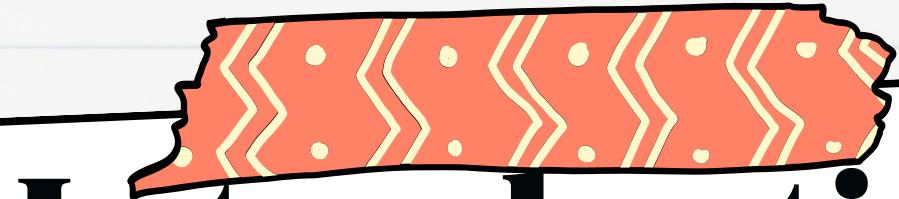


Capstone Project



Introduction

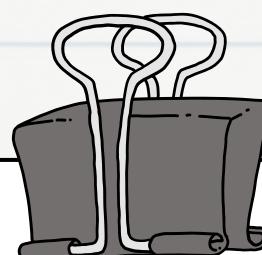
The primary objective of this project is to assist individuals in the thorough exploration of superior facilities within their vicinity, particularly in the Scarborough region of Toronto. The purpose is to empower individuals in making judicious and informed decisions when selecting an optimal neighborhood from a myriad of alternatives.

Given the increasing migration of people to various provinces in Canada, there arises a substantial need for comprehensive research concerning housing affordability and reputable educational institutions for families with children. This project is expressly designed to cater to the requirements of those individuals who are actively seeking neighborhoods with superior amenities. The focus encompasses facile access to essential services such as cafes, schools, supermarkets, medical and grocery stores, malls, theaters, hospitals, as well as fostering connections with like-minded individuals.

The overarching goal of this project is to furnish a nuanced analysis of pertinent features for individuals contemplating a move to Scarborough, facilitating a meticulous comparative examination of different neighborhoods. The considered features encompass an evaluation of median housing prices, educational institutions with commendable ratings, crime rates specific to each area, road infrastructure, prevailing weather conditions, efficacy of emergency management, availability of freshwater resources, sewage infrastructure for waste disposal, and recreational amenities.

By providing a comprehensive overview of these critical aspects, the project aspires to equip individuals with a heightened awareness of their prospective areas and neighborhoods before embarking on a relocation, be it for professional pursuits or to commence a new chapter in life.





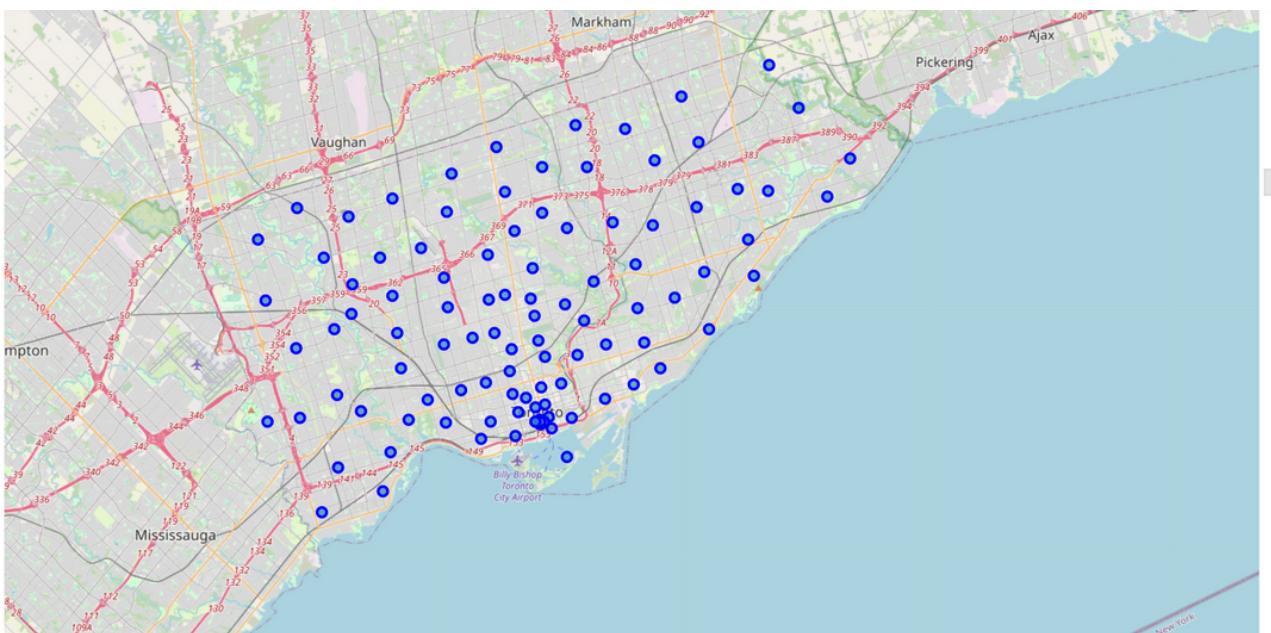
Dataset

use Scarborough dataset which we scrapped from wikipedia on Week 3.

