

YOUR HOME AWAY FROM HOME

Finding a new home in your neighbourhood, but in another city!

By Sameh A. Rasoul



RELATING QUALITY OF LIFE IN PROPERTY LISTINGS TO YOUR HOME TOWN

- Which property to select when one is relocating to an entirely different city.?
 - Which region?
 - How to go about it?
 - should I ask friends? or
 - better look-up reviews?
- Why not scrape the listings results page the user is looking at and then cluster them together with the neighbourhoods in the home city of the user, showing the user some relevant details about the listing, similar neighborhoods back home and additional features all together in an interactive map.



DATA ACQUISITION & CLEANING

- Data Sources:

1. The property listings website results page and the page of each individual listing in the results page
2. Wikipedia page including Toronto's postal codes and an online csv archive of geolocation of Toronto's neighborhoods.
3. Foursquare venue data for all property listings and Toronto's neighborhoods

Formatting & Concatenating



LOCATIONS DATASET

- The locations dataset including property listings: (25 listings)

```
1 DF.sample(5)
```

	Postal Code	Borough	Neighbourhood	Latitude	Longitude	Link	Pic
61	NaN	NaN	1,100,000 AED	25.206125	55.343607	https://www.propertyfinder.ae/en/buy/apartmen...	https://www.propertyfinder.ae/property/fc117ae...
54	NaN	NaN	1,199,000 AED	25.087251	55.145574	https://www.propertyfinder.ae/en/buy/apartmen...	https://www.propertyfinder.ae/property/7ea868e...
18	M4N	Central Toronto	Lawrence Park	43.728	-79.3888	NaN	NaN
14	M6K	West Toronto	Brockton, Parkdale Village, Exhibition Place	43.6368	-79.4282	NaN	NaN
44	NaN	NaN	1,300,000 AED	25.189836	55.27576	https://www.propertyfinder.ae/en/buy/apartmen...	https://www.propertyfinder.ae/property/8455cdb...



VENUES DATA SET

- 2912 venues data were harvested

```
[44]: print(all_venues.shape)  
      all_venues.sample(10)
```

```
(2912, 7)
```

```
[44]:
```

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
2379	1,100,000 AED	25.082217	55.142624	The Radisson Blu Residence, Dubai Marina	25.078288	55.143216	Hotel
1984	1,049,987 AED	25.191107	55.26991	Gazebo	25.190844	55.266950	Indian Restaurant
641	Little Portugal, Trinity	43.6479	-79.4197	Ufficio	43.649439	-79.423014	Italian Restaurant
487	Richmond, Adelaide, King	43.6506	-79.3846	Booster Juice	43.648898	-79.383351	Juice Bar
199	St. James Town	43.6515	-79.3754	Downtown Camera	43.653107	-79.375120	Camera Store
1911	1,250,000 AED	25.073151	55.136982	Zafran Indian Bistro	25.076842	55.139459	Indian Restaurant
1419	St. James Town, Cabbagetown	43.668	-79.3677	China Gourmet	43.664180	-79.368359	Chinese Restaurant
1539	Church and Wellesley	43.6659	-79.3832	Como En Casa	43.665160	-79.384796	Mexican Restaurant
4	Regent Park, Harbourfront	43.6543	-79.3606	Impact Kitchen	43.656369	-79.356980	Restaurant
2404	1,199,000 AED	25.087251	55.145574	Habtoor Grand Resort, Autograph Collection	25.085991	55.141161	Resort



BUILDING THE MODEL

- Frequency of occurrence for 2912 venues around 64 unique locations (listings & neighbourhoods)

----1,299,999 AED----

	venue	freq
0	Café	0.09
1	Resort	0.06
2	Restaurant	0.06
3	Italian Restaurant	0.06
4	Coffee Shop	0.06

----1,300,000 AED----

	venue	freq
0	Coffee Shop	0.09
1	Middle Eastern Restaurant	0.07
2	Hotel	0.07
3	Café	0.05
4	Restaurant	0.04

----Central Bay Street----

	venue	freq
0	Coffee Shop	0.18
1	Café	0.06
2	Italian Restaurant	0.04
3	Sandwich Place	0.04
4	Japanese Restaurant	0.03

----Christie----

	venue	freq
0	Grocery Store	0.25
1	Café	0.19
2	Park	0.12
3	Coffee Shop	0.06
4	Italian Restaurant	0.06



