

The background of the slide features a series of concentric circles and curved lines in a light gray color, creating a subtle, abstract pattern. A large, solid red rectangle is positioned in the center, containing the title and subtitle text.

Toronto or Montreal: The Battle of the Neighborhoods

How Similar or Dissimilar are the cities' neighborhoods?

Introduction

- This project will focus on three types of potential newcomers and advise where to move or not move based on their characteristics.
 1. A couple who either is looking to start a family or already has young kids. This family will not want to live close to the nightlife scene but would appreciate parks and playgrounds.
 2. A young professional who's not a fan of their cooking; they want plenty of restaurant choices around them.
 3. An athlete with an active lifestyle. They'd prefer sports fields and courts and gyms.
- To capture each neighborhood's essence, the unsupervised machine learning algorithm k- means clustering was used, along with two values for k, five, and ten. The results will be calculated based on the most effective value.

The background of the slide features several sets of thin, curved lines in light gray, some solid and some dashed, creating a sense of motion or data flow. A large red speech bubble is positioned on the left side, containing the title text.

Data: Datasets

- Since the project focuses on the two cities' neighborhoods, the datasets weren't complicated or full of details. They included the postal code and neighborhood provided by Wikipedia pages [here](#) and [here](#). Furthermore, the Toronto longitude and latitude were available by the CSV file provided by IBM [here](#). While a defined function inserted the Montreal longitude and latitude via Geocoder.

The background of the slide features several sets of concentric, curved lines in a light gray color, some solid and some dashed, creating a sense of depth and movement. A large red speech bubble is positioned on the left side, containing the title text.

Data: Foursquare

- To capture the essence of each neighborhood, it's essential to focus on the most popular venues using Foursquare API.
- Parameters:
 - Limit: 100 venues.
 - Version: is the date. Capturing the neighborhoods' essence, it was critical to focus on before COVID-19.
 - Longitude and Latitude: based on the postal codes from datasets.
 - The radius: 500-meter radius

K-Means Clustering

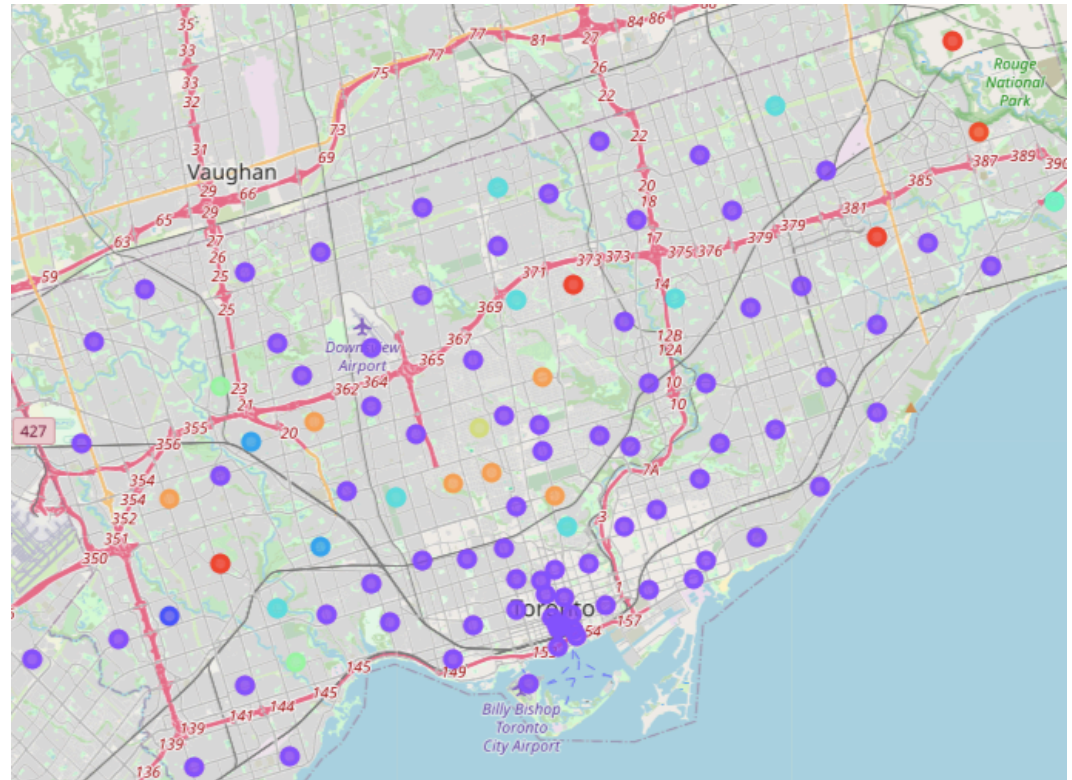
- K-means clustering is a type of portioning clustering. K-means can group data based on their similarities or dissimilarities into non-overlapping segments. Objects within the same cluster are similar, whereas objects across different clusters are dissimilar. Deciding the value for k is considered a hard problem because there's no clear-cut answer. In this case, we tested the values of k as five and ten. Comparing the results from $k = 5$ and $k = 10$, it was clear that $k = 10$ was more accurate as the clusters are more homogenous.

