



Capstone Project Report

Comparison of neighborhoods in major cities of United States and Canada

28 Feb 2020

By : Yogesh Kuvelkar



Comparison of neighborhoods in major cities of United States and Canada

- Background

United States of America is a great place for the people across the world to dream visiting or getting a job or starting their own business and eventually settling down. I think there are many resources available that differentiate the various major cities in USA / Canada in terms of the weather, population, demographics etc. Such readily available resources help aspirants planning to move to USA to decide their cities of preference based on weather, demographics, etc. I think what can help is a readily available resource that also describes how the major cities are similar or dissimilar comparing the neighborhoods in each city.

- Problem

Such a resource will require access to data of various neighborhoods and the venues in each neighborhood. Once the data is accessible, it requires analysis and co-relation of various venues in various cities. The project aims to study the neighborhoods of major cities in United States and Canada and the venues around these cities and identify the pattern or similarity or dissimilarity of venues that are observed in cities.

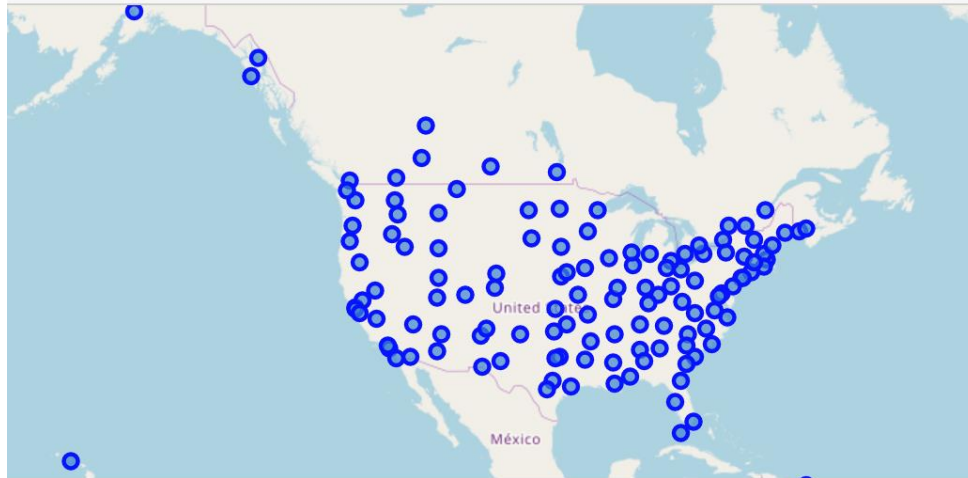
- Interest

Such a resource will find interest from visitors, entrepreneurs, job-seekers and immigrants who are either new or are planning to move to United States/Canada. For individuals planning to start a restaurant business, this data will provide insights on in which city their cuisine of interest will flourish. For individuals looking for a job, this data will provide insights on neighborhoods where they can reside and grow their family.



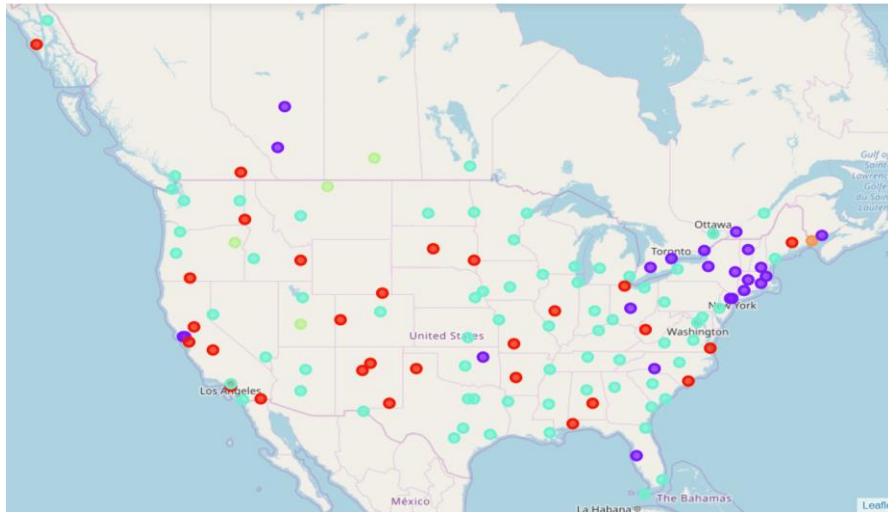
Dataset used

- A list of major cities in United States and Canada was used and using the location coordinates of the cities the same were plotted on the map as below. The for each of the city, FourSquare “Explore” API was called to retrieve the neighborhood information, including the venue name, venue category, average ratings, and location.



