THE BATTLE OF NEIGHBORHOODS

INTRODUCTION / BUSINESS PROBLEM

- A resident of Clairville in Etobicoke of Toronto, Canada has been offered a job opportunity on the other side of Toronto at Highland Creek in Scarborough.

- He wants to find out more on whether the services and amenities available in Highland Creek or in nearby neighborhoods would be comparable or better than those in Clairville.

INTRODUCTION / BUSINESS PROBLEM

- Problem: Compare the top 10 venues of each neighborhood based on the frequency of each venue in each neighborhood

- Target audience: Resident of Clairville in Etobicoke

INTRODUCTION / BUSINESS PROBLEM

- Characteristics and interests of the resident:

  - dog owner

  - typically eats out with preference for Turkish cuisine

  - enjoys social drinking with friends and colleagues on regular basis

  - occasionally rents a car to drive out of town to have short vacations over the weekend or while on leave

DATA

- Borough and neighborhood data from Wikipedia;

- Geospatial data from the CSV file; and

- Top 10 venues in each neighborhood pulled via FourSquare API to find out the the services and amenities for each neighborhood.

DATA

- Fields needed from data:

  - Boroughs and neighborhoods extracted from Wikipedia

  - Longitudes and latitudes of neighborhoods extracted from geospatial data

  - Longitudes and latitudes of venues per neighborhood extracted via FourSquare API

  - Type of each venue in each neighborhood extracted via FourSquare API

DATA SANITISATION

- Use pandas to extract the boroughs and neighborhood data from Wikipedia into a data frame

- Drop all rows with boroughs 'Not assigned'

- Extract geospatial data from the CSV

- Merge data together with original data frame to form a new data frame containing latitude and longitude for each unique postal code

METHODOLOGY

- Create API request URL and GET request to pull nearby venues of each neighborhood from FourSquare based on its name, latitude and longitude

- Process each neighborhood via one-hot encoding

- Group rows by neighborhoods and by taking the mean of the frequency of occurrence of each category

	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	Agincourt	Lounge	Latin American Restaurant	Breakfast Spot	Skating Rink	Escape Room	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Dim Sum Restaurant	Ethiopian Restaurant
1	Alderwood, Long Branch	Pizza Place	Pharmacy	Pub	Gym	Sandwich Place	Coffee Shop	Skating Rink	Dessert Shop	Dim Sum Restaurant	Diner
2	Bathurst Manor, Wilson Heights, Downsview North	Bank	Coffee Shop	Fried Chicken Joint	Frozen Yogurt Shop	Restaurant	Supermarket	Sushi Restaurant	Middle Eastern Restaurant	Ice Cream Shop	Deli / Bodega
3	Bayview Village	Café	Japanese Restaurant	Bank	Chinese Restaurant	Women's Store	Dim Sum Restaurant	Discount Store	Distribution Center	Dog Run	Doner Restaurant
4	Bedford Park, Lawrence Manor East	Italian Restaurant	Coffee Shop	Sandwich Place	Liquor Store	Juice Bar	Indian Restaurant	Restaurant	Sushi Restaurant	Café	Pub

TABLE 1: EXTRACT OF DATA FRAME CONTAINING TOP 10 MOST COMMON VENUES FOR NEIGHBORHOODS

METHODOLOGY

- Run k-means clustering of the neighborhoods of Toronto into 7 clusters via a reiterative process

- K-means clustering assists in zooming in on neighborhoods in the same cluster as Clairville, and allows analysis to be made on whether Highland Creek is sufficiently similar to to Clairville

- Allows decisions to be made quickly on whether the resident should take up the job opportunity if his new neighborhood or nearby neighborhoods is similar to Clairville

