COMPARING THE NEIGHBORHOOD OF TORANTO AND NEW YORK

DEEPSHIKA JANARTHANAN

10 JUNE 2021

INTRODUCTION

To compare the neighborhood clusters of the Toronto and the new York city to find the similarities and dissimilarities.

METHODOLOGY

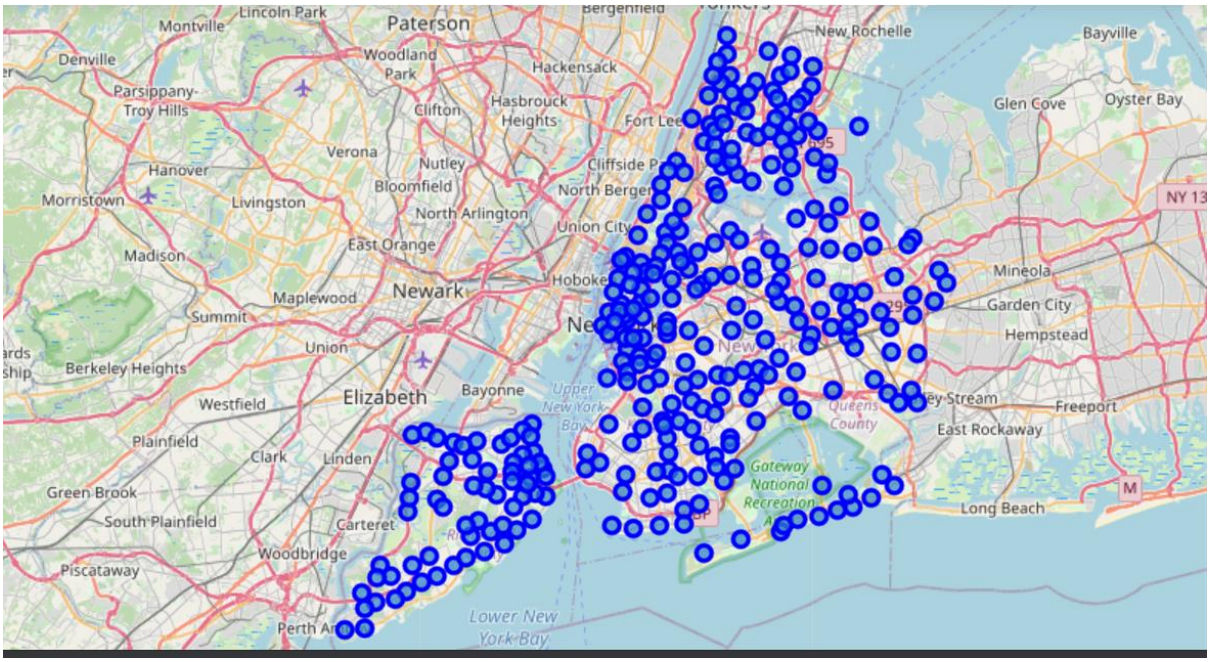
DATA

The data is extracted from the URL links using the requests and the beautiful soup libraries. Then the neighborhood, borough and the postal code fields are extracted from the json data and stored in the respective variables. The geo spatial data is then extracted which contains the latitude and the longitude of the data along with the postal code.

The data frames are merged on the postal code column the folium map is created and the latitudes and the longitudes are mapped.

The New York data is also extracted from the URL and the latitudes and the longitudes are mapped.

