

SIMILARITY OF 10 MOST MULTICULTURAL CITIES

IBM DATA SCIENCE CAPSTONE PROJECT

MULTICULTURAL CITIES

- The population of multicultural cities are made up of different ethnic groups from all over the world, with different religions, speaking different languages.
- Locals and immigrants created diverse cultural environments, like different restaurants and art scenes.
- By comparing the venues in the cities, we'll be able to tell the similarity of the cities.

DATA ACQUISITION

- The 10 Most Multicultural Cities in the World
(theculturaltrip.com)
- Coordinates of the cities (Python Geocoder)
- Nearby venues (Foursquare)

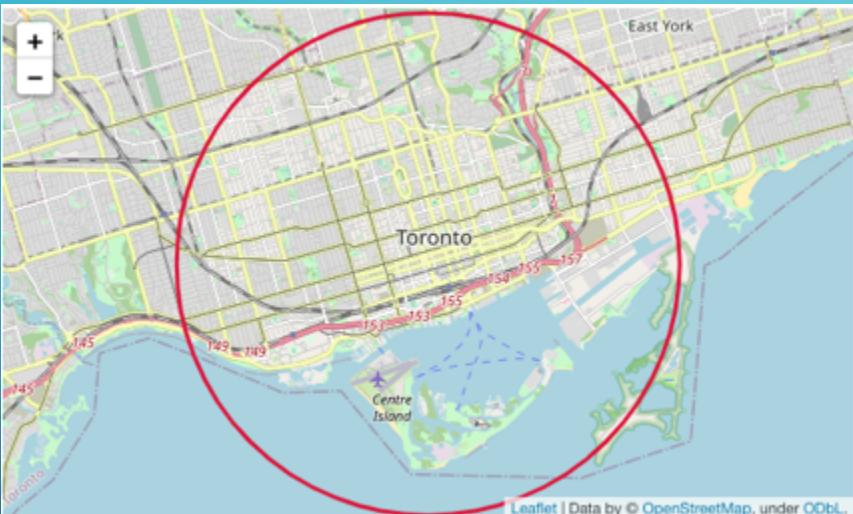
10 MOST MULTICULTURAL CITIES

City Name	Population	Multicultural Symbols
Toronto	2.7 m	Half the population born outside Canada, with 200 ethnic groups, over 140 languages
Sydney	5.3 m	Approximately 250 languages
Singapore	5.7 m	Chinese, Malay, Indian and other ethnic groups, 4 official languages
Sao Paulo	12m	Immigrants from Africa, Italy, Lebanon
San Francisco	9.6m	Immigrants from China, Germany, Italy, Mexico and Indian.
New York City	8.2m	More languages spoken in the Queens than anywhere else in the world
Paris	12.5m	14% to 20% are immigrants
Log Angeles	18m	People from 140 countries and speaking 86 languages
London	8.9m	1/3 Londoner are foreign-born, over 200 languages are spoken
Amsterdam	8.1m	178 different cultural backgrounds

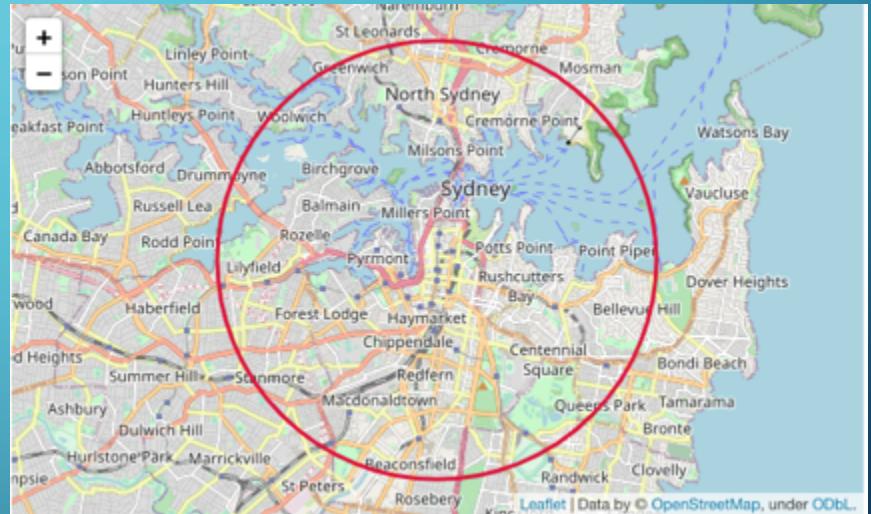
Data Source: Cultural Trip, Wikipedia

CITY CENTRAL AREA

Use Geocoder to get the coordinates of the city and use it as the Center Point.



