

# **Clustering Canadian Cities into Groups based on Facility Availability**

by Almaz Sabyr

# How is Clustering helpful?

- The cities in Canada are clustered into multiple groups based on availability of essential facilities like: health, education, food, services, transport, etc.
- This clustering of the cities helps both private & public organizations who want to expand their business.
- This helps in identifying cities that have yet to experience major growth. It is cheaper for the organizations to set up business here, due to comparatively lower initial cost & lesser competitors.
- It also helps government to provide basic healthcare and public transport facilities in these cities.

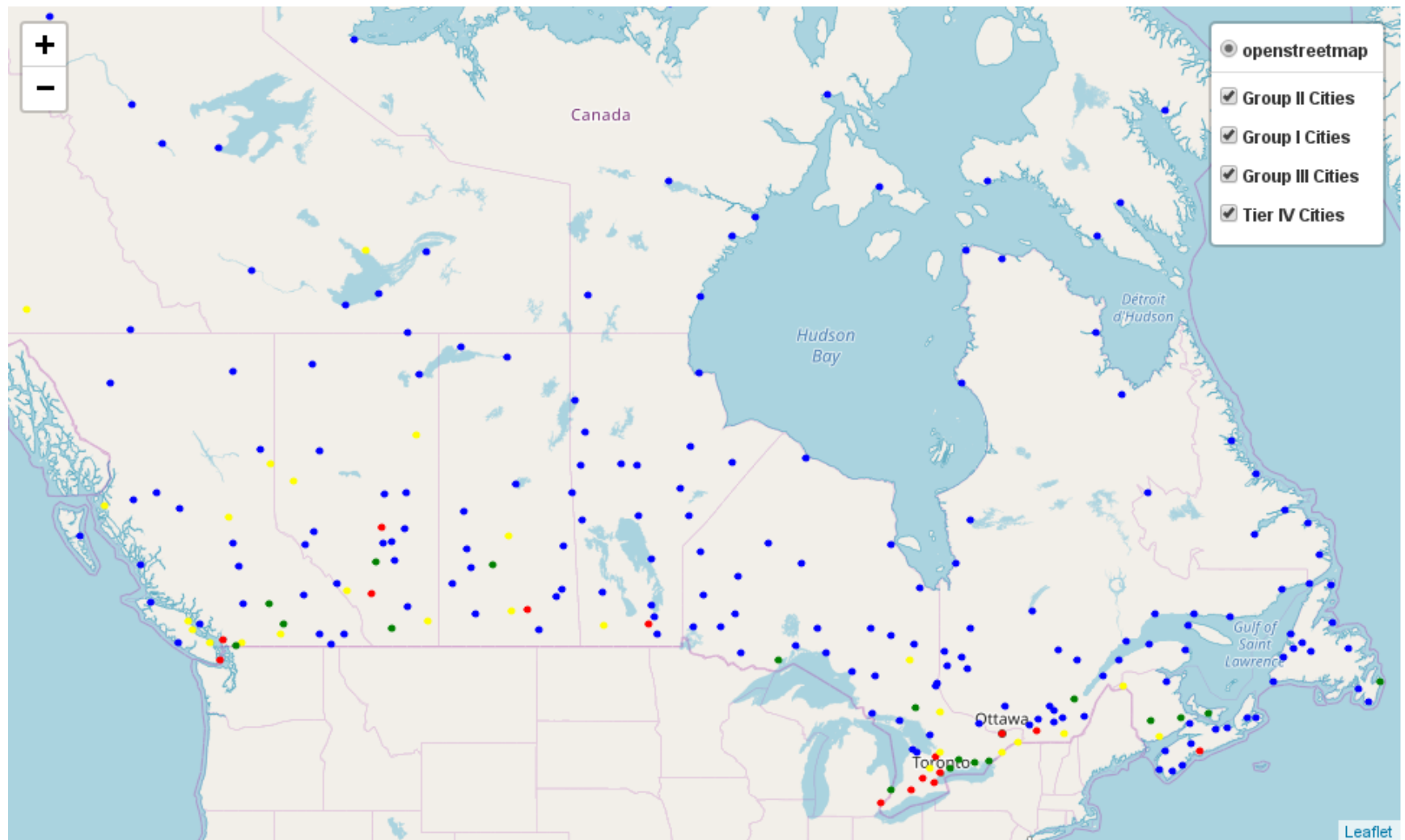
# Data Acquisition

- The dataset of 250 Canadian Cities is collected from [https://simplemaps.com/static/data/world-cities/basic/simplemaps\\_worldcities\\_basicv1.5.zip](https://simplemaps.com/static/data/world-cities/basic/simplemaps_worldcities_basicv1.5.zip). This contains basic information like population, state, latitude, longitude
- Using Foursquare API we get a list of venues categorized as Medical Center, College & University, Food, Shop & Service, Travel & Transport within a radius of 5km for each city
- Higher the number of venues for a city better the level of group

# K-Means Algorithm

- We use K-Means to cluster the cities since we do not have predefined categories.
- Based on the distribution of features we select 4 clusters(groups)
- K-means predicts the labels for the cities as: 0,1,2,3
- The face value of the label is not representative of the group. We have to check the value of the features to assign a group to a label. Higher the values better the level group.
- Once we ascertain the values groups are assigned to the labels.

