Cousera Capstone Project Report

Best Neighborhood to Rent in Hamilton

Author: Zhi Zhang

August, 2020

1. Introduction

Hamilton is a Canadian port city on the western tip of Lake Ontario, it is famous for its amazing escarpment views and numerous waterfalls, it is a great place to live in. Many people working in Toronto are moving to Hamilton because of the great natural views, cheaper rent and convenient transportation. The objective of this project is to analyze the neighborhoods in Hamilton based on the featured venues and the average rent price to help people who are considering moving to Hamilton to find the best place to rent that suit their needs.

2. Data

The average rent data is collected from Zumper webpage(https://www.zumper.com/rent-research/hamilton-on), it contains the median rent for 1 bedroom apartment for all neighborhoods in Hamilton. Geopandas and Geopy libraries will be used to obtain the coordinates of the neighborhoods, and then Foursquare API will be used to explore the neighborhoods in Hamilton. I will use the explore function to get the most common venue categories in each neighborhood and then use this feature combined with the average rent to group the neighborhoods into clusters using the k-means clustering algorithm. Finally, I will use the Folium library to visualize the neighborhoods in Hamilton and their merging clusters.

3. Methodology

3.1 Download and Explore Dataset

Since the rent data is on a webpage and not downloadable, I created a python dictionary to store the average rent of each neighborhood in Hamilton, then transformed it into a pandas dataframe. Then using the geocodes function from Geopy library, I obtained the latitude and longitude for the neighborhoods, and dropped the rows where the location data is missing. Here's the rent dataframe after cleaning.

	index	Neighbourhood	MedianRent	Latitude	Longitude
0	0	Durand	1300	43.250247	-79.875734
1	1	Central Hamilton	1675	43.256080	-79.872858
2	3	Beasley	1395	43.259204	-79.861012
3	4	Corktown	1375	43.250681	-79.868619
4	5	Gibson	1199	43.257866	-79.839098
5	6	Kirkendall North	1720	39.977308	-86.047118
6	9	Westdale South	720	43.261881	-79.905921
7	10	Riverdale West	1450	43.228332	-79.759820
8	12	Stinson	1310	43.246953	-79.852747
9	13	Strathcona	1272	43.265244	-79.883693
10	14	St. Clair	1250	39.128103	-84.517178
11	15	Rosedale	1424	43.226083	-79.807812
12	16	Raleigh	1348	43.204820	-79.842024
13	18	North End East	1200	43.269583	-79.857872
14	19	Waterdown	2150	43.331421	-79.895668
15	22	University Gardens	1375	43.263222	-79.936684
16	23	Hill Park	1115	-37.857526	175.679765
17	26	Greenford	1650	43.225812	-79.769345

3.2 Explore Neighborhoods in Hamilton, ON

After obtaining the location data, I utilized the Foursquare API to explore the neighborhoods. I defined getNearbyVenues function that takes the name, latitude, longitude of the neighborhood and returns top 100 venues within 500 meters. Running this function on each neighborhood, I got the results and created a hamilton_venues dataframe that returns all satisfied venues.

```
hamilton_venues.head()
```

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Durand	43.250247	-79.875734	One Duke Restaurant and Lounge	43.251812	-79.871852	Seafood Restaurant
1	Durand	43.250247	-79.875734	Red Crow Coffee	43.250061	-79.871915	Café
2	Durand	43.250247	-79.875734	Durand Coffee	43.251520	-79.878845	Café
3	Durand	43.250247	-79.875734	Bronzie's Place	43.250541	-79.871633	Italian Restaurant
4	Durand	43.250247	-79.875734	The Pheasant Plucker	43.251970	-79.870248	Pub

3.3 Analyze Each Neighborhood

In order to analyze the hamilton_venues data, Venue has to be converted from categorical variable to dummy/indicator variables. By applying get_dummies function from Pandas, the dataframe is changed to the following:

	Neighborhood	Adult Boutique	American Restaurant	Arcade	Art Gallery	Arts & Crafts Store	Asian Restaurant	Auto Garage	Bagel Shop	Bakery	Bank	Bar	Beer Store	Big Box Store	Bookstore	Breakfast Spot	Brewery
0	Durand	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Durand	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Durand	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Durand	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Durand	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Then sort the venues in descending order and only store the top 10 common venues for each neighborhood to cluster the neighborhoods.

```
neighborhoods_venues_sorted.head()
```