Toronto



Sydney

NEARBY VENUES

Use Foursquare to get the nearby venues, within 5km from the Center Point of the city

	Name	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Toronto, Canada	43.64869	-79.38544	Byblos Toronto	43.647615	-79.388381	Mediterranean Restaurant
1	Toronto, Canada	43.64869	-79.38544	Downtown Toronto	43.653232	-79.385296	Neighborhood
2	Toronto, Canada	43.64869	-79.38544	Adelaide Club Toronto	43.649279	-79.381921	Gym / Fitness Center
3	Toronto, Canada	43.64869	-79.38544	Delta Hotels by Marriott Toronto	43.642882	-79.383949	Hotel
4	Toronto, Canada	43.64869	-79.38544	Nathan Phillips Square	43.652270	-79.383516	Plaza

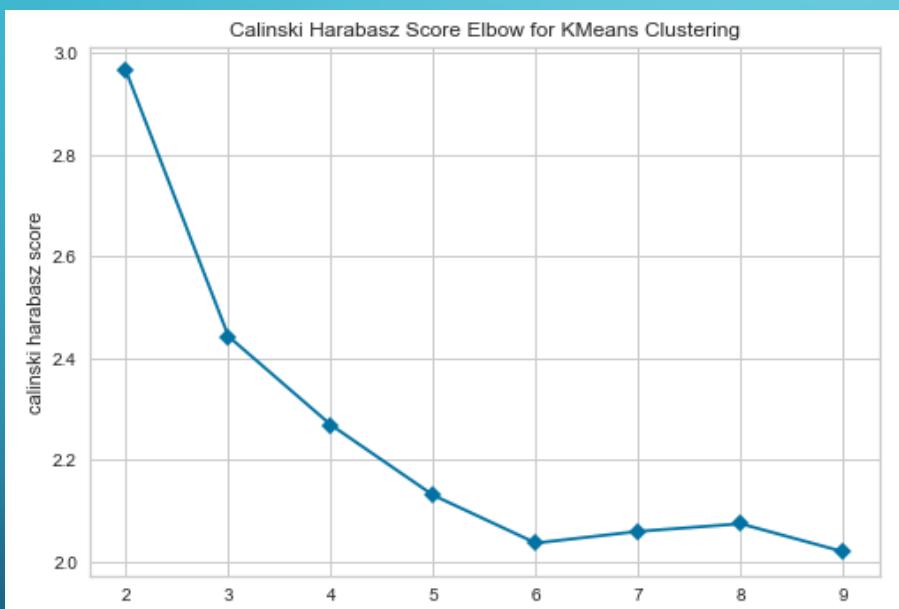
10 MOST COMMON VENUE CATEGORIES

Group the data with city name and select the 10 most common venue categories

	Name	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Amsterdam, Netherlands	Hotel	Coffee Shop	Ice Cream Shop	Bookstore	Canal	Bar	Breakfast Spot	Cocktail Bar	Bakery	Plaza
1	London, UK	Hotel	Theater	Park	Plaza	Grocery Store	Hotel Bar	Wine Bar	Art Gallery	History Museum	Bookstore
2	Los Angeles, US	Coffee Shop	Brewery	American Restaurant	Taco Place	Plaza	Theater	Bakery	Italian Restaurant	Art Gallery	Concert Hall
3	New York City, US	Park	Memorial Site	Bakery	Italian Restaurant	Theater	Garden	Hotel	Ice Cream Shop	Yoga Studio	Dance Studio
4	Paris, France	Plaza	Hotel	Ice Cream Shop	Bookstore	Garden	Wine Bar	Art Museum	Sandwich Place	Fountain	Coffee Shop
5	San Francisco, US	Bakery	Coffee Shop	Grocery Store	Pizza Place	Yoga Studio	Park	Gym / Fitness Center	Dance Studio	Marijuana Dispensary	Restaurant
6	Singapore	Hotel	Waterfront	Coffee Shop	Wine Bar	Event Space	Park	Performing Arts Venue	Art Gallery	Garden	Bistro
7	Sydney, Australia	Café	Park	Cocktail Bar	Hotel	Coffee Shop	Thai Restaurant	Ice Cream Shop	Theater	Bakery	Bookstore
8	São Paulo, Brazil	Ice Cream Shop	Italian Restaurant	Art Museum	Theater	Pizza Place	Bookstore	Park	Restaurant	Hotel	Dance Studio
9	Toronto, Canada	Coffee Shop	Park	Bakery	Café	Dessert Shop	Sandwich Place	Restaurant	Dance Studio	Japanese Restaurant	Farmers Market

K-MEANS CLUSTERING

By using Elbow Method, 6 is the suitable value for K-means Clustering



	Name	Cluster Labels
0	Toronto, Canada	2
1	Sydney, Australia	3
2	Singapore	0
3	São Paulo, Brazil	1
4	San Francisco, US	2
5	New York City, US	2
6	Paris, France	4
7	Los Angeles, US	5
8	London, UK	0
9	Amsterdam, Netherlands	4

RESULTS AND DISCUSSION

- Location is the most important factor for the formation of the cities

eg.

North American cluster: Toronto, San Francisco, New York City

European cluster: Paris, Amsterdam

- History also plays important role

eg.

UK cluster: London, Singapore

FUTURE IMPROVEMENT

- Add more cities into the analysis
- Divide the central area into smaller circulars so that more venues can be found
- Adjust the center point and radius based on the city area
- Group similar venue categories