- There were no significant changes in the clusters for additional clusters created
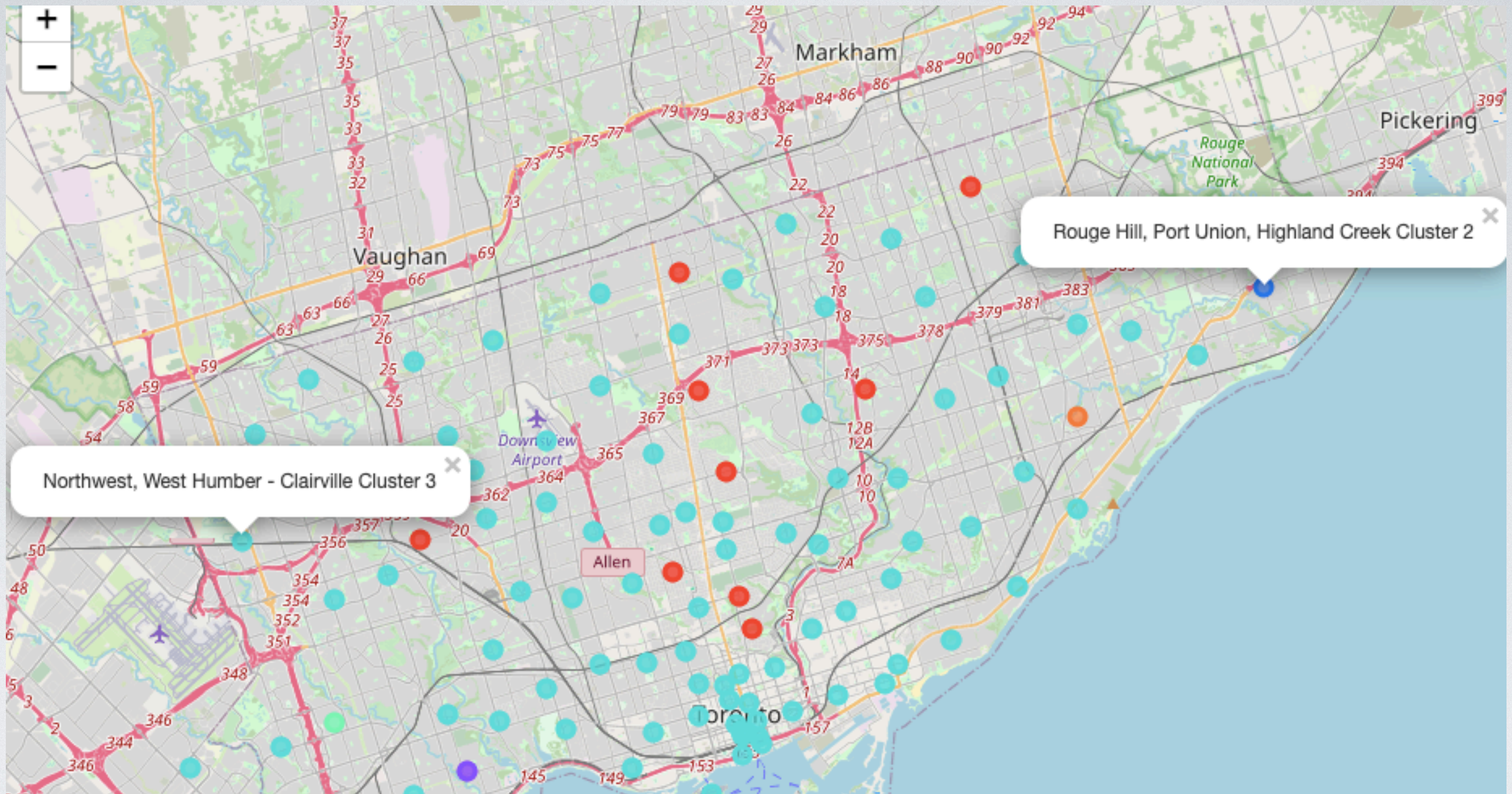
FIGURE 1: MAP SHOWING THE CLUSTERING OF THE NEIGHBORHOODS IN TORONTO

	Borough	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	9th Most Common Venue	10th Most Common Venue
94	Etobicoke	3	Bar	Rental Car Location	Drugstore	Truck Stop	Women's Store	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaura

TABLE 2: TOP 10 MOST COMMON VENUES FOR CLAIRVILLE

	Borough	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	9th Most Common Venue	10th Most Common Venue
12	Scarborough	2	Bar	Women's Store	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Farm

TABLE 3: TOP 10 MOST COMMON VENUES FOR HIGHLAND CREEK

RESULTS AND DISCUSSION

- Clairville in Etobicoke (Cluster 3) and Highland Creek in Scarborough (Cluster 2) belongs to separate clusters

- Advantages: Highland Creek has venues which addresses most of the needs of the resident, such as the following:

  - Dog runs are 6th most common, allowing the resident to send his dog outside to play and expend energy while he is out at work

  - Diners are 3rd most common with doner restaurants being 7th most common, providing the resident with options to eat out with choices for Turkish cuisine

  - Bars are the most common venue, thus the resident can continue to enjoy social drinking with friends and colleagues on a regular basis

RESULTS AND DISCUSSION

- Disadvantage: Highland Creek does not have rental car locations, which may make it difficult for to drive out for short vacations

- Alternative solution: To look at nearby neighborhoods.

    - West Hill in Scarborough is the nearest neighborhood from Highland Creek that also belongs to the same cluster as Clairville.