The background of the slide features several sets of curved, concentric lines in a light gray color, resembling a stylized map or a series of orbits. These lines are more prominent on the left side and fade out towards the right.

Toronto Clusters

- Moving from $k = 5$ to $k = 10$ was expected to modify the dominant cluster much more than it did. This indicates that the dominant cluster is more homogenous than initially thought. Furthermore, it was expected that the clusters would be along the lines of or geography. However, the further you get from downtown Toronto, the neighborhood's essence does not change as much as initially thought. This proves that Toronto neighborhoods' core isn't based on location at all.

Toronto Cluster Map when $k=10$

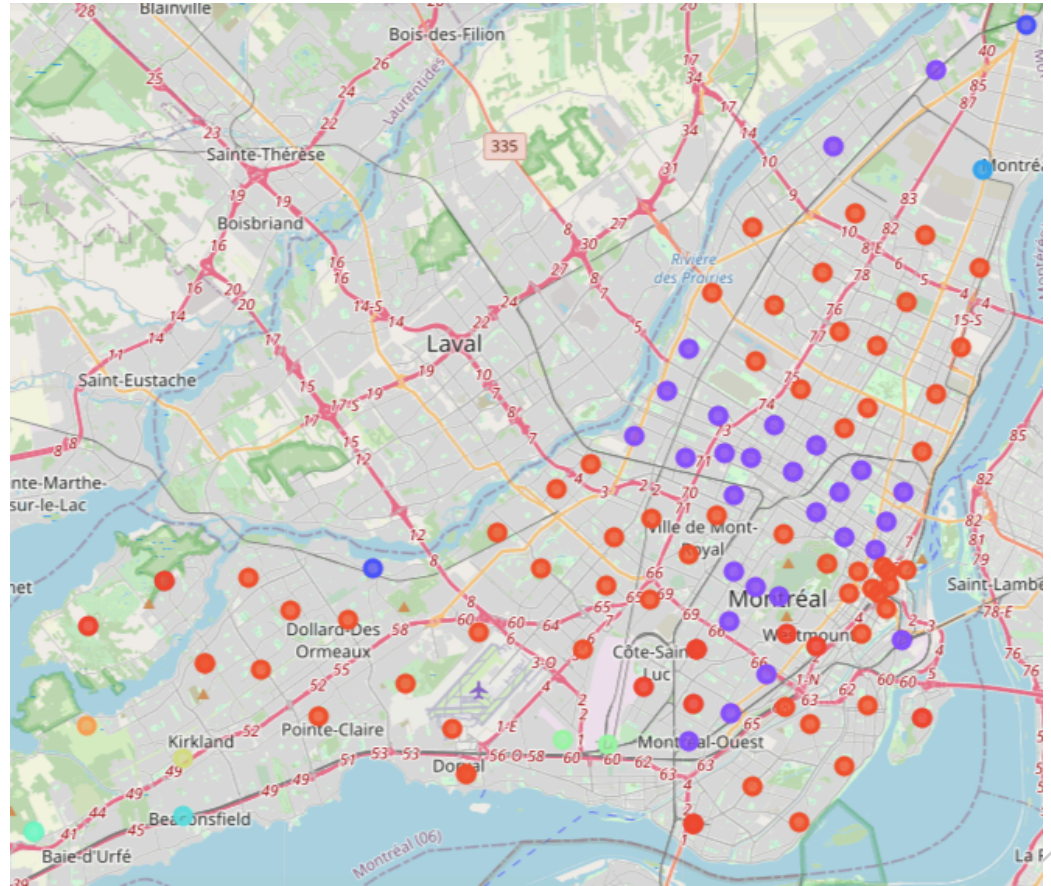


The background of the slide features several sets of thin, curved lines in light gray and white, some solid and some dashed, creating a sense of movement and depth. A large, solid red rectangle is positioned on the left side, containing the title text.

Montreal Clusters

- Unlike Toronto, Montreal changed a lot when $k = 10$. The red marks in figure 5 split into three groups, red markers, purple markers, and light red markers as predicted Montreal's clusters did partition one geographic line to a certain extent. The west island and east-end is heterogeneous compared to the rest of the island.
- Moreover, the second cluster (purple markers) breaks up the ninth (light red markers), by cutting through the island entirely, proving that is some segmentation based on geographical lines.

Montreal Cluster Map when $k = 10$





Results: The Young Family Neighborhoods

- What they want: a family-friendly neighborhood with parks and playgrounds for their children.
- Where to look: fifth and ninth cluster in Toronto.
 - The fifth cluster includes the following neighborhoods; Parkwoods, Caledonia-Fairbanks, Willowdale, Newtonbrook, York Mills West, Milliken, Agincourt North, Steeles East, Rosedale, The Kingsway, Montgomery Road, and Old Mill North.
 - The ninth cluster includes the following neighborhoods; Humewood-Cedarvale, North Park, Map Leaf Park, Upwood Park, Lawrence Park, Forest Hill North & West, Forest Hill Road Park, Kingsview Village, St. Phillips, Martin Grove, Moore Park, and Summerhill East.

