



Photo : Colourbox / #47677819

The Battle of Neighbourhoods Montreal

Capstone Project
Applied Data Science
Seif El Kilany

Introduction

- ▶ Montreal is the second most populated city in Canada.
- ▶ Young families tend to leave the downtown for more family oriented neighbourhoods.
- ▶ Work from home due to the COVID pandemic is speeding up the exodus.



Question

- ▶ The question that young families and real estate agents alike are always trying to address is :

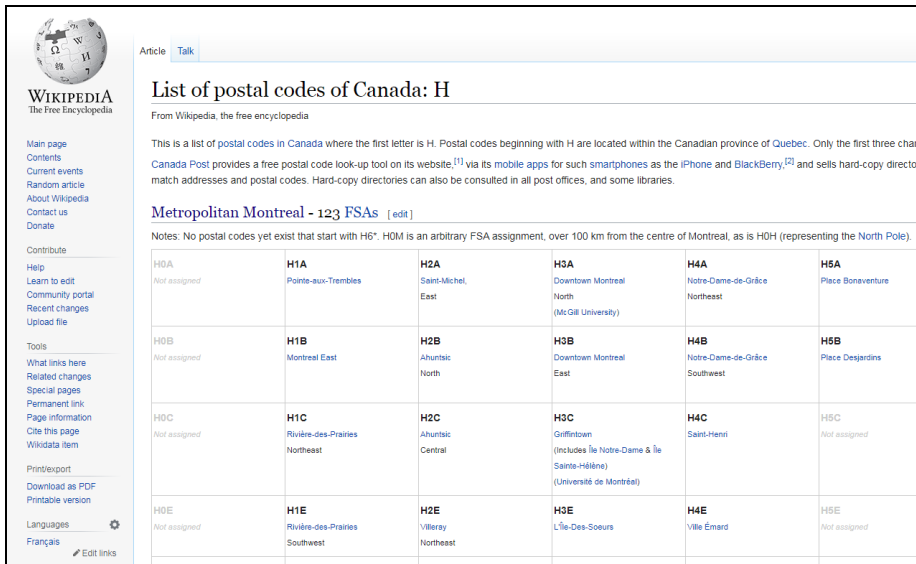
Where should they be moving to?



Data

Montreal Postal Codes

- ▶ We need geographical data of the city to identify each neighbourhood, for which postal codes is key.
- ▶ A list of the city's postal codes and neighbourhoods is available on Wikipedia.



Article Talk

List of postal codes of Canada: H

From Wikipedia, the free encyclopedia

This is a list of **postal codes in Canada** where the first letter is H. Postal codes beginning with H are located within the Canadian province of Quebec. Only the first three characters of the postal code are used to identify the neighbourhood. Canada Post provides a free postal code look-up tool on its website,^[1] via its **mobile apps** for such **smartphones** as the **iPhone** and **BlackBerry**,^[2] and sells hard-copy directories that match addresses and postal codes. Hard-copy directories can also be consulted in all post offices, and some libraries.

Metropolitan Montreal - 123 FSAs [[edit](#)]

Notes: No postal codes yet exist that start with H6*. H0M is an arbitrary FSA assignment, over 100 km from the centre of Montreal, as is H0H (representing the North Pole).

H0A <i>Not assigned</i>	H1A Pointe-aux-Trembles	H2A Saint-Michel, East	H3A Downtown Montreal North (McGill University)	H4A Notre-Dame-de-Grâce Northeast	H5A Place Bonaventure
H0B <i>Not assigned</i>	H1B Montreal East	H2B Ahuntsic North	H3B Downtown Montreal East	H4B Notre-Dame-de-Grâce Southwest	H5B Place Desjardins
H0C <i>Not assigned</i>	H1C Rivière-des-Prairies Northeast	H2C Ahuntsic Central	H3C Griffintown (Includes Île Notre-Dame & Île Sainte-Hélène) (Université de Montréal)	H4C Saint-Henri	H5C <i>Not assigned</i>
H0E <i>Not assigned</i>	H1E Rivière-des-Prairies Southwest	H2E Villeray Northeast	H3E L'Île-Des-Sœurs	H4E Ville Émard	H5E <i>Not assigned</i>