The clustering of neighborhood in Toronto

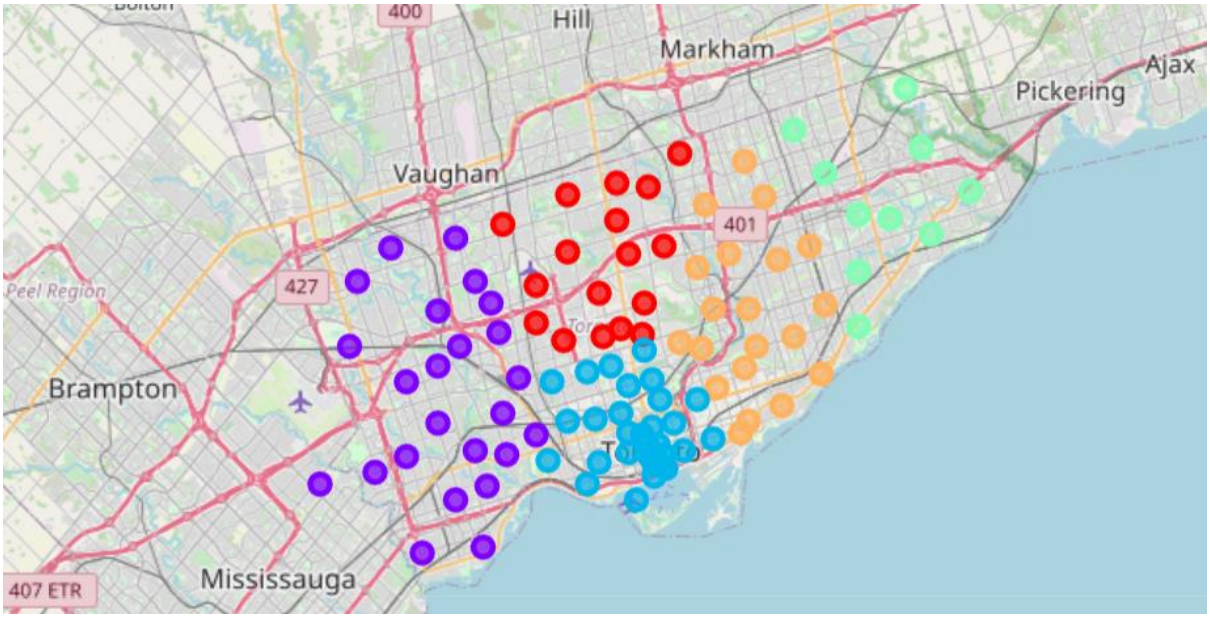


Clustering of the neighborhood of New York city

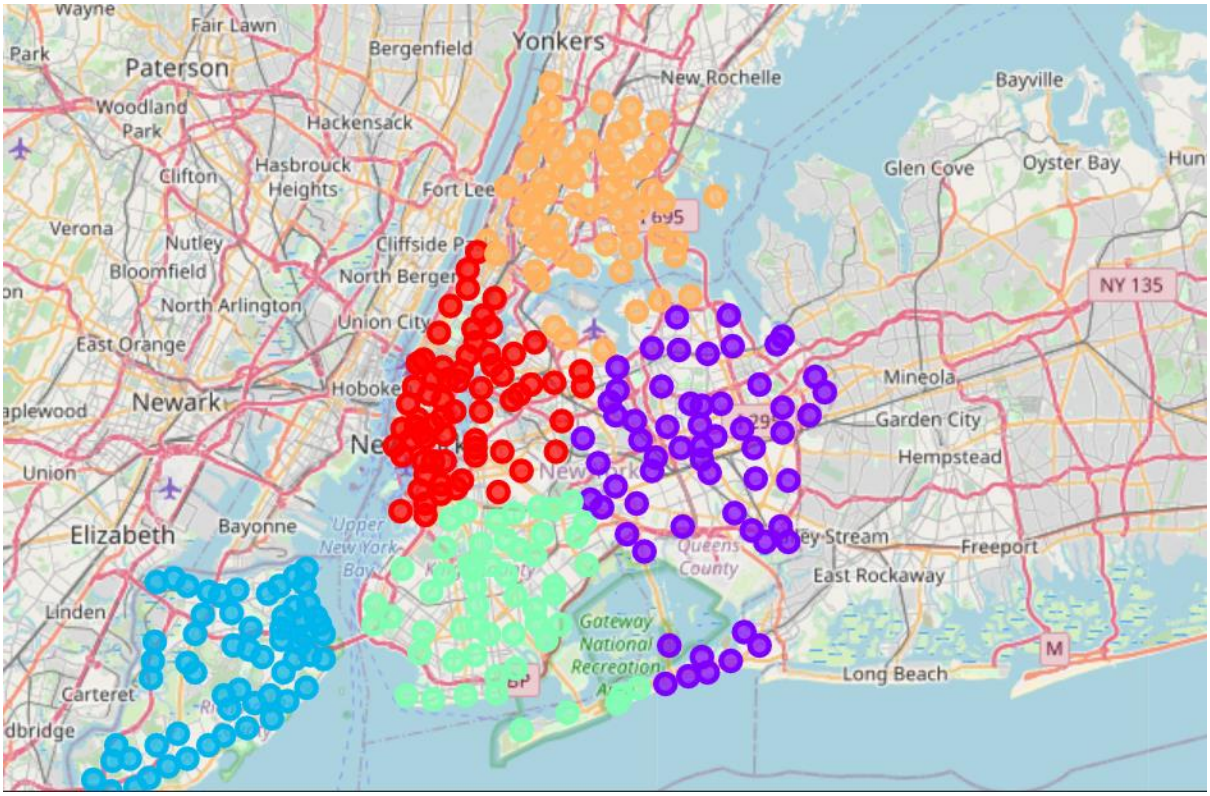
We could see that the New York city has more neighborhood and more strong cluster pattern when compared with the Toronto.

CLUSTERING MODEL

The latitudes and the longitudes are then clusters using the K means clustering algorithm. The k means algorithm Is the clustering algorithm which groups the nearest data points with the help of the centriods. The main objective of this algorithm is to minimize the intra distance in the clusters and to increase the inter distance between the clusters.



After grouping the clusters of Toronto using the k means algorithm the clusters will look as this.



We could clearly see that the algorithm is trying to minimize the intra distance and maximize the inter distance between the clusters.

RESULTS

The problem statement was to find if the New York city and the Toronto has any similarities between them in terms of the neighbourhood clusters. The results of the map defines that the both the cities are unique of their kind and do not have much similarity.

DISCUSSION

May be further analysis can be carried out comparing the other cities with the New York to find some other cities can match the clustering pattern of it.

CONCLUSION

The final insights of this experiments is that the New York city has more number of neighborhood and groupings as compared to the Toronto city and they both are dissimilar from each other.