Comparison of the Neighborhoods of Manhattan and Toronto with the “Top 19 Hippest mid-size US cities”

Introduction

With many jobs moving towards remote work during the pandemic, many urbanites have been reconsidering life in the big metropolises of the world and mulling over the idea of relocating to a smaller, less crowded and less expensive city.¶

Trip.com recently published a list of the the "Top 19 Hippest Mid-Size US Cities". This list is a good place to start for looking for mid-size cities that might offer a similar diverse and upscale lifestyle to big cities like New York and Toronto.

This analysis will examine both he various boroughs/neighborhoods of Manhattan and Toronto and compare the the availability and diversity of restaurants and entertainment venues with those of the mid-size cities on the [trip.com](http://trip.com) list. The result will allow anyone in Manhattan or Toronto to identify which, if any of the 19 mid-size cities in the analysis would provide a similar food/entertainment experience to their current home.

Data

In previous lab exercises, we clustered neighborhoods in Manhattan and Toronto to determine which areas of each city provide a similar mix of venues/businesses. The cluster analysis was performed on a list of venue types for each locale. Using the same FourSquare data for the 19 cities on the [trip.com](http://trip.com) “Hippest” list, a k-means cluster analysis was performed to determine which, if any, mid-size US cities would provide a comparable mix of food/entertainment venues for those wishing to relocate from the various boroughs of the larger cities.¶

Analyzing individual boroughs makes sense in NYC where, even with a robust public transportation system, it take a long time to get other parts of town. Mid-sized cities, on the other hand, typically have less traffic and are not as spread out geographically. You can often get from one side of town to the other in a half-hour or less. Additionally, the relative abundance of free parking make driving to the venue of your choice much more practical than in most big cities.¶

For this reason, city-wide FourSquare venue queries were run on the mid-sized cities, in comparison with neighborhood-by-neighborhood queries in Manhattan and Toronto. This allowed the identification of cities that offer similar mixes of entertainment and dining venues as the various boroughs we have already identified in the two large cities.

Methodology

The selection of mid-sized for this analysis cities was based on an article from [trip.com](http://trip.com) which included a list of the “Top 19 Hippest Mid-Size US Cities”. The first step in the process therefore was to use web-scraping techniques to import the list of cities into the project. A get command was run to download the target page, and then Beautiful Soup 4 was used to parse the HTML and strip the data from the table of cities. This list of cities was then entered into a Pandas data frame.

Get Requests were then sent to the FourSquare API for each city in the list, returning a json file containing 100 venues for each city, which is the maximum number of venues that can be retrieved in a single call with the license in use. The json file was then parsed and venue categories were extracted and sent to a Pandas data frame, along with venue names and their associated city.

A data frame listing the proportionate representation of each venue category in each city was constructed, matching those which were previously prepared for the neighborhood of Manhattan and Toronto. The Manhattan and Toronto data was exported from previous labs as json files, and imported into this project as Panda data frames. The three resulting data frames (Manhattan, Toronto, Mid-Size Cities) were then concatenated into a single data frame and all NaN values (resulting when venue categories present in one file were not present in another) were converted to zeros.

A k-means closer analysis was then performed on the data using the Scikitlearn library. 10 clusters were created and all neighborhoods and mid-size cities were assigned to a cluster.

Results

The k-means analysis resulted in the assignment of each neighborhood and mid-size city to one of 10 clusters. While each of the 10 clusters was represented by at least one neighborhood/city, the classification was widely skewed, with 97 locations placed in Cluster 8. The majority of the diversity represented by the cluster assignments was seen in the neighborhoods of Toronto, with most of the Manhattan neighborhoods and all but one of the mid-size cities (Vancouver, WA) falling into Cluster 8.

Discussion

The results of this analysis show that a move from most any neighborhood in Manhattan to one of Americas’s “19 Hippest” mid-size cities would not be a huge culture shock. Based on the cluster analysis of the venue data returned by FourSquare, there is a similar mix of restaurants and entertainment venues in 18 of the 19 cities as is present in most of the boroughs of Manhattan.

Toronto is another story, however. Not only did the analysis find much more diversity in the neighborhoods of Toronto than those of Manhattan or even across the mid-size cities, but that diversity was enough to result in most of Manhattan and 18 of the 19 mid-size cities to be placed in a single cluster. This suggests that a relocation from almost anywhere in Toronto to a mid-size US city, no matter how “hip” it may be, could come with a bit of a culture shock.

The only mid-size city in the analysis not to fall into Cluster 8 was Vancouver, Wa. Vancouver, Wa was placed in Cluster 3, which also included multiple neighborhoods in Toronto. Could this be a result of the proximity of Vancouver, Wa to Canada and a bleeding over of Canadian culture into the Pacific Northwest?

One limitation of this analysis was the 100 venue limit when making a request to the FourSquare API. When analyzing a mid-size city like Santa Rosa, CA or Albuquerque, NM, 100 venues is likely not representative of the full diversity and culinary/cultural mix of the city. Access to a larger dataset may aid in further dividing the mid-size cities and neighborhoods of Manhattan that were lumped into Cluster 8.

Conclusions

The surge in work-from-home employment during the Covid-19 pandemic has resulted in large numbers of workers leaving expensive large cities to set down roots in more slow-paced and affordable smaller cities. This move, of course, comes with trade-offs, as lower population densities generally result in reduced quantities and varieties of entertainment and dining venues.

The result of this study show that there are a number of mid-size cities that could potentially provide a similar lifestyle as the various boroughs of Manhattan. This k-means cluster analysis placed 18 of the 19 cities on the [trip.com](http://trip.com) list of the “Top 19 Hippest Mid-Size US Cities” in a shared cluster with the majority of the Manhattan neighborhoods included in the analysis.

On the other hand, the analysis showed significant differences between many of the boroughs of Toronto and the mid-size US cities. Perhaps this is to be expected, because as close as the USA and Canada are, there are significant cultural differences between the two nations. Perhaps a similar analysis with mid-size Canadian cities would demonstrate similar congruencies between the neighborhoods of Toronto and various “hip” mid-size Canadian cities.