Website before scrapping

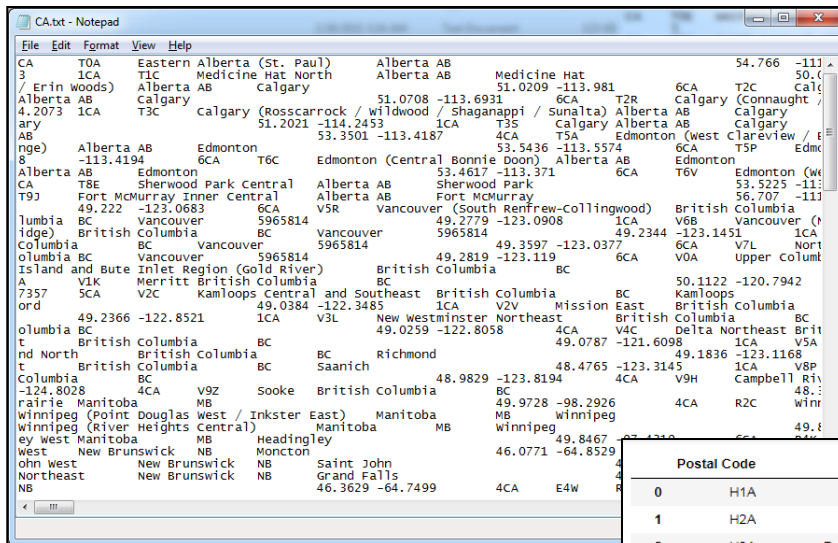
Dataframe after wrangling

	Postal Code	Neighbourhood
1	H1A	Pointe-aux-Trembles
2	H2A	Saint-Michel, East
3	H3A	Downtown Montreal North (McGill University)
4	H4A	Notre-Dame-de-Grâce Northeast
5	H5A	Place Bonaventure
6	H7A	Duvernay-Est
8	H9A	Dollard-des-Ormeaux Northwest
10	H1B	Montreal East
11	H2B	Ahuntsic North
12	H3B	Downtown Montreal East
13	H4B	Notre-Dame-de-Grâce Southwest

Data

Montreal Coordinates

- ▶ To be able to plot the neighbourhoods on a map we need their coordinates.
- ▶ We can get a list of all Canadian postal codes and their coordinates at GeoNames.



```
CA.txt - Notepad
File Edit Format View Help
CA TOA Eastern Alberta (St. Paul) Alberta AB 54.766 -113.111
3 ICA TIC Medicine Hat North Alberta AB 51.0209 -113.981
/ Erin Woods) Alberta AB Calgary 51.0708 -113.6931
Alberta AB Calgary 51.0708 -113.6931
4.2073 ICA T3C Calgary (Rosscarrock / wildwood / shaganappi / sunalta) Alberta AB
ary AB 51.2021 -114.2453 ICA T3S Calgary Alberta AB
AB 53.3501 -113.4187 ICA T5A Edmonton (west) Calgary / t
nge) Alberta AB Edmonton 53.5436 -113.5574 GCA TSP Edm
8 -113.4194 GCA T6C Edmonton (Central Bonnie Doon) Alberta AB Edmonton (W
Alberta AB Edmonton 53.4617 -113.371 GCA T6V Edmonton (W
CA T8E Sherwood Park Central Alberta AB Sherwood Park 53.225 -113
T9J Fort McMurray Inner Central Alberta AB Fort McMurray 56.707 -111
49.222 -123.0683 GCA V5R Vancouver (South Renfrew-Collingwood) British Columbia
lumbia BC Vancouver 5965814 49.2779 -123.0908 ICA V68 Vancouver (N
idge) British Columbia BC Vancouver 5965814 49.2344 -123.1451 ICA
olumbia BC Vancouver 5965814 49.3597 -123.0377 GCA V7L Nort
olumbia BC Vancouver 5965814 49.2819 -123.119 GCA V0A Upper Colum
Island and But Inlet Region (Gold River) British Columbia BC 50.1122 -120.7942
A V1K Merritt British Columbia BC Kamloops 50.1122 -120.7942
7357 SCA V2C Kamloops central and Southeast British Columbia BC
ord 49.2366 -122.8521 ICA V3L New Westminster Northeast British Columbia BC
olumbia BC British Columbia BC BC Saanich 48.9829 -123.8194
nd North British Columbia BC BC Saanich 48.9829 -123.8194
t Columbia BC BC Saanich 48.9829 -123.8194
-124.8028 ICA V9Z Sooke British Columbia BC 49.9728 -98.2926
rairie Manitoba MB Winnipeg (Point Douglas west / Inkster East) Manitoba MB
winnipeg (River Heights Central) Manitoba MB Winnipeg (River Heights Central)
ey West Manitoba MB Winnipeg (River Heights Central) Manitoba MB
west New Brunswick NB Moncton 46.0771 -64.8529
ohn west New Brunswick NB Saint John 46.3629 -64.7499
Northeast New Brunswick NB Grand Falls 46.3629 -64.7499
NB 46.3629 -64.7499 ICA E4W
```