	Neighborhood	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
0	Beasley	Middle Eastern Restaurant	Coffee Shop	Vietnamese Restaurant	Fast Food Restaurant	Pharmacy	Beer Store	Asian Restaurant	Theater	Sushi Restaurant	Dog Run
1	Central Hamilton	Coffee Shop	Pub	Fast Food Restaurant	Sandwich Place	Bar	Café	Middle Eastern Restaurant	Indian Restaurant	Hotel	Burrito Place
2	Corktown	Pub	Italian Restaurant	Park	Sandwich Place	Fast Food Restaurant	Pizza Place	Mexican Restaurant	Restaurant	Coffee Shop	Seafood Restaurant
3	Durand	Pub	Café	Italian Restaurant	Pharmacy	Ethiopian Restaurant	Fast Food Restaurant	Breakfast Spot	Seafood Restaurant	Bank	Pizza Place
4	Gibson	Restaurant	Coffee Shop	Gas Station	Library	Fast Food Restaurant	Cupcake Shop	Deli / Bodega	Department Store	Dessert Shop	Diner

3.4 Cluster Neighborhoods

K-means algorithm is being used to cluster neighborhoods because it is fast and easy to implement. Using the K-means cluster function imported from sklearn library, I set the number of clusters to be 5, and I obtained the cluster labels for each neighborhood. Adding the cluster labels to the dataframe, and join this dataframe with the rent dataframe, I got the final dataframe that will be used to visualize the clustered map.

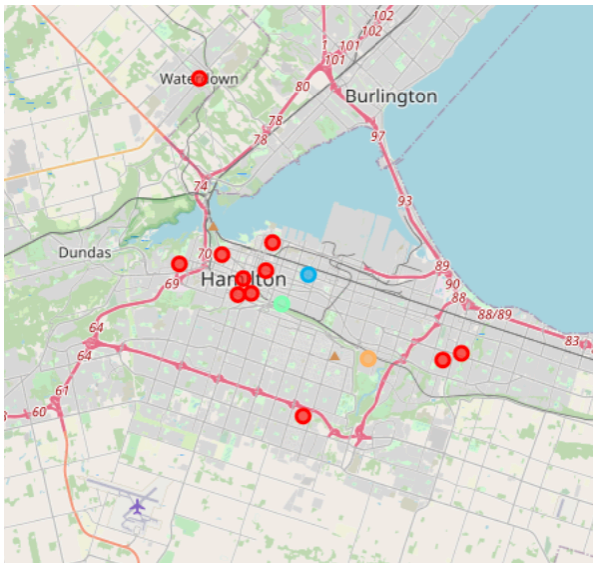
```
hamilton_merged.head()
```

	Neighborhood	MedianRent	Latitude	Longitude	Cluster Labels	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	9t Co
0	Durand	1300	43.250247	-79.875734	0	Pub	Café	Italian Restaurant	Pharmacy	Ethiopian Restaurant	Fast Food Restaurant	Breakfast Spot	Seafood Restaurant	
1	Central Hamilton	1675	43.256080	-79.872858	0	Coffee Shop	Pub	Fast Food Restaurant	Sandwich Place	Bar	Café	Middle Eastern Restaurant	Indian Restaurant	
3	Beasley	1395	43.259204	-79.861012	0	Middle Eastern Restaurant	Coffee Shop	Vietnamese Restaurant	Fast Food Restaurant	Pharmacy	Beer Store	Asian Restaurant	Theater	Res
4	Corktown	1375	43.250681	-79.868619	0	Pub	Italian Restaurant	Park	Sandwich Place	Fast Food Restaurant	Pizza Place	Mexican Restaurant	Restaurant	
5	Gibson	1199	43.257866	-79.839098	2	Restaurant	Coffee Shop	Gas Station	Library	Fast Food Restaurant	Cupcake Shop	Deli / Bodega	Department Store	

4. Results

Folium library is being used to visualized the clustered neighborhoods.

Before Clustering



After Clustering

When click on each neighborhood, a popup message will display the name of the neighborhood, the cluster group and the average rent in this neighborhood.