Dataset: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

The information obtained per venue:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category



Methodology

Using K-Means Clustering Approach

```
▶ neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

Scarborough_merged =df_2.iloc[:16,:]

# merge toronto_grouped with toronto_data to add latitude/longitude for each neighborhood
Scarborough_merged = Scarborough_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'), on='Neighborhood')

Scarborough_merged.head()# check the last columns!
```

| | Postalcode | Borough | Neighborhood | 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 | 9th Most Common Venue |
|---|------------|-------------|--|-----------|------------|----------------|-----------------------|----------------------------|-----------------------|----------------------------|-----------------------------|-----------------------|-----------------------------|-----------------------|-----------------------------|
| 0 | M1B | Scarborough | Rouge, Malvern | 43.811525 | -79.195517 | 0 | Zoo Exhibit | Financial or Legal Service | Fast Food Restaurant | Construction & Landscaping | Fish & Chips Shop | Filipino Restaurant | Field | Fish Market | Farmers Market |
| 1 | M1C | Scarborough | Highland Creek, Rouge Hill, Port Union | 43.785665 | -79.158725 | 0 | Bar | Falafel Restaurant | Donut Shop | Dumpling Restaurant | Eastern European Restaurant | Electronics Store | Elementary School | Ethiopian Restaurant | Event Space |
| 2 | M1E | Scarborough | Guildwood, Morningside, West Hill | 43.765815 | -79.175193 | 2 | Park | Gym / Fitness Center | Pool | Fried Chicken Joint | Indian Restaurant | Athletics & Sports | Ethiopian Restaurant | Donut Shop | Dumpling Restaurant |
| 3 | M1G | Scarborough | Woburn | 43.768369 | -79.217590 | 0 | Coffee Shop | Fast Food Restaurant | Business Service | Park | Yoga Studio | Dumpling Restaurant | Eastern European Restaurant | Electronics Store | Elementary School |
| 4 | M1H | Scarborough | Cedarae | 43.769688 | -79.239440 | 0 | Flower Shop | Athletics & Sports | Thai Restaurant | Bank | Bakery | Caribbean Restaurant | Hakka Restaurant | Indian Restaurant | Eastern European Restaurant |