# Result of K-Means



# Tier I Cities

```
#Group I Cities
tier1=ca_cities_stats[ca_cities_stats[:]['tier']==2]
tier1
```

	city	lat	lng	admin_name	population	medical	education	food	shops	transport	tier
0	Toronto	43.7000	-79.4200	Ontario	5213000.0	50.0	50.0	50.0	50.0	50.0	2
1	Montréal	45.5000	-73.5833	Québec	3678000.0	50.0	50.0	50.0	50.0	25.0	2
2	Vancouver	49.2734	-123.1216	British Columbia	2313328.0	50.0	50.0	50.0	50.0	50.0	2
3	Ottawa	45.4167	-75.7000	Ontario	1145000.0	50.0	24.0	50.0	50.0	31.0	2
4	Calgary	51.0830	-114.0800	Alberta	1110000.0	50.0	27.0	50.0	50.0	37.0	2
5	Edmonton	53.5500	-113.5000	Alberta	1058000.0	50.0	44.0	50.0	50.0	37.0	2
6	Hamilton	43.2500	-79.8300	Ontario	721053.0	50.0	43.0	50.0	50.0	12.0	2
7	Winnipeg	49.8830	-97.1660	Manitoba	632063.0	50.0	48.0	50.0	50.0	30.0	2
10	Kitchener	43.4500	-80.5000	Ontario	417001.0	50.0	29.0	50.0	50.0	17.0	2
11	Halifax	44.6500	-63.6000	Nova Scotia	359111.0	50.0	19.0	50.0	50.0	20.0	2
12	London	42.9700	-81.2500	Ontario	346765.0	50.0	29.0	50.0	50.0	14.0	2
13	Windsor	42.3333	-83.0333	Ontario	319246.0	50.0	44.0	50.0	50.0	18.0	2
14	Victoria	48.4333	-123.3500	British Columbia	289625.0	50.0	20.0	50.0	50.0	18.0	2
16	Barrie	44.3838	-79.7000	Ontario	182041.0	37.0	50.0	25.0	50.0	11.0	2
17	Regina	50.4500	-104.6170	Saskatchewan	176183.0	50.0	22.0	50.0	50.0	15.0	2

# Tier II Cities

```
#Group II Cities
tier2=ca_cities_stats[ca_cities_stats[:]['tier']==1]
tier2
```

	city	lat	lng	admin_name	population	medical	education	food	shops	transport	tier
8	Québec	46.8400	-71.2456	Québec	624177.0	33.0	19.0	4.0	25.0	0.0	1
9	Oshawa	43.8800	-78.8500	Ontario	450963.0	50.0	18.0	28.0	44.0	10.0	1
15	Saskatoon	52.1700	-106.6700	Saskatchewan	198958.0	49.0	23.0	25.0	47.0	13.0	1
18	Sudbury	46.5000	-80.9666	Ontario	157857.0	36.0	45.0	14.0	33.0	3.0	1
19	Abbotsford	49.0504	-122.3000	British Columbia	151683.0	46.0	15.0	21.0	44.0	7.0	1
20	Sarnia	42.9666	-82.4000	Ontario	144172.0	50.0	24.0	11.0	41.0	1.0	1
22	St. John's	47.5850	-52.6810	Newfoundland and Labrador	131469.0	33.0	6.0	8.0	39.0	7.0	1
23	Kelowna	49.9000	-119.4833	British Columbia	125109.0	50.0	26.0	22.0	34.0	7.0	1
25	Kingston	44.2337	-76.4833	Ontario	114195.0	30.0	18.0	23.0	33.0	3.0	1
26	Thunder Bay	48.4462	-89.2750	Ontario	99334.0	27.0	24.0	9.0	38.0	6.0	1
27	Moncton	46.0833	-64.7667	New Brunswick	90635.0	50.0	9.0	13.0	45.0	9.0	1
30	Peterborough	44.3000	-78.3333	Ontario	83627.0	21.0	19.0	20.0	40.0	3.0	1
32	Red Deer	52.2666	-113.8000	Alberta	74857.0	33.0	7.0	17.0	30.0	7.0	1
33	Lethbridge	49.7005	-112.8333	Alberta	70617.0	26.0	14.0	14.0	33.0	5.0	1
34	Kamloops	50.6667	-120.3333	British Columbia	68714.0	33.0	5.0	14.0	30.0	6.0	1
39	Fredericton	45.9500	-66.6333	New Brunswick	52337.0	43.0	11.0	12.0	27.0	4.0	1
45	Belleville	44.1667	-77.3833	Ontario	43990.0	14.0	35.0	14.0	33.0	8.0	1
46	Charlottetown	46.2493	-63.1313	Prince Edward Island	42402.0	26.0	16.0	18.0	32.0	5.0	1

