The Battle of Neighborhoods

IBM Data Science Professional Certificate - Capstone Project

Background

We hypothesize to be a Start Up which has to open his headquarters either in Toronto or in New York. We want to ensure that our employers have the higher standard of life possible. Therefore, we want to assure that our employers live and work in a place which has enough places and facilities for everyone's taste.

Problem Statement

❖ Understand the similarities and differences between Toronto and New York

❖ Find the best Neighborhood for a Start Up which wants to move its headquarters based on venues

Objective

❖ Collect the neighborhoods top venues using Foursquare

❖ From neighborhood based on venue categories using unsupervised k-means clustering algorithm

❖ identify and understand the similarities and differences between Downtown Toronto and Manhattan neighborhoods to obtain insights

Data Description

To Analyze the Neighborhoods in Toronto, we will use the Wikipedia page utilized in past week's assignment and merge it with the location data (latitude-longitude) provided in previous weeks.

In order to segment and cluster Neighborhoods in New York, we'll use the data provided by the New York University's open data site.

Then, Foursquare API search feature will be used in order to collect neighborhood venue information, providing insights about the quality of a neighborhood.

Limitations

❖ This project doesn't consider employers with family and children, which may find more important have facilities such as schools, kindergartens or sport associations. However, we are considering a Start Up, which probably has employers aged between 19 and 30, more interested in meeting points.

❖ Although cost of living in Toronto and New York is approximately the same, this project doesn't specifically consider it in the analysis. This cost could potentially affect the quality of living for employers.

Compariso between Neighborhoods

❖ Downtown Toronto has 75 distinct venues in 205 categories

❖ the most common venues are bar, coffee shops and restaurants

❖ Manhattan has 2287 distinct venues in 341 unic categories

❖ in venues, Manhattan has more gyms, parks, theaters, so it has more variety and choice

Conclusion

❖ Besed on the quantity and variety of venues Manhattan is a better choice.