, bowling alley, dance studio, sports club, football stadium, gym, fitness center, event space, park.

Determine the optimal number of clusters using the elbow method:



You can't determine how many clusters to take from the elbow method, so we use the silhouette coefficient method and Davies-Bouldin score:



The range of Silhouette score is  $[-1, 1]$ . Its analysis is as follows –  
+1 Score – Near +1 Silhouette score indicates that the sample is far away from its neighboring cluster.

0 Score – 0 Silhouette score indicates that the sample is on or very close to the decision boundary separating two neighboring clusters.

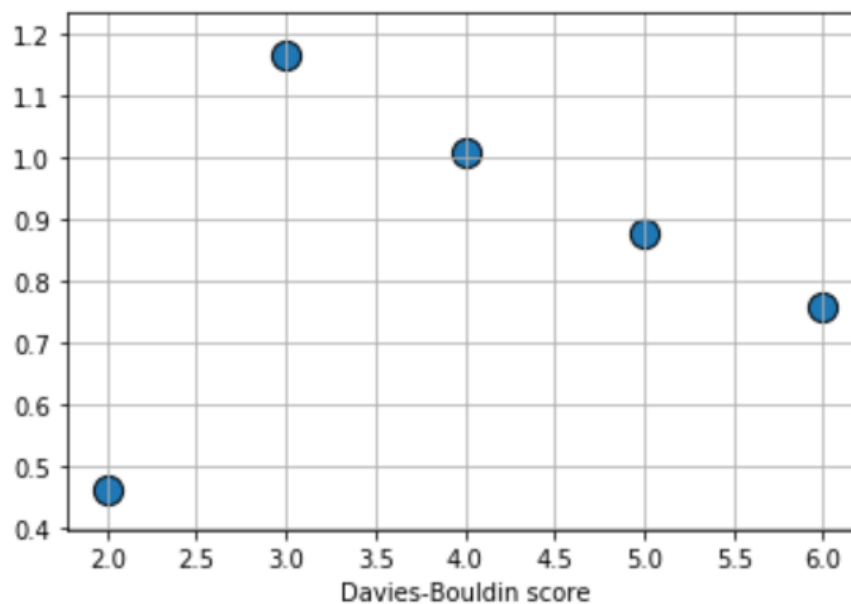
-1 Score – 1 Silhouette score indicates that the samples have been assigned to the wrong clusters.

The calculation of Silhouette score can be done by using the following formula:

$$\text{silhouettescore} = (p - q) / \max(p, q)$$

Here,  $p$  = mean distance to the points in the nearest cluster

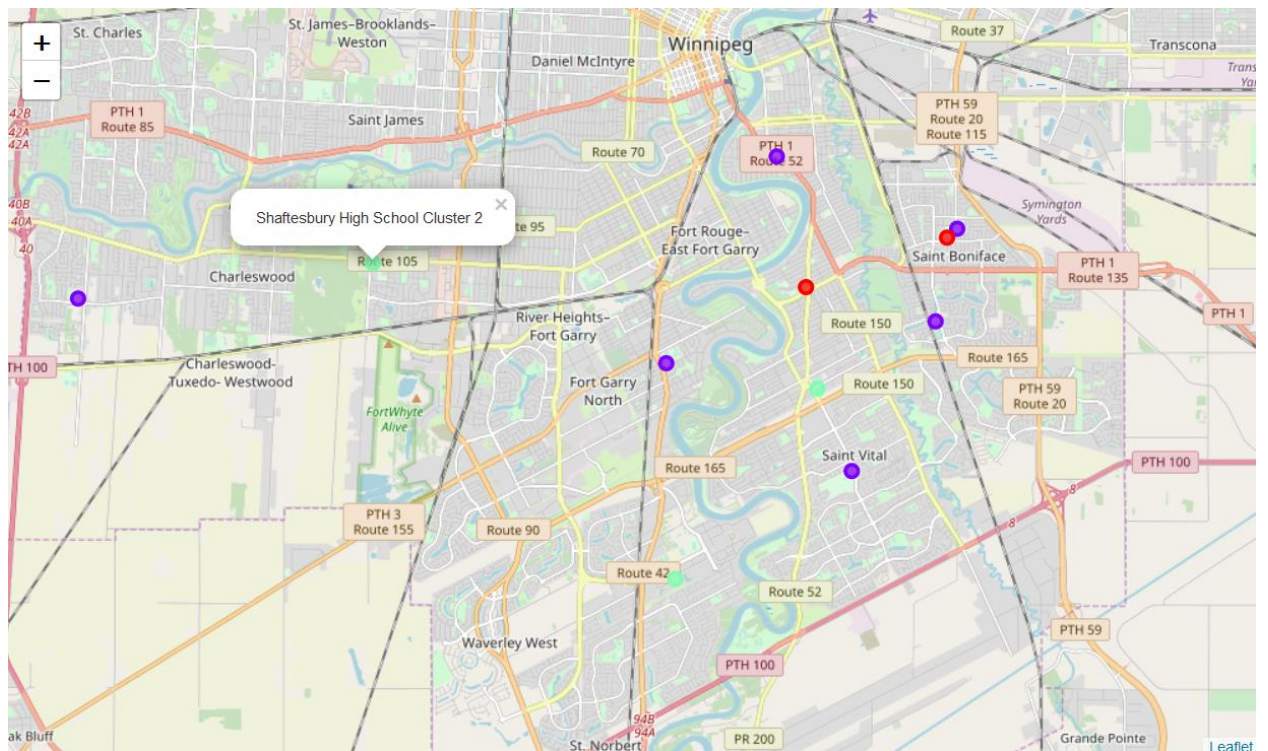
And,  $q$  = mean intra-cluster distance to all the points.