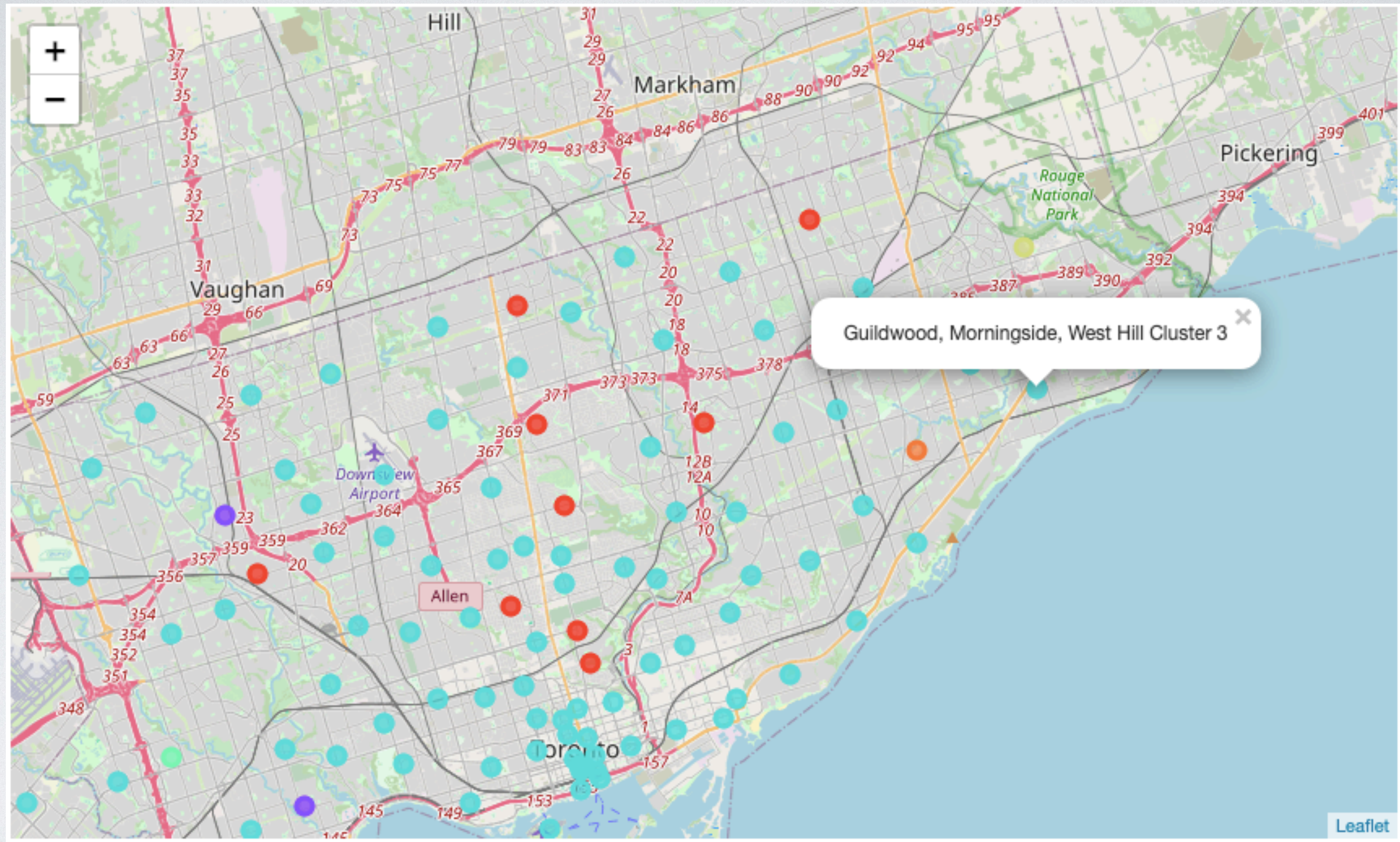   - The two neighborhoods are only approximately 3-4km apart

FIGURE 2: LOCATION OF WEST HILL IN SCARBOROUGH

	Borough	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	9th Most Common Venue	10th Most Common Venue
18	Scarborough	3	Bank	Electronics Store	Restaurant	Breakfast Spot	Medical Center	Rental Car Location	Mexican Restaurant	Women's Store	Discount Store	Distribution Center

TABLE 4: TOP 10 MOST COMMON VENUES FOR WEST HILL

RESULTS AND DISCUSSION

- Limitation of alternative:

  - While West Hill belongs to the same cluster as Clairville, top common venues only addresses one interest of the resident

  - More analyses would need to be performed on whether the other venues in West Hill other than those in the top 10 are sufficiently common in addressing the needs of the resident

CONCLUSION

- Recommendation: Resident should take up the job opportunity and to move to Highland Creek

  - Job opportunity is offered at Highland Creek

  - Top 10 most common venues in Highland Creek better satisfy the needs of the resident as compared to West Hill

  - Short distance between Highland Creek and West Hill allows the resident to have a wide range of rental car locations at his disposal at West Hill with short traveling time to meet his remaining needs