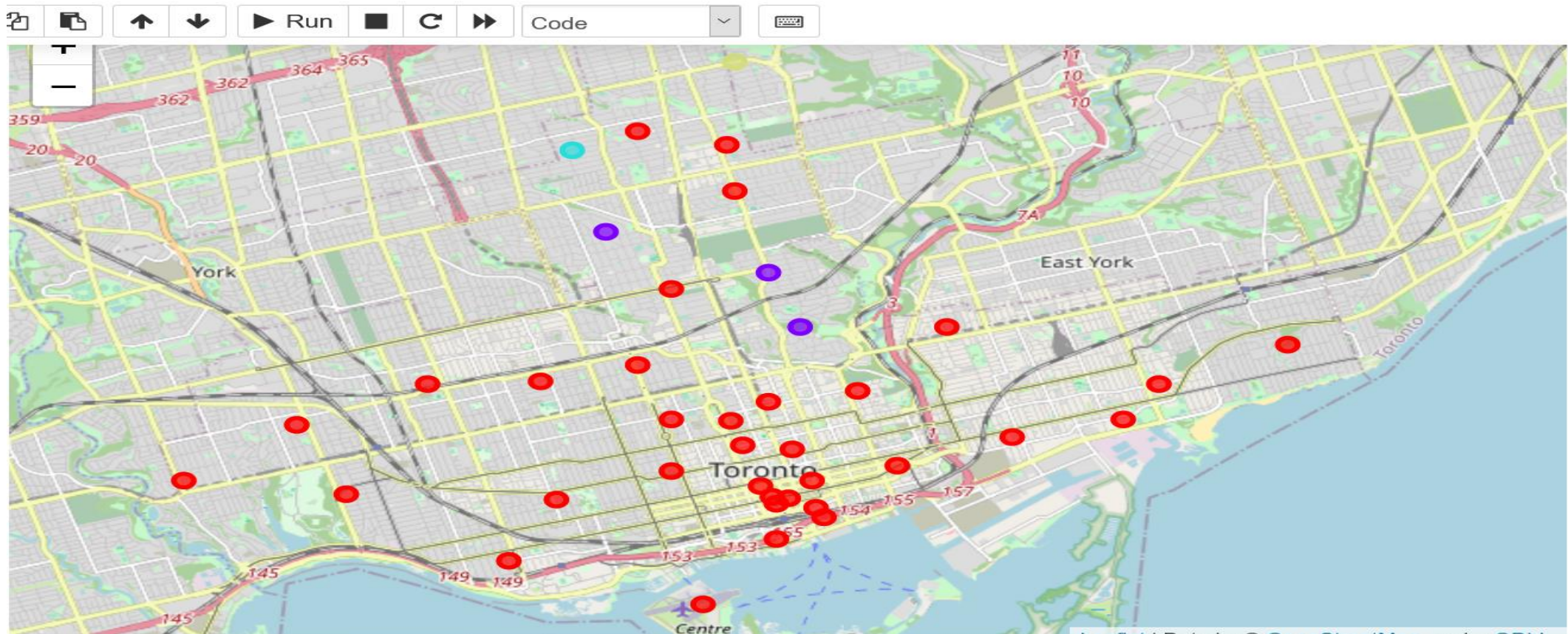
K-MEANS CLUSTERING

- Dubai Properties: with $K = 4$, All selected properties fit in the same cluster