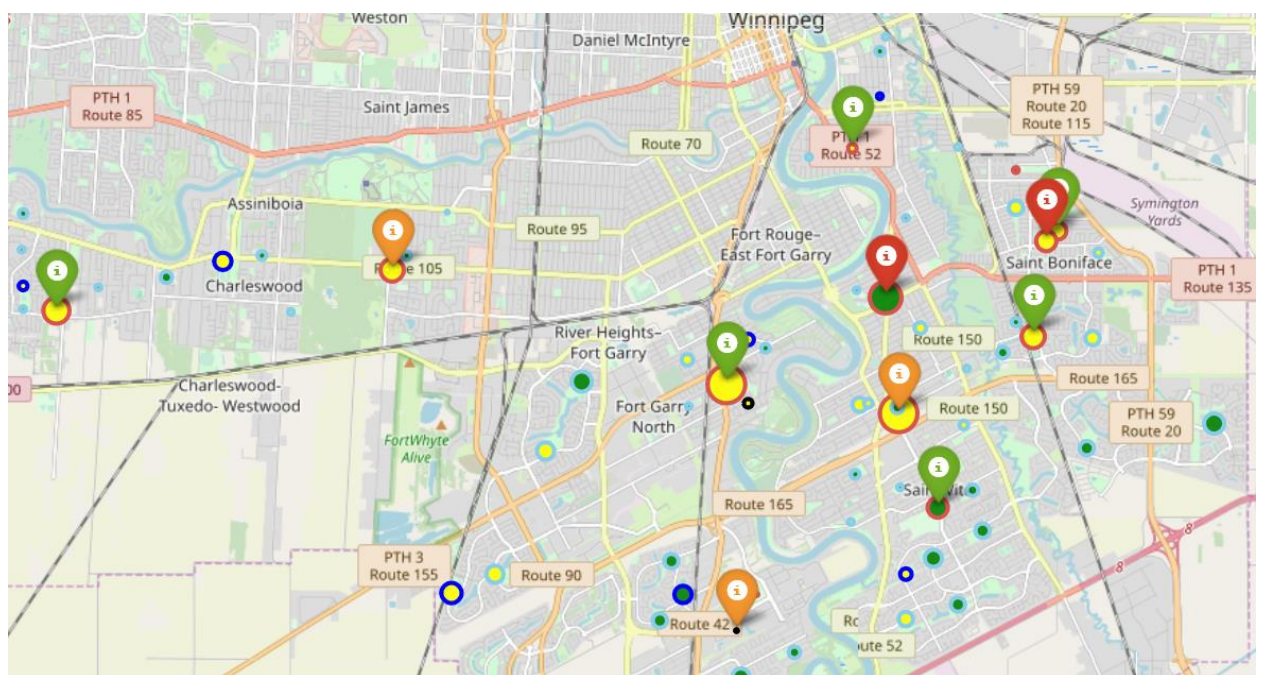
This is an internal evaluation scheme, where the validation of how well the clustering has been done is made using quantities and features inherent to the dataset. This has a drawback that a good value reported by this method does not imply the best information retrieval. Due to the way it is defined, as a function of the ratio of the within cluster scatter, to the between cluster separation, a lower value will mean that the clustering is better. It happens to be the average similarity between each cluster and its most similar one, averaged over all the clusters, where the similarity is defined as  $S_i$  above. This affirms the idea that no cluster has to be similar to another, and hence the best clustering scheme essentially minimizes the Davies–Bouldin index. ( [https://en.wikipedia.org/wiki/Davies%E2%80%93Bouldin\\_index](https://en.wikipedia.org/wiki/Davies%E2%80%93Bouldin_index) )

Therefore, the most appropriate number of clusters is 3.