# Tier III Cities

```
#Group III Cities
tier3=ca_cities_stats[ca_cities_stats[:]['tier']==3]
tier3
```

	city	lat	lng	admin_name	population	medical	education	food	shops	transport	tier
21	Sherbrooke	45.4000	-71.9000	Québec	139652.0	10.0	6.0	2.0	11.0	1.0	3
28	Saint John	45.2670	-66.0767	New Brunswick	87857.0	30.0	3.0	11.0	21.0	3.0	3
29	Nanaimo	49.1460	-123.9343	British Columbia	84905.0	23.0	3.0	15.0	17.0	3.0	3
35	Prince George	53.9167	-122.7667	British Columbia	65558.0	23.0	23.0	11.0	19.0	4.0	3
36	Medicine Hat	50.0333	-110.6833	Alberta	63138.0	21.0	6.0	7.0	21.0	5.0	3
40	Chilliwack	49.1666	-121.9500	British Columbia	51942.0	14.0	3.0	10.0	18.0	6.0	3
41	North Bay	46.3000	-79.4500	Ontario	50170.0	14.0	3.0	9.0	18.0	10.0	3
43	Cornwall	45.0171	-74.7333	Ontario	48821.0	24.0	6.0	10.0	22.0	2.0	3
48	Grand Prairie	55.1666	-118.8000	Alberta	41462.0	20.0	7.0	6.0	24.0	5.0	3
49	Penticton	49.5004	-119.5833	British Columbia	37721.0	14.0	6.0	7.0	14.0	2.0	3
51	Orillia	44.6000	-79.4167	Ontario	37483.0	7.0	9.0	12.0	15.0	0.0	3
53	Timmins	48.4666	-81.3333	Ontario	34974.0	5.0	2.0	4.0	20.0	2.0	3
54	Prince Albert	53.2000	-105.7500	Saskatchewan	34609.0	22.0	0.0	6.0	12.0	2.0	3
55	Campbell River	50.0171	-125.2500	British Columbia	33430.0	14.0	2.0	9.0	11.0	1.0	3
56	Courtenay	49.6833	-125.0000	British Columbia	32793.0	14.0	4.0	14.0	17.0	5.0	3
57	Orangeville	43.9171	-80.0833	Ontario	32640.0	13.0	3.0	5.0	11.0	2.0	3
58	Moose Jaw	50.4000	-105.5500	Saskatchewan	32166.0	12.0	2.0	7.0	10.0	3.0	3
59	Brandon	49.8333	-99.9500	Manitoba	28418.0	24.0	4.0	8.0	22.0	2.0	3
60	Brockville	44.5893	-75.6953	Ontario	26458.0	8.0	2.0	3.0	18.0	4.0	3
64	Whitehorse	60.7167	-135.0500	Yukon	23276.0	8.0	3.0	5.0	9.0	5.0	3
66	Fort McMurray	56.7333	-111.3833	Alberta	21863.0	19.0	1.0	13.0	24.0	1.0	3
72	Yellowknife	62.4420	-114.3970	Northwest Territories	19234.0	6.0	1.0	7.0	12.0	1.0	3
75	Edmundston	47.3794	-68.3333	New Brunswick	17894.0	12.0	0.0	1.0	10.0	2.0	3
81	Prince Rupert	54.3167	-130.3300	British Columbia	14708.0	8.0	1.0	3.0	13.0	1.0	3
92	Dawson Creek	55.7670	-120.2333	British Columbia	10802.0	9.0	8.0	5.0	7.0	2.0	3
105	Banff	51.1780	-115.5719	Alberta	7502.0	5.0	0.0	4.0	15.0	1.0	3