Original CSV file

Dataframe after reading file and wrangling

	Postal Code	Neighbourhood	Place Name	Province	Province Code	Latitude	Longitude
0	H1A	Pointe-aux-Trembles	Pointe-Aux-Trembles	Quebec	QC	45.6753	-73.5016
1	H2A	Saint-Michel,East	Saint-Michel East	Quebec	QC	45.5618	-73.5990
2	H3A	Downtown Montreal North(McGill University)	Downtown Montreal North	Quebec	QC	45.5040	-73.5747
3	H4A	Notre-Dame-de-GrâceNortheast	Notre-Dame-de-Grâce Northeast	Quebec	QC	45.4717	-73.6149
4	H5A	Place Bonaventure	Place Bonaventure	Quebec	QC	45.4992	-73.5646
5	H7A	Duvernay-Est	Duvernay-Est	Quebec	QC	45.6739	-73.5924
6	H9A	Dollard-des-OrmeauxNorthwest	Dollard-Des-Ormeaux Northwest	Quebec	QC	45.4948	-73.8317
7	H1B	Montreal East	Montreal East	Quebec	QC	45.6320	-73.5075

Data

Montreal Venues

- ▶ We need to explore each neighbourhood for points of interest to help in the decision making.
- ▶ Foursquare API is used to pool such vast data.

FOURSQUARE /developers

Blog Docs

Documentation Places API

Search Documentation

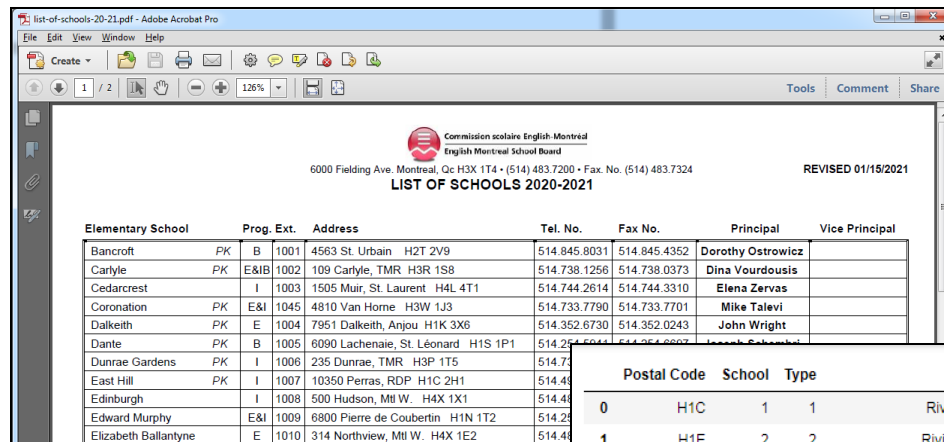
	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Pointe-aux-Trembles	45.6753	-73.5016	Parc-nature de la Pointe-aux-Prairies	45.678834	-73.501162	Park
1	Pointe-aux-Trembles	45.6753	-73.5016	AMT Gare Pointe-aux-Trembles	45.674882	-73.504908	Train Station
2	Pointe-aux-Trembles	45.6753	-73.5016	Parc Yves-Thériault	45.678675	-73.502037	Park
3	Saint-Michel,East	45.5618	-73.5990	Bar Zoe	45.559673	-73.597542	Karaoke Bar
4	Saint-Michel,East	45.5618	-73.5990	Marché Aux Pucés Saint-Michel	45.562502	-73.605079	Flea Market
5	Saint-Michel,East	45.5618	-73.5990	STM Station Saint-Michel	45.559425	-73.599749	Metro Station
6	Saint-Michel,East	45.5618	-73.5990	Restaurant Kim Hour	45.561836	-73.605112	Chinese Restaurant
7	Saint-Michel,East	45.5618	-73.5990	Petro-Canada	45.560984	-73.602396	Gas Station

Dataframe after API importing and wrangling

Data

Montreal English Schools

- ▶ Since families are a big stakeholder here, we need the location of all English elementary and high schools in each neighbourhood.
- ▶ A list is available on the website of the English Montreal School Board.



