Comparison of the Neighborhoods of Manhattan and Toronto with the "Top 19 Hippest Mid-Size US Cities"

IBM Data Science Capstone Project

Justin Fitch - June 27, 2021

Introduction

De-Urbanization During Covid-19

- Many jobs moving towards remote work
- Less incentive to live in expensive crowded large cities
- Many looking to relocate to a slower-paced and more affordable small- to mid-size city
- trip.com: “Top 19 Hippest Mid-Size US Cites”
- Cluster analysis on Manhattan/Toronto neighborhoods and 19 cities from list
- Cluster neighborhoods/cities on FourSquare dining/entertainment venue data

Data

trip.com and Wikipedia Web Scraping

- Extracted Manhattan/Toronto neighborhoods and list of hip mid-size cities via web-scraping with BeautifulSoup 4 for Python
- Created Pandas data frames with lat/long of each neighborhood, and the list of cities
- Cycled through list of neighborhoods/cities, sending get requests for each locale to the FourSquare API

Methodology

FourSquare API

- Get requests were sent to the FourSquare API to obtain venue data for each locale
- Results were returned via a json file, listing venues, locations, and types
- The json files were parsed, and the venue types were extracted to create a dataframe indicating the proportion of venue types for each neighborhood/city
- A k-means cluster analysis was performed on this dataset, using the sk-learn k-means function.
- All neighborhoods and mid-size cities were assigned to one of 10 clusters

Results

10 Uneven Clusters

- Each neighborhood or city was placed into one of 10 clusters
- 18 or 19 mid-size cites, as well as almost all Manhattan suburbs, were placed in Cluster 8
- The analysis found much more variety within the neighborhoods of Toronto than those of Manhattan or the mid-size cities
- Multiple clusters were represented within the Toronto neighborhoods
- The only mid-size city not placed in Cluster 8 was Vancouver, WA

Discussion

More options for Manhattan Ex-Urbanites

- Results show that move from most Manhattan neighborhoods to one of America's hippest mid-size cities should not provide too much of a culture shock
- Analysis found that the neighborhoods of Toronto are much more diverse than those of Manhattan
- The neighborhoods of Toronto also have less in common with the analyzed mid-size US cities than those of Manhattan
- Vancouver, WA was the only mid-size city NOT placed in Cluster 8…Proximity to Canada?
- 100 venue limit from FourSquare could be a severely limiting factor to this analysis, especially when looking at entire mid-size cities; 100 venues is hardly representative

Conclusions

Opportunities for Further Investigation

- Results show that the neighborhoods of Toronto are more diverse than those of Manhattan
- Neighborhoods of Manhattan have more in common with “hip” mid-size US cities than those of Toronto
- Further analysis might include mid-size Canadian cities, as the cultural differences between the US and Canada may contribute to these results
- A paid FourSquare developer license would allow for extraction of larger datasets, and possibly significantly alter the results of this study