# Tier IV Cities

```
# Group IV Cities
tier4=ca_cities_stats[ca_cities_stats[:]['tier']==0]
tier4
```

	city	lat	lng	admin_name	population	medical	education	food	shops	transport	tier
24	Trois-Rivières	46.3500	-72.5499	Québec	119693.0	6.0	8.0	0.0	5.0	0.0	0
31	St-Jerome	45.7666	-74.0000	Québec	78439.0	6.0	0.0	0.0	7.0	0.0	0
37	Drummondville	45.8833	-72.4834	Québec	59489.0	3.0	7.0	0.0	2.0	0.0	0
38	Chicoutimi	48.4333	-71.0667	Québec	53940.0	3.0	0.0	0.0	3.0	0.0	0
42	Shawinigan	46.5504	-72.7333	Québec	49161.0	3.0	5.0	0.0	2.0	0.0	0
44	Joliette	46.0333	-73.4333	Québec	45361.0	2.0	0.0	0.0	2.0	1.0	0
47	Victoriaville	46.0504	-71.9667	Québec	41500.0	3.0	1.0	0.0	2.0	0.0	0
50	Sydney	46.0661	-60.1800	Nova Scotia	37538.0	2.0	0.0	0.0	0.0	0.0	0
52	Rimouski	48.4337	-68.5167	Québec	35584.0	0.0	0.0	0.0	2.0	0.0	0
61	Saint-Georges	46.1171	-70.6667	Québec	26149.0	1.0	0.0	0.0	1.0	0.0	0
62	Sept-Îles	50.3161	-66.3600	Québec	25686.0	0.0	0.0	0.0	0.0	0.0	0
63	Rouyn-Noranda	48.2500	-79.0332	Québec	24602.0	4.0	0.0	0.0	1.0	0.0	0
65	Owen Sound	44.5666	-80.8500	Ontario	22625.0	1.0	1.0	1.0	1.0	1.0	0
67	Corner Brook	48.9500	-57.9333	Newfoundland and Labrador	20791.0	2.0	5.0	1.0	8.0	1.0	0
68	Val d'Or	48.1166	-77.7666	Québec	20625.0	1.0	0.0	0.0	0.0	0.0	0
69	New Glasgow	45.5833	-62.6333	Nova Scotia	20322.0	4.0	0.0	1.0	11.0	5.0	0
70	Terrace	54.5000	-128.5833	British Columbia	19443.0	8.0	2.0	2.0	8.0	1.0	0
71	North Battleford	52.7666	-108.2833	Saskatchewan	19440.0	4.0	2.0	6.0	6.0	2.0	0
73	Fort St. John	56.2500	-120.8333	British Columbia	18776.0	4.0	1.0	7.0	7.0	1.0	0
74	Cranbrook	49.5167	-115.7667	British Columbia	18610.0	4.0	2.0	3.0	8.0	2.0	0
76	Rivière-du-Loup	47.8333	-69.5333	Québec	16403.0	2.0	0.0	0.0	1.0	0.0	0
77	Camrose	53.0167	-112.8166	Alberta	15808.0	6.0	1.0	7.0	8.0	2.0	0
78	Pembroke	45.8503	-77.1166	Ontario	15551.0	1.0	2.0	5.0	5.0	4.0	0
79	Yorkton	51.2171	-102.4665	Saskatchewan	15172.0	5.0	1.0	4.0	5.0	1.0	0
80	Swift Current	50.2837	-107.7666	Saskatchewan	14906.0	5.0	1.0	1.0	2.0	0.0	0
82	Williams Lake	52.1166	-122.1500	British Columbia	14168.0	6.0	0.0	6.0	3.0	0.0	0
83	Brooks	50.5671	-111.9000	Alberta	14163.0	5.0	1.0	3.0	4.0	2.0	0
84	Quesnel	52.9837	-122.4833	British Columbia	13788.0	6.0	2.0	3.0	2.0	1.0	0
85	Thompson	55.7499	-97.8666	Manitoba	13727.0	4.0	1.0	2.0	4.0	0.0	0
86	Dolbeau	48.8666	-72.2333	Québec	13337.0	0.0	0.0	0.0	0.0	0.0	0

# Statistics of the Tiers

```
#Stats of each tier  
tier_stats
```

	Tier	Avg_Population	Avg_Medical	Avg_Education	Avg_Food	Avg_Shops	Avg_Transport
0	Tier 1	1197361	49	36	48	50	25
1	Tier 2	151394	36	18	15	36	5
2	Tier 3	41518	14	4	7	15	2
3	Tier 4	7204	1	0	0	1	0

# Conclusion

- Thus idea to cluster cities reveals a major disparity in the availability of facilities across the Groups
- Lower level Groups of Cities have yet to experience growth which makes them potentially rewarding for setting up businesses since even these cities have a healthy population.