Predictive Modelling

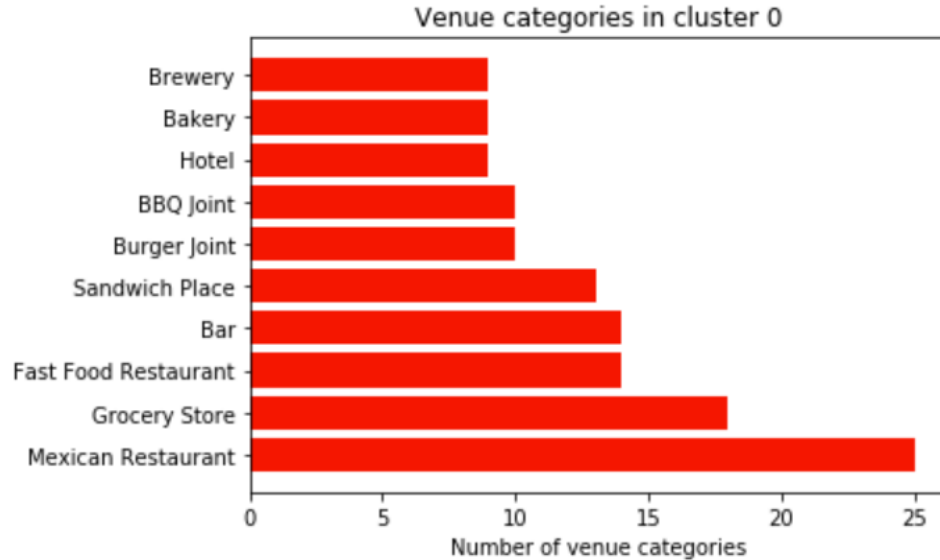
Clustering machine learning technique was used to determine which cities are similar or dissimilar to each other.



The different colored dots represent these clusters:

- Cluster 0
- Cluster 1
- Cluster 2
- Cluster 3
- Cluster 4
- Cluster 5

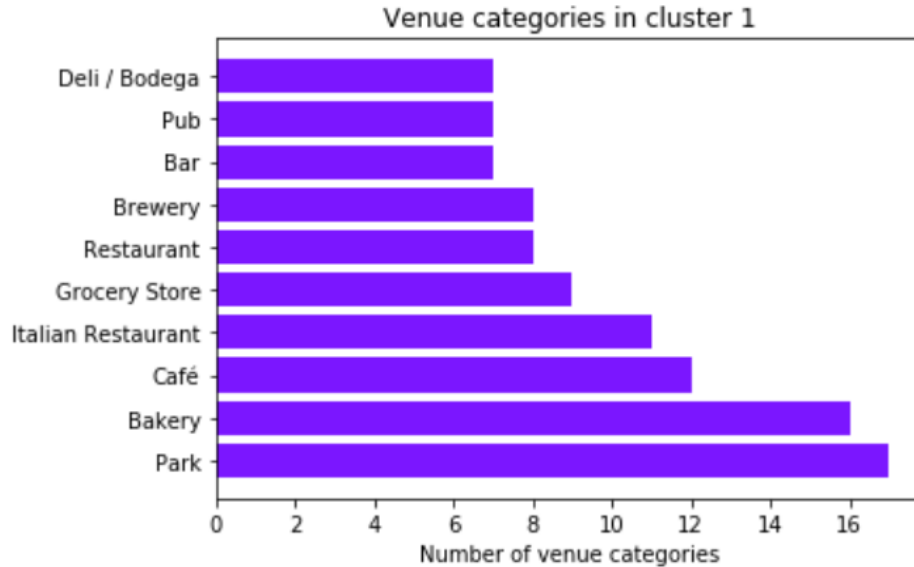
City Neighbourhood Similarities



['Albuquerque, N.M.',	'Amarillo, Tex.',
'Bangor, Maine',	'Carlsbad, N.M.',
'Charleston, W. Va.',	'Cheyenne, Wyo.',
'El Centro, Calif.',	'Fresno, Calif.',
'Grand Junction, Colo.',	'Hot Springs, Ark.',
'Idaho Falls, Idaho',	'Klamath Falls, Ore.',
'Lewiston, Idaho',	'Long Beach, Calif.',
'Mobile, Ala.',	'Montgomery, Ala.',
'Nelson, B.C., Can.',	'Pierre, S.D.',
'Sacramento, Calif.',	'San Jose, Calif.',
'Santa Fe, N.M.',	'Sioux Falls, S.D.',
'Sitka, Alaska',	'Springfield, Ill.',
'Springfield, Mo.',	'Toledo, Ohio',
'Virginia Beach, Va.',	'Wilmington, N.C.']

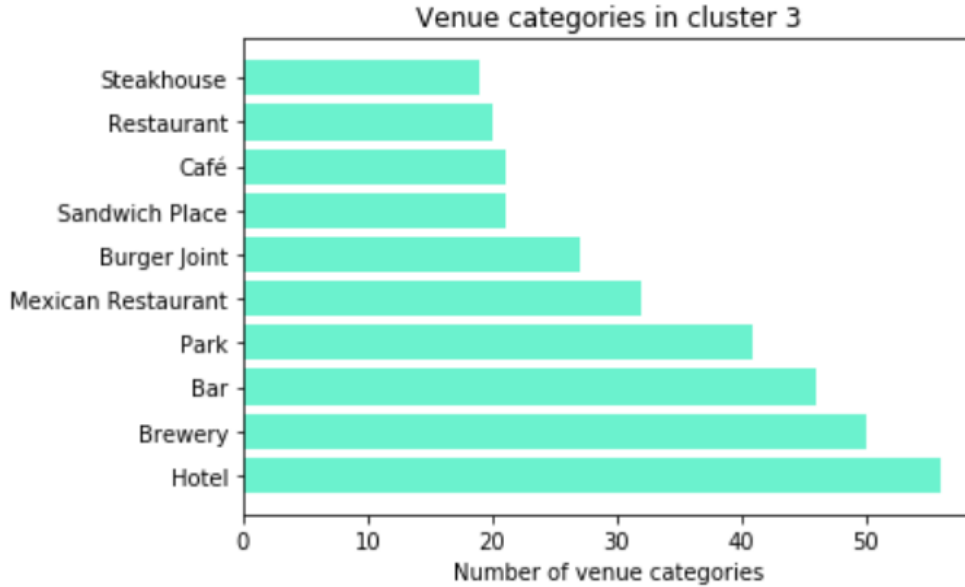
:

City Neighbourhood Similarities



['Albany, N.Y.',
'Calgary, Alba., Can.',
'Columbus, Ohio',
'Kingston, Ont., Can.',
'Manchester, N.H.',
'Montreal, Que., Can.',
'New York, N.Y.',
'Oakland, Calif.',
'San Francisco, Calif.',
'St. John, N.B., Can.',
'Tampa, Fla.',
'Tulsa, Okla.'],
'Boston, Mass.',
'Charlotte, N.C.',
'Edmonton, Alb., Can.',
'London, Ont., Can.',
'Montpelier, Vt.',
'New Haven, Conn.',
'Newark, N.J.',
'Providence, R.I.',
'Springfield, Mass.',
'Syracuse, N.Y.',
'Toronto, Ont., Can.'],