Elementary School	Prog. Ext.	Address	Tel. No.	Fax No.	Principal	Vice Principal
Bancroft	PK B	1001 4563 St Urbain H2T 2V9	514.845.8031	514.845.4352	Dorothy Ostrowicz	
Carlyle	PK E&IB	1002 109 Carlyle, TMR H3R 1S8	514.738.1256	514.738.0373	Dina Vourdouis	
Cedarcrest	I 1003	1505 Muir, St. Laurent H4L 4T1	514.744.2614	514.744.3310	Elena Zervas	
Coronation	PK E&I	1045 4810 Van Horne H3W 1J3	514.733.7790	514.733.7701	Mike Talevi	
Dalkeith	PK E	1004 7951 Dalkeith, Anjou H1K 3X6	514.352.6730	514.352.0243	John Wright	
Dante	PK B	1005 6090 Lachenaie, St. Léonard H1S 1P1	514.254.6644	514.254.6667	Joseph Sabourin	
Dunrae Gardens	PK I	1006 235 Dunrae, TMR H3P 1T5	514.738.1256	514.738.0373	Dina Vourdouis	
East Hill	PK I	1007 10350 Perras, RDP H1C 2H1	514.481.1111	514.481.1111		
Edinburgh	I 1008	500 Hudson, Mtl W. H4X 1X1	514.481.1111	514.481.1111		
Edward Murphy	E&I	1009 6800 Pierre de Coubertin H1N 1T2	514.254.6644	514.254.6667	Joseph Sabourin	
Elizabeth Ballantyne	E	1010 314 Northview, Mtl W. H4X 1E2	514.481.1111	514.481.1111		

Dataframe after reading pdf and wrangling

	Postal Code	School	Type	Neighbourhood	Place Name	Province	Province Code	Latitude	Longitude
0	H1C	1	1	Rivière-des-PrairiesNortheast	Rivière-des-Prairies Northeast	Quebec	QC	45.6656	-73.5367
1	H1E	2	2	Rivière-des-PrairiesSouthwest	Rivière-Des-Prairies Southwest	Quebec	QC	45.6342	-73.5842
2	H1G	2	2	Montréal-NordNorth	Montreal North North	Quebec	QC	45.6109	-73.6211
3	H1H	1	1	Montréal-NordSouth	Montreal North South	Quebec	QC	45.5899	-73.6389
4	H1K	1	1	AnjouEast	Anjou East	Quebec	QC	45.6097	-73.5472
5	H1N	1	1	MercierSoutheast	Mercier Southeast	Quebec	QC	45.5779	-73.5304
6	H1P	1	1	Saint-LéonardNorth	Saint-Léonard North	Quebec	QC	45.5966	-73.5928
7	H1R	1	1	Saint-LéonardWest	Saint-Léonard West	Quebec	QC	45.5864	-73.6082

Original pdf file

Methodology

Data Collection

- ▶ **Requests** library to read data from the websites. It was used in 3 out of our 4 data sets: Wikipedia, GeoNames and Foursquare API.
 - ▶ **Tabula** library to read data from a pdf file, used in our fourth data set from the English Montreal School Board.
 - ▶ The **ZipFile** and **BytesIO** libraries to import and unzip the compressed folder for our GeoNames data.
 - ▶ **Pandas** library to load the data read from all sources into dataframes and analyze the data.
 - ▶ **NumPy** library to further work with and analyze our arrays.
-



Methodology

Initial Review and Stats

- ▶ Montreal has 123 unique postal codes
- ▶ Foursquare provided venue data for 116 of them.
- ▶ Venues fall in to 260 categories.
- ▶ Montreal has 51 schools affiliated with the English Montreal School Board. They are located in 36 of the neighbourhoods.