Methodology

Most Common venues near Neighborhood

```
indicators = ['st', 'nd', 'rd']

columns = ['Neighborhood']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = Scarborough_grouped['Neighborhood']

for ind in np.arange(Scarborough_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(Scarborough_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

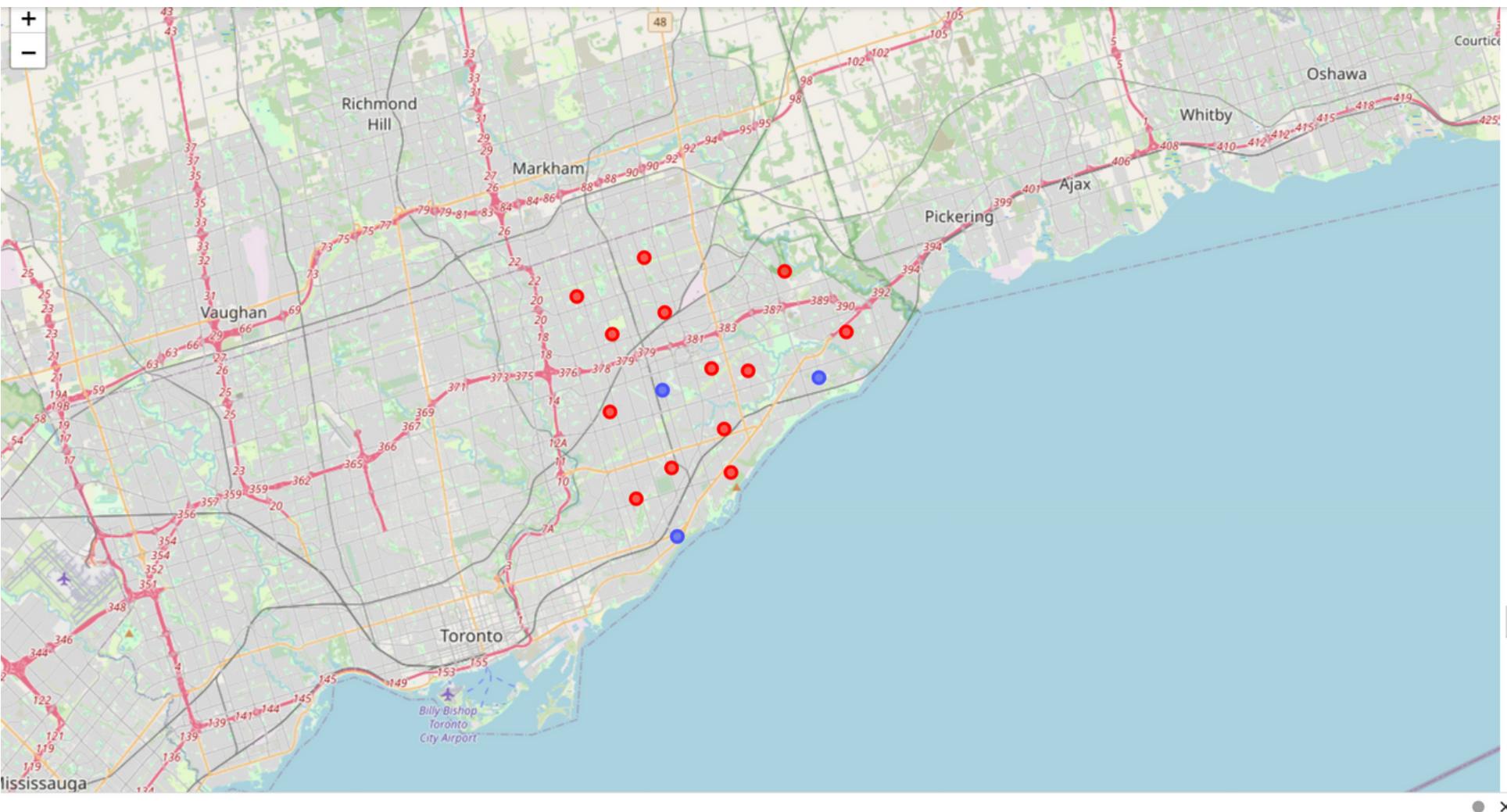
| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------------|--------------------------|-----------------------|------------------------|
| 0 | Adelaide, King, Richmond | Coffee Shop | Café | Hotel | Gastropub | Burger Joint | Asian Restaurant | Bar | Restaurant | American Restaurant | Steakhouse |
| 1 | Agincourt | Chinese Restaurant | Shopping Mall | Pizza Place | Supermarket | Sushi Restaurant | Breakfast Spot | Print Shop | Mediterranean Restaurant | Coffee Shop | Pool |