The background of the slide features several sets of thin, curved lines in light gray, some solid and some dashed, creating a sense of motion or flow. A large red speech bubble is positioned on the left side, containing the title text.

Results: Young Professional

- What they want: a neighborhood with many restaurant choices.
- Where to look: Toronto's second cluster, such as Glencairn, Bedford Park, and Lawrence Manor East, all of which have their top four most common venues like restaurants.
- Montreal's neighborhoods in the tenth cluster, such as Lasalle, Saint Laurent, Montreal-Nord, and Dollard-des-Ormeaux. All of which have at least three of their top four top venues as some type of restaurant.



Results: The Athlete

- **What they want:** They're interested in sports courts and fields. They aren't interested in the nightlife scene.
- **Where to look:** a neighborhood in Toronto's seventh cluster; Humberlea, Emery, Old Mill South, King's Mill Park, and Sunnylea, which have a baseball field as their most common venue. They also have options in Montreal's first cluster, such as Nun's Island with a tennis court or Ile-Bizard with a hockey arena.

Conclusion

- To capture the essence of each neighborhood in the Canadian cities, unsupervised machine learning algorithm k-means clustering was used. Two values for k were tested, five and ten. Ultimately $k = 10$ was the most effective.
- The young family of newcomers should choose a neighborhood within the Toronto fifth and ninth clusters. The young profession who likes to eat out has the most choices in both cities as their options are in the largest clusters. The last newcomer is an athlete. Their lifestyle would fit Toronto's seventh cluster and Montreal's first cluster.

The background of the slide features several sets of curved, concentric lines in a light gray color, creating a sense of motion or a stylized globe. These lines are more prominent on the left and right sides of the slide.

Works Cited

- Cheatham, A. (2020, August 3). What Is Canada's Immigration Policy? Retrieved August 17, 2020, from <https://www.cfr.org/backgrounder/what-canadas-immigration-policy>
- List of postal codes of Canada: H. (2020, June 10). Retrieved August 17, 2020, from https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_H
- List of postal codes of Canada: M. (2020, June 10). Retrieved August 17, 2020, from https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_H