The data obtained is shown on the map of Winnipeg:



On this map is shown areas where you can rent or buy housing, depending on the requirements for the school (for example, what language classes are taught or whether there is a Park nearby), to make clusters visible they are drawn as icons:



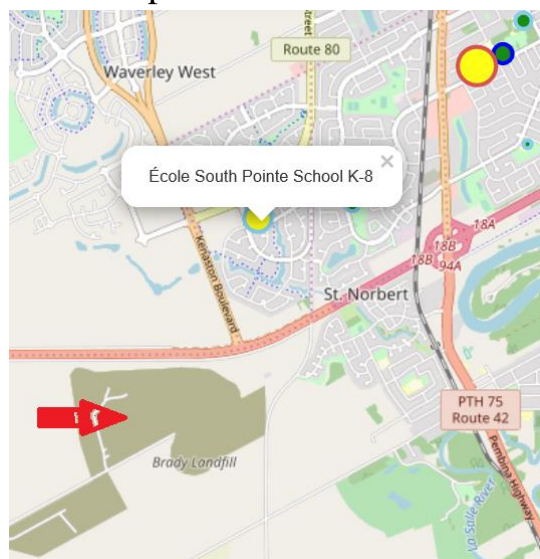


#### 4. Discussion section.

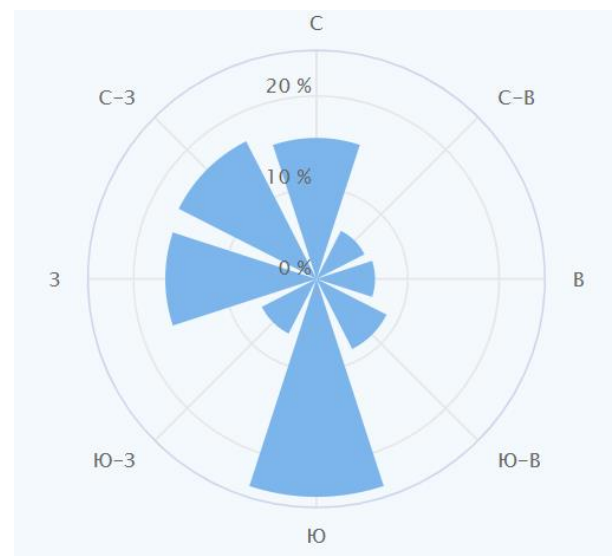
When choosing a place of residence, we must also pay attention to parameters that cannot be measured.

For example:

1. Proximity of the house to the railway road and railway crossing (for example, houses located near J. H. Bruns Collegiate).
2. The house is close to a busy motorway (Institut collégial Vincent Massey Collegiate).
3. Is there a threat of flooding in the neighborhood (for example, many buildings in Elm Park (the neighborhood where the Glenlawn Collegiate is located) were damaged by the 1950 Red River Flood. The neighborhood was one of the worst hit areas in the city.)
4. According to the wind rose, it can be noted that 23.8% of the wind blows from the South, where the Brady Road Landfill is located, so the South Pointe area and the school (École South Pointe) located in it can feel the smell, which is not pleasant.



<https://www.winnipeg.ca/waterandwaste/garbage/bradyRoad.stm>



<https://world-weather.info/archive/canada/winnipeg/>

5. The number of mosquitoes increases closer to the Northern areas of the city.
6. Housing and rent prices (École secondaire Oak Park High School and Shaftesbury High School located in areas with very high rents and home values).
7. It is also necessary to consider who will be neighbors, whether there are noisy places nearby (for example, night clubs), windows look at the yard or on the road, and many other factors that affect a comfortable stay.

## **Conclusions**

In this study, I analyzed the Winnipeg districts in terms of school locations. I hope that it will be useful for people who decide to live in Winnipeg. Each area has its own advantages and disadvantages and you need to know about them before moving to a new place. On the basis of this information, decide whether this place is suitable for you and your family, whether you are ready to put up with these shortcomings every day or not.