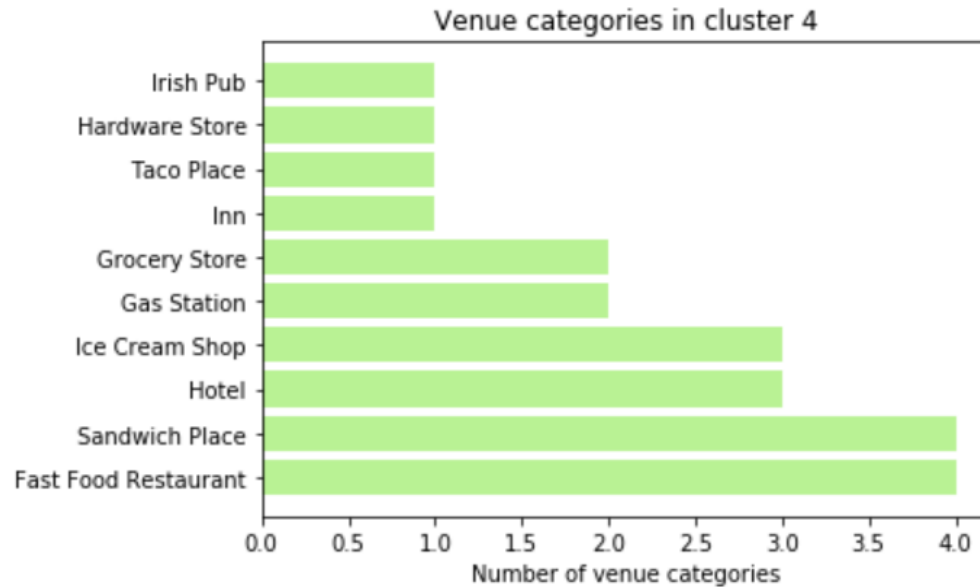
City Neighbourhood Similarities



Restricted - External

'Anchorage, Alaska',	'Atlanta, Ga.',	'Austin, Tex.',
'Baltimore, Md.',	'Birmingham, Ala.',	'Bismarck, N.D.',
'Boise, Idaho',	'Buffalo, N.Y.',	'Charleston, S.C.',
'Chicago, Ill.',	'Cincinnati, Ohio',	'Cleveland, Ohio',
'Columbia, S.C.',	'Dallas, Tex.',	'Denver, Colo.',
'Des Moines, Iowa',	'Detroit, Mich.',	'Dubuque, Iowa',
'Duluth, Minn.',	'El Paso, Tex.',	'Eugene, Ore.',
' Fargo, N.D.',	'Flagstaff, Ariz.',	'Fort Worth, Tex.',
'Grand Rapids, Mich.',	'Helena, Mont.',	'Honolulu, Hawaii',
'Houston, Tex.',	'Indianapolis, Ind.',	'Jackson, Miss.',
'Jacksonville, Fla.',	'Juneau, Alaska',	'Kansas City, Mo.',
'Key West, Fla.',	'Knoxville, Tenn.',	'Las Vegas, Nev.',
'Lincoln, Neb.',	'Los Angeles, Calif.',	'Louisville, Ky.',
'Memphis, Tenn.',	'Miami, Fla.',	'Milwaukee, Wis.',
'Minneapolis, Minn.',	'Nashville, Tenn.',	'New Orleans, La.',
'Oklahoma City, Okla.',	'Omaha, Neb.',	'Ottawa, Ont., Can.',
'Philadelphia, Pa.',	'Phoenix, Ariz.',	'Pittsburgh, Pa.',
'Portland, Maine',	'Portland, Ore.',	'Quebec, Que., Can.',
'Raleigh, N.C.',	'Reno, Nev.',	'Richmond, Va.',
'Roanoke, Va.',	'Salt Lake City, Utah',	'San Antonio, Tex.',
'San Diego, Calif.',	'San Juan, P.R.',	'Savannah, Ga.',
'Seattle, Wash.',	'Shreveport, La.',	'Spokane, Wash.',
'St. Louis, Mo.',	'Vancouver, B.C., Can.',	'Victoria, B.C., Can.',

City Neighbourhood Similarities



'Baker, Ore.',
'Havre, Mont.',
'Moose Jaw, Sask., Can.',
'Richfield, Utah'

Conclusion and future directions

- Based on the above, the aspirants should get a good sense of similarities and dissimilarities of the major cities in USA / Canada. Below are some guidelines that the aspirant could use.
 - Residential: less shops and restaurants than elsewhere
 - Residential with services: well connected with shops and restaurants
 - Cultural & going out places: well connected with lots of cultural places and clubs
 - Activity centres: large activity and large number of events
 - Shopping areas: more shops and less cultural/going out places
 - Universities: well connected with public services and university facilities
 - Randall's Island: alone within its cluster of parkings and entertainment
- Above insights combined with other readily available resources such as weather, population, demographics can help aspirants planning to come to United States make a good informed decision.

Reference

Project Report - https://github.com/kuvelky/Coursera_Capstone/blob/master/CapstoneAssignmentReport28Feb.pdf

Code - https://github.com/kuvelky/Coursera_Capstone/blob/master/DataScienceProj.ipynb