| 2 | Agincourt North, L'Amoreaux East, Milliken, St... | Pharmacy | Sandwich Place | Sushi Restaurant | Doner Restaurant | Donut Shop | Dumpling Restaurant | Eastern European Restaurant | Electronics Store | Elementary School | Ethiopian Restaurant |
| 3 | Albion Gardens, Beaumont Heights, Humbergate, ... | Grocery Store | Park | Sandwich Place | Discount Store | Japanese Restaurant | Fried Chicken Joint | Beer Store | Hardware Store | Pizza Place | Fast Food Restaurant |
| 4 | Alderwood, Long Branch | Convenience Store | Pub | Sandwich Place | Coffee Shop | Gas Station | Dance Studio | Gym | Pharmacy | Pizza Place | Falafel Restaurant |

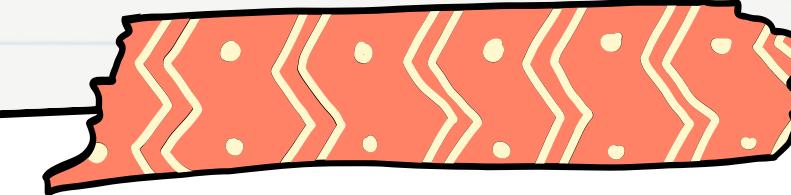




Results Section

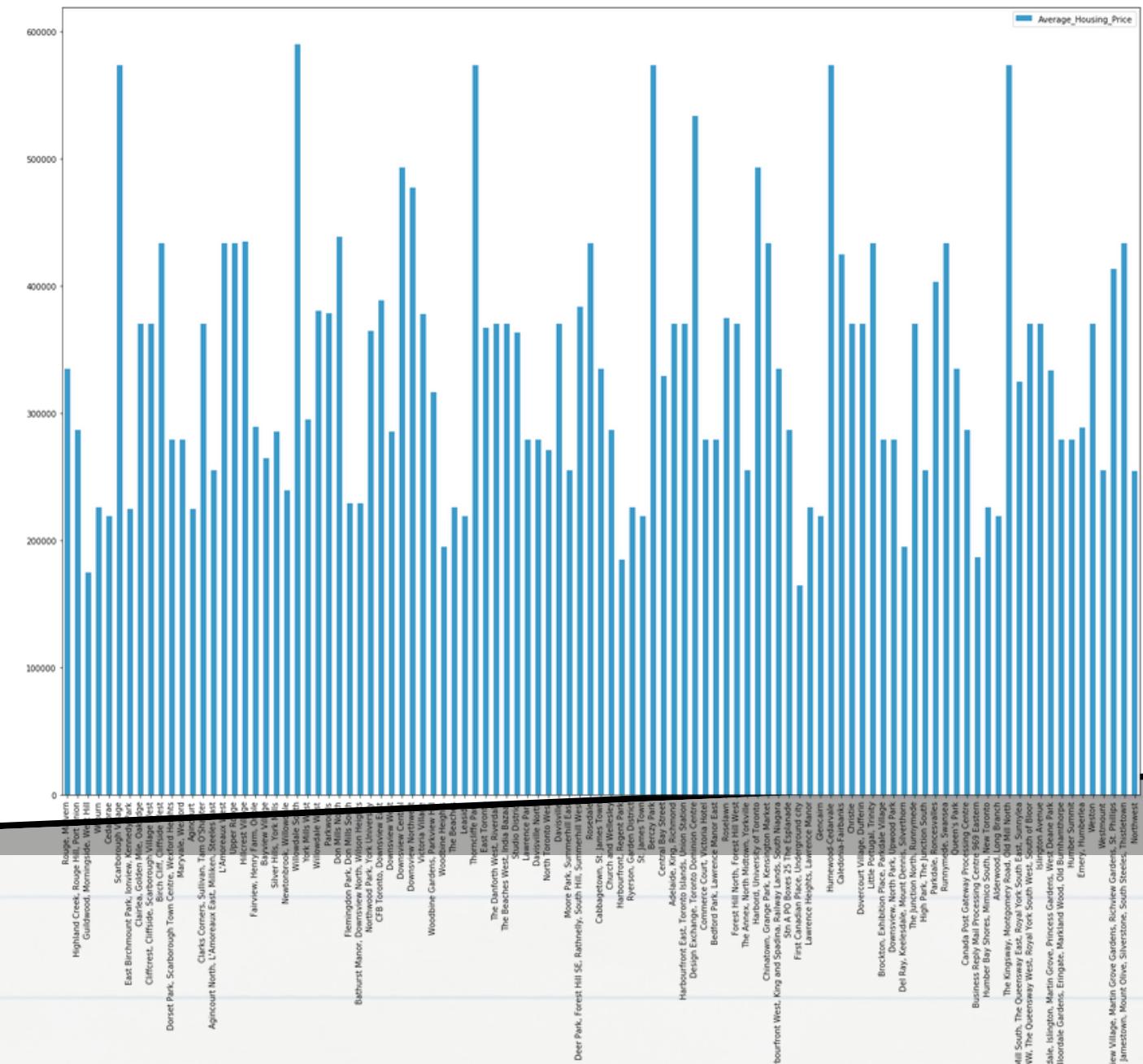
Map of Clusters in Scarborough

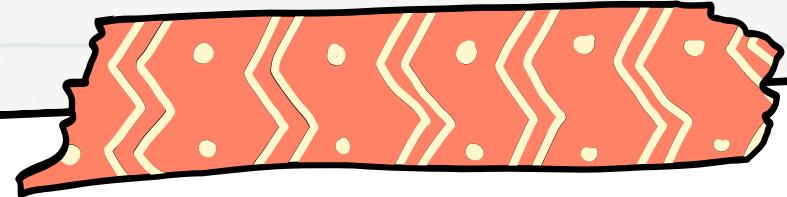




Results Section

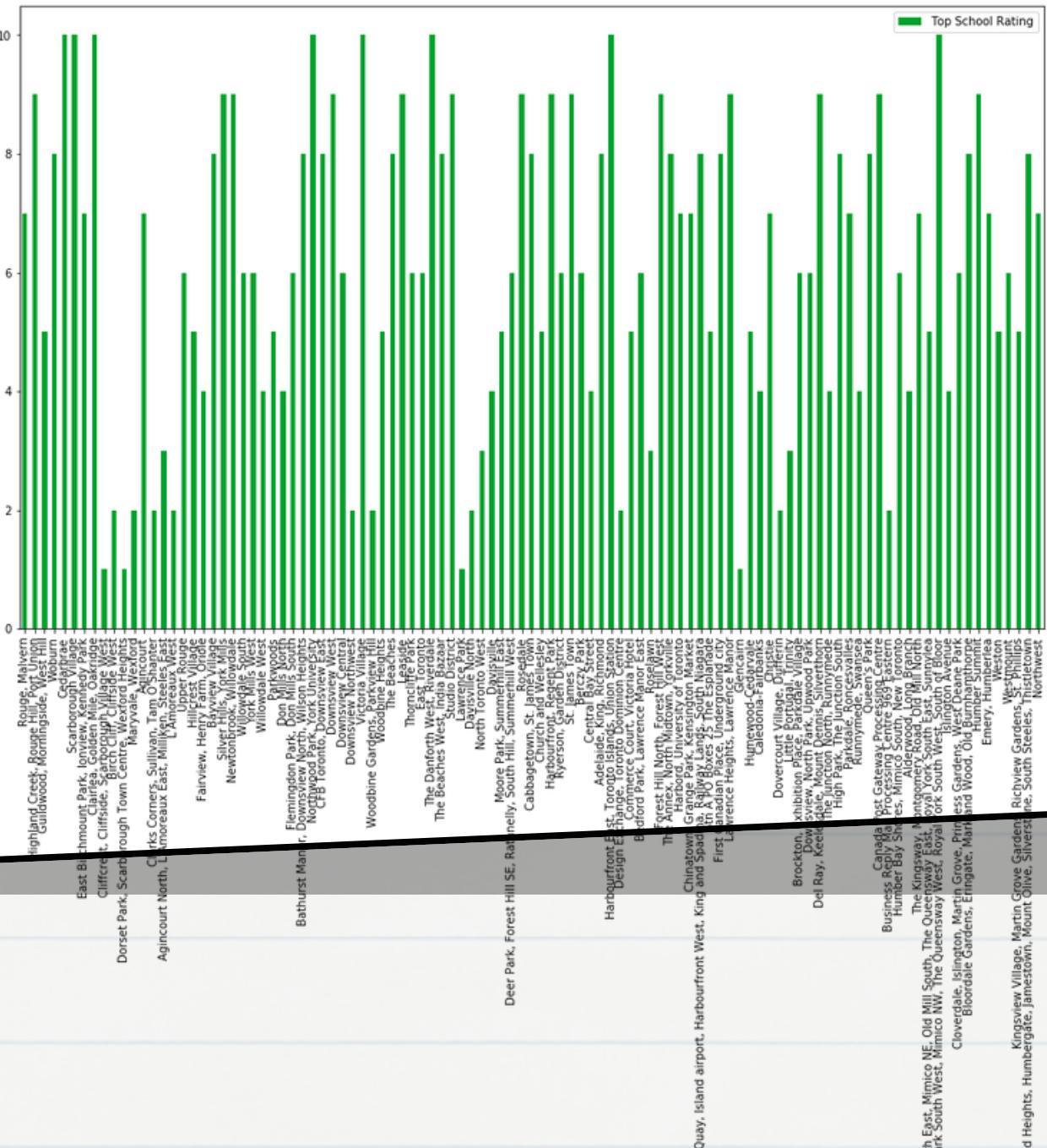
Average Housing Price by Clusters in Scarborough





Results Section

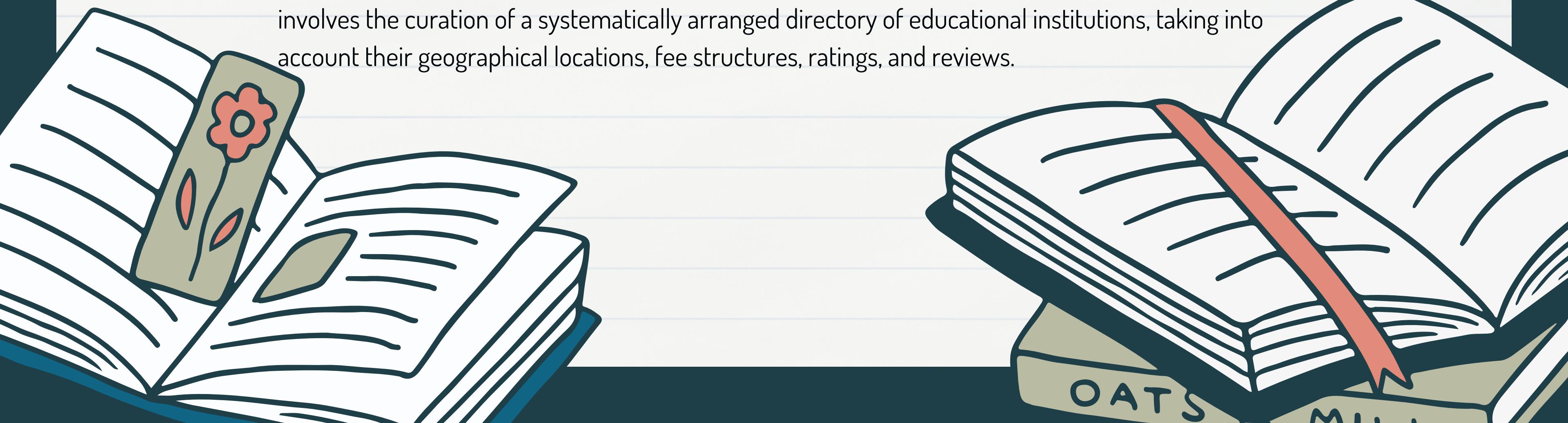
School Ratings by Clusters in Scarborough



Discussion