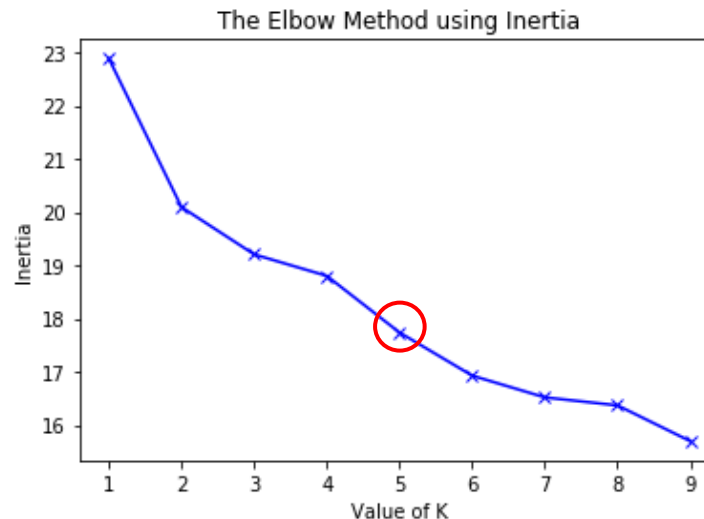
Methodology

Clustering and Segmentation

- ▶ **Sklearn** library is used for K-means clustering.
- ▶ **Folium** library for mapping.
- ▶ Steps:
 1. We run a One Hot encoding process, to prepare the data.

```
montreal_onehot = pd.get_dummies(montreal_venues[['Venue Category']], prefix="", prefix_sep="")
```

2. Using the Elbow method we arrive at 5 clusters.



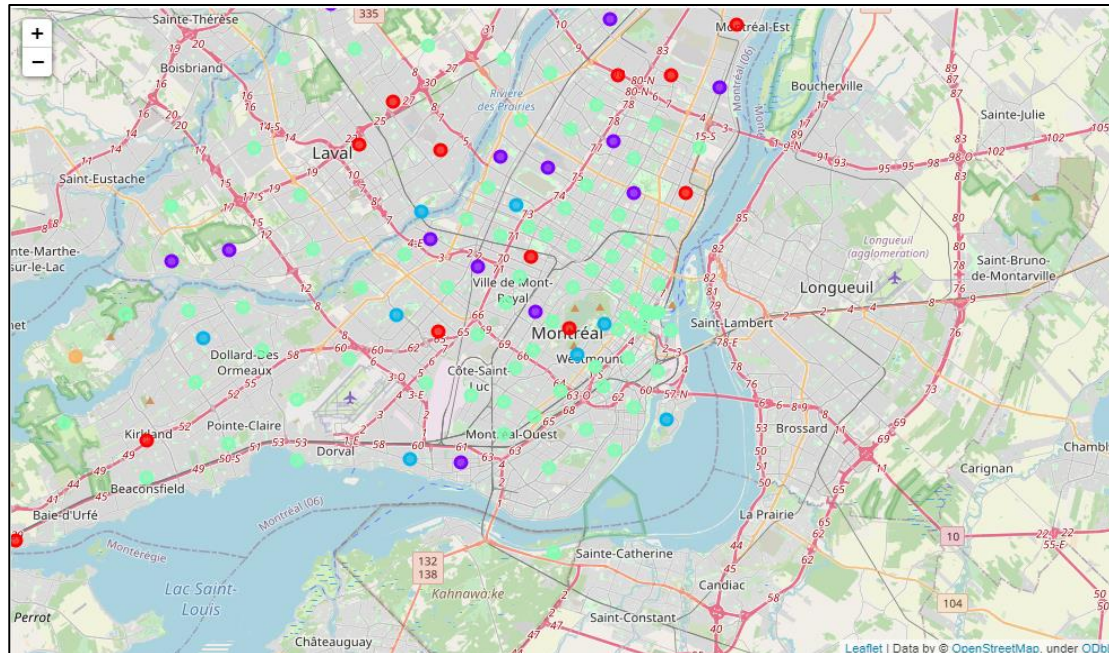
Methodology

Clustering and Segmentation

3. Run Kmeans to cluster our data.

```
# with initialization method Kmeans++  
# and n_init to run 40 times versus default of 10  
kmeans = KMeans(init="k-means++", n_clusters=kclusters, n_init=40, random_state=0).fit(montreal_grouped_clustering)
```