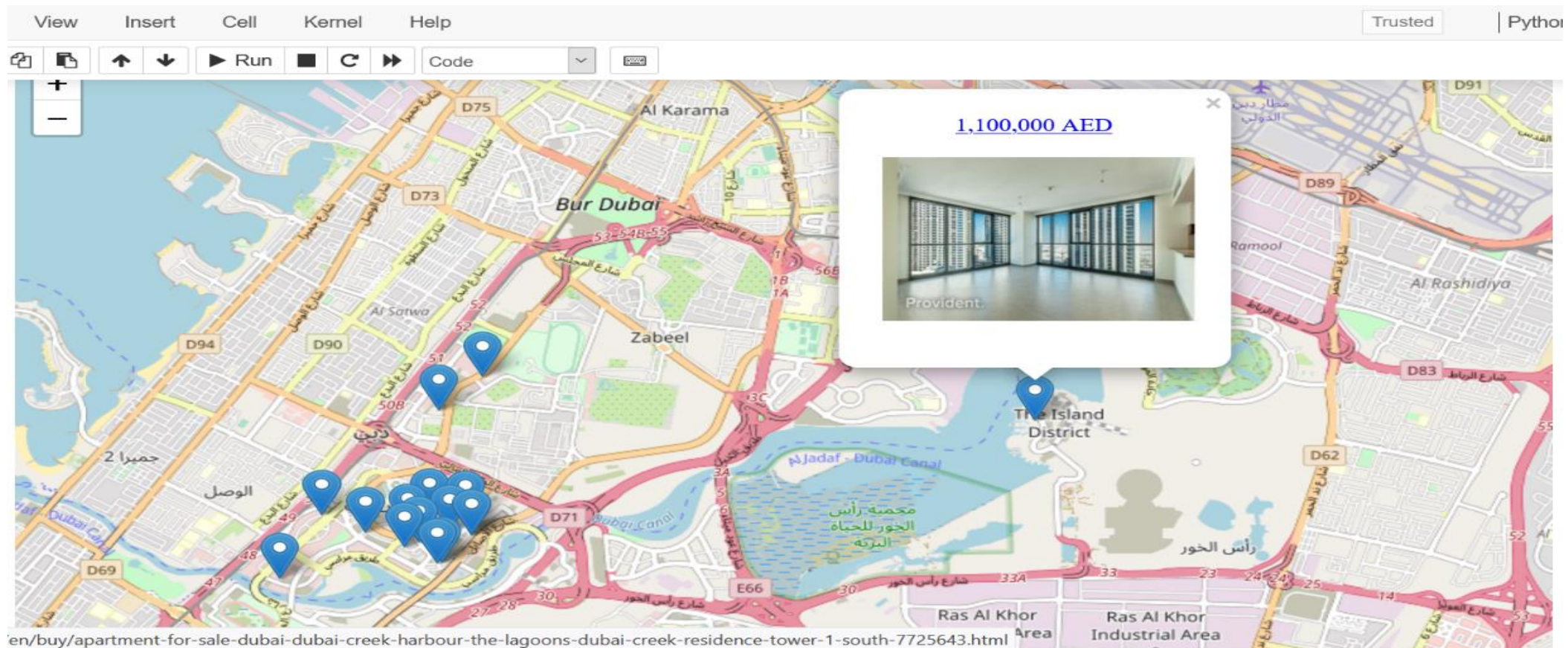
TORONTO CLUSTERS

- Comparison with home town neighbourhoods in Toronto clustering results



INTERACTIVE MAP FUNCTIONALITY

- HTML pop-ups with hyperlinked details.



SAMPLE LOCATIONS

- In depth comparison of most common venues

```
1 locations_venues_sorted.sample(5)
```

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
37	0	Parkdale, Roncesvalles	Breakfast Spot	Gift Shop	Cuban Restaurant	Eastern European Restaurant	Bookstore	Movie Theater	Dessert Shop	Italian Restaurant	Bar	Restaurant
29	0	Harbourfront East, Union Station, Toronto Islands	Coffee Shop	Aquarium	Hotel	Café	Restaurant	Fried Chicken Joint	Brewery	Scenic Lookout	Park	Baseball Stadium
22	0	Commerce Court, Victoria Hotel	Coffee Shop	Restaurant	Café	Hotel	Gym	American Restaurant	Japanese Restaurant	Deli / Bodega	Seafood Restaurant	Bakery
1	0	1,040,000 AED	Café	Spa	Cocktail Bar	Coffee Shop	Harbor / Marina	Chinese Restaurant	Hotel	Middle Eastern Restaurant	Sports Bar	Latin American Restaurant
10	0	1,217,819 AED	Hotel	Café	Restaurant	Middle Eastern Restaurant	Coffee Shop	Seafood Restaurant	Breakfast Spot	Lounge	Chinese Restaurant	American Restaurant



CONCLUSIONS

- The relevant neighbourhood from Toronto was assigned based on surrounding venues categories.
- conclusion was reached based on a predictive clustering model applied to a selection of listings on a results page from propertyfinder.ae.
- The model can be easily reapplied to any other results page from different cities on the same website, and easily adapted to scrape any listings website.

▪ *FUTURE DIRECTION*

- Expand the dataset to encompass other sources of data in addition to foursquare, such as google maps reviews, trip-advisor, etc. and also official data such as crime rate, insurance rates, etc.
- Test more clustering models.