The primary objective of this undertaking is to propose an optimal residential locale within a novel urban setting for individuals in the process of relocating. Considerations encompass the societal integration through the presence of congenial individuals with shared interests. Furthermore, the project emphasizes the importance of seamless accessibility to critical infrastructure such as airports, bus terminals, city centers, markets, and proximate amenities that cater to daily needs.

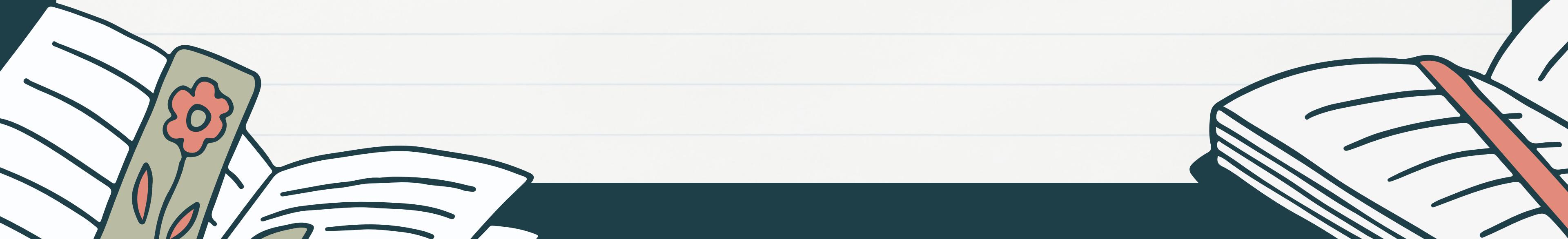
The project entails the compilation of a meticulously organized inventory of residences, categorized based on housing prices in either an ascending or descending order. Additionally, the endeavor involves the curation of a systematically arranged directory of educational institutions, taking into account their geographical locations, fee structures, ratings, and reviews.



Conclusion

In this undertaking, the k-means clustering algorithm was employed to partition the neighborhoods into ten discrete clusters. Utilizing latitude and longitude coordinates from a dataset encompassing 103 distinct locations, neighborhoods exhibiting pronounced similarities were identified. The ensuing visual representations encapsulate a comprehensive portrayal of specific neighborhoods, taking into account both average housing prices and school ratings.

The sense of fulfillment derived from these endeavors is substantial, affirming the conviction that this course, encompassing a broad spectrum of topics, is deserving of commendation. This project has not only served as a practical manifestation but has also demonstrated the efficacy of leveraging Data Science tools to address genuine situations with discernible personal and financial implications. The integration of Folium for mapping purposes emerges as a potent technique, enhancing the consolidation of information and facilitating more confident analysis and decision-making.



Libraries

Pandas

Folium

Scikit Learn

JSON

XML

Geocoder

Beautiful Soup and Requests

Matplotlib