4. Plot the neighbourhoods, color coded per cluster, on a map of Montreal.



Results

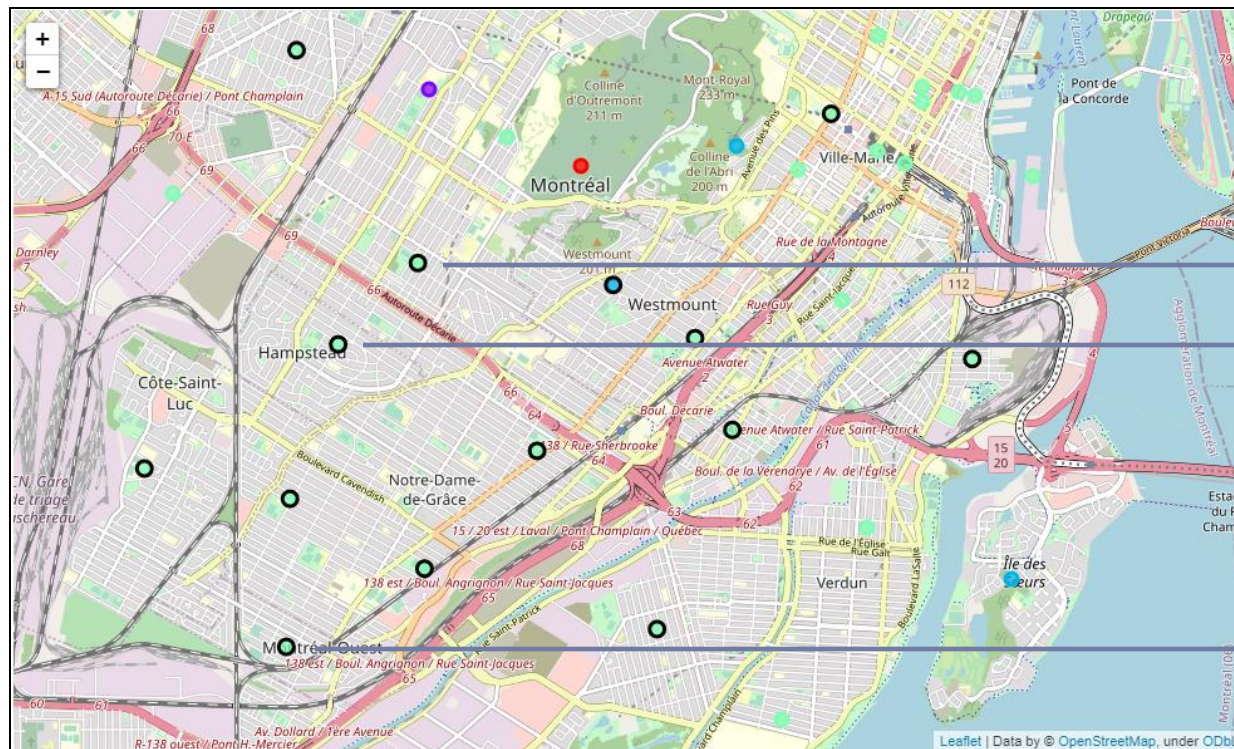
- ▶ Lets merge the neighbourhood data with the school data to see which neighbourhoods could be targeted.

	Postal Code	School	Neighbourhood_x	Cluster Labels
35	H4X	3	Montreal West	3
23	H3X	3	Hampstead / Côte Saint-Luc	3
22	H3W	3	Côte-des-NeigesSouthwest	3
19	H3N	2	Parc-Extension	0
31	H4M	2	Saint-LaurentEast	0
30	H4L	2	Saint-LaurentInner Northeast	3
8	H1S	2	Saint-LéonardSoutheast	1
1	H1E	2	Rivière-des-PrairiesSouthwest	1
17	H3A	2	Downtown Montreal North(McGill University)	3
16	H2T	2	Plateau Mont-RoyalWest	3
33	H4V	2	Côte Saint-LucEast	3
2	H1G	2	Montréal-NordNorth	3
34	H4W	1	Côte Saint-LucWest	3

- ▶ H4X, H3X, H2X stand out as having the most schools and all in the same cluster.

Results

- ▶ Mapping our top targets we can see two are geographically adjacent.



H3W

H3X

H4W

Conclusion

- ▶ Young families move out of downtown Montreal, could focus on the Côte-des-Neiges Southwest, Hampstead / Côte Saint-Luc along with the adjacent, same cluster, area of Westmount South.
- ▶ Real estate agents could find plenty of selling points for the neighbourhoods:
 - ▶ Away from the congested downtown
 - ▶ Convenient for those whose work is still downtown.
 - ▶ Access to services such pharmacies, convenience stores, and restaurants.
 - ▶ Family oriented facilities like: parks, pool, athletic venues.
 - ▶ Most importantly, several elementary and high schools.
- ▶ Overall an ecosystem to allow young families to grow.