In cluster 1, there are 13 neighborhoods,

	Neighborhood	MedianRent	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
0	Durand	1300	Pub	Italian Restaurant	Café	Breakfast Spot	Fast Food Restaurant	Mexican Restaurant	Ethiopian Restaurant	Pharmacy	Pizza Place	Seafood Restaurant
1	Central Hamilton	1675	Coffee Shop	Pub	Fast Food Restaurant	Restaurant	American Restaurant	Middle Eastern Restaurant	Indian Restaurant	Sandwich Place	Hotel	Nightclub
2	Beasley	1395	Middle Eastern Restaurant	Pharmacy	Sushi Restaurant	Vietnamese Restaurant	Beer Store	Hostel	Theater	Fast Food Restaurant	Coffee Shop	Farm
3	Corktown	1375	Pub	Italian Restaurant	Park	Sandwich Place	Historic Site	Sports Bar	Gas Station	Fast Food Restaurant	Ethiopian Restaurant	Mexican Restaurant
6	Westdale South	720	Coffee Shop	Sandwich Place	Fried Chicken Joint	Burrito Place	Burger Joint	Flower Shop	Indie Movie Theater	Supermarket	Mediterranean Restaurant	Restaurant
7	Riverdale West	1450	Fast Food Restaurant	Frozen Yogurt Shop	Pizza Place	Sandwich Place	Chinese Restaurant	Shopping Mall	Bus Station	Liquor Store	Lighting Store	Breakfast Spot
9	Strathcona	1272	Yoga Studio	Theater	Coffee Shop	Fast Food Restaurant	Park	Gas Station	Gastropub	Hotel	Pharmacy	History Museum
10	St. Clair	1250	Bar	Sushi Restaurant	Fast Food Restaurant	Pizza Place	Indian Restaurant	Dessert Shop	Sandwich Place	Yoga Studio	Deli / Bodega	Ethiopian Restaurant
12	Raleigh	1348	Wings Joint	Restaurant	Sandwich Place	Coffee Shop	Pharmacy	Yoga Studio	Fast Food Restaurant	Department Store	Dessert Shop	Diner
13	North End East	1200	Convenience Store	Restaurant	Coffee Shop	Skating Rink	Breakfast Spot	Brewery	Fish & Chips Shop	Dessert Shop	Diner	Discount Store
14	Waterdown	2150	Grocery Store	Ice Cream Shop	Bank	Snack Place	Fast Food Restaurant	Farmers Market	Diner	Cosmetics Shop	Park	Pub
16	Hill Park	1115	Historic Site	Museum	Garden	Lake	Farm	Exhibit	Pub	Trail	Tree	Auto Garage
17	Greenford	1650	Pizza Place	Fast Food Restaurant	Restaurant	Burger Joint	Gas Station	Bus Station	Sandwich Place	Italian Restaurant	Fried Chicken Joint	Coffee Shop

In cluster 2-5, there is only one neighborhood for each cluster:

Cluster 2

```
hamilton_merged.loc[hamilton_merged['Cluster Labels'] == 1, hamilton_merged.columns[[0]+[1] + list(range(5, hamilton_me
```

	Neighborhood	MedianRent	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
5	Kirkendall North	1720	Park	Golf Course	Dog Run	Yoga Studio	Fish & Chips Shop	Department Store	Dessert Shop	Diner	Discount Store	Ethiopian Restaurant

Cluster 3

```
hamilton_merged.loc[hamilton_merged['Cluster Labels'] == 2, hamilton_merged.columns[[0]+[1] + list(range(5, hamilton_me
```

	Neighborhood	MedianRent	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
4	Gibson	1199	Restaurant	Coffee Shop	Gas Station	Library	Yoga Studio	Fish & Chips Shop	Dessert Shop	Diner	Discount Store	Dog Run

Cluster 4

```
hamilton_merged.loc[hamilton_merged['Cluster Labels'] == 3, hamilton_merged.columns[[0]+[1] + list(range(5, hamilton_me
```

	Neighborhood	MedianRent	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
8	Stinson	1310	Park	Trail	Coffee Shop	Lawyer	Locksmith	Yoga Studio	Fish & Chips Shop	Dessert Shop	Diner	Discount Store

Cluster 5

```
hamilton_merged.loc[hamilton_merged['Cluster Labels'] == 4, hamilton_merged.columns[[0]+[1] + list(range(5, hamilton_me
```

	Neighborhood	MedianRent	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
11	Rosedale	1424	Park	Convenience Store	Chinese Restaurant	Burrito Place	Yoga Studio	Fish & Chips Shop	Diner	Discount Store	Dog Run	Ethiopian Restaurant

5. Discussion

The number of neighborhoods in clusters is uneven, where cluster 1 has the most neighborhoods while the rest clusters only have one neighborhood each. Examining this on the clustered map, it makes sense because the neighborhoods in cluster 1 are mainly located in Downtown Hamilton.

Looking at the specific date for each cluster, we will find that neighborhoods in cluster 1 have many restaurants, and there's a wide range of average rent, from $720 to $2150, so any food-lovers can find rent in a neighborhood within cluster 1 that can fit their budget.

The neighborhood in Cluster 2 has a relatively high average rent, it is suitable for those who has a higher budget and love sports like golf, soccer, yoga, etc.. It is also a great neighborhood for dog-owners because it has many dog runs.

The neighborhood in Cluster 3 is a very good place to live in because the average rent is reasonable, and there are many restaurants, coffee shops and gas stations, which makes it ideal for commuters and workers.

The neighborhoods in Cluster 4, 5 both have similar average rent price, and they have many stores, restaurants and parks. It will be suitable for people with a budget around $1300 to $1400 and do not want to live in crowded areas.

6. Conclusion

In this project, I collected the location data for neighborhoods in Hamilton, ON, and clustered them based on their top 10 common venues, combined with the average rent data, the results give potential renters in Hamilton a better idea of which neighborhood they should choose. The result of this project could help people who are not familiar with Hamilton to get a better idea of where to rent in Hamilton.