**Data Science Capstone Project Report:**

**Clustering and Segmenting and Analysis the New York and Toronto neighbourhood**

Contents

[1. Business Understanding 3](#_Toc25398627)

[2. Problem Definition and Target audience 3](#_Toc25398628)

This file contains information about the project being executed, in this case, the Clustering and Analysis of Neighbourhood data of New York and Toronto. It is organized according to the capstone project process flow.

NOTE: This is a sample for a tutorial, so scope, plan etc., does not necessarily correspond to an actual data science project addressing a specific business question. In an actual project, the problem definition, scope, plan, personnel sections are likely to be much more detailed, based on discussions with the client (or business owner), the structure of the data science team etc.

# Business Background

The study include 2 of most popular tourist and industrial centre in the world.

Toronto

Toronto is the provincial capital of Ontario and the most populous city in Canada,

Toronto is a prominent centre for music, theatre, motion picture production, and television production, and is home to the headquarters of Canada's major national broadcast networks and media outlets.

Its varied cultural institutions, which include numerous museums and galleries, festivals and public events, entertainment districts, national historic sites, and sports activities, attract over 43 million tourists each year. Toronto is known for its many skyscrapers and high-rise buildings, in particular the tallest free-standing structure in the Western Hemisphere, the CN Tower.

Toronto has a diverse array of public spaces, from city squares to public parks overlooking ravines. Nathan Phillips Square is the city's main square in downtown, and forms the entrance to City Hall.

Large parks in the outer areas managed by the city include High Park, Humber Bay Park, Centennial Park, Downsview Park, Guild Park and Gardens, and Morningside Park. Toronto also operates several public golf courses. Most ravine lands and river bank floodplains in Toronto are public parklands.

The Royal Ontario Museum is a museum of world culture and natural history. The Toronto Zoo is home to over 5,000 animals representing over 460 distinct species. The Art Gallery of Ontario contains a large collection of Canadian, European, African and contemporary artwork, and also plays host to exhibits from museums and galleries all over the world. The Gardiner Museum of ceramic art is the only museum in Canada entirely devoted to ceramics, and the Museum's collection contains more than 2,900 ceramic works from Asia, the Americas, and Europe. The city also hosts the Ontario Science Centre, the Bata Shoe Museum, and Textile Museum of Canada.

City of New York

The City of New York, usually referred to as either New York City (NYC) or simply New York (NY), is the most populous city in the United States. New York is also the most densely populated major city in the United States. Located at the southern tip of the state of New York, the city is the center of the New York metropolitan area, the largest metropolitan area in the world by urban landmass and one of the world's most populous megacities.

New York City is often referred to collectively as the five boroughs, and in turn, there are hundreds of distinct neighborhoods throughout the boroughs, many with a definable history and character to call their own. If the boroughs were each independent cities, four of the boroughs (Brooklyn, Queens, Manhattan, and the Bronx) would be among the ten most populous cities in the United States

New York City is a global hub of business and commerce, as a center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theater, fashion, and the arts in the United States; while Silicon Alley, metonymous for New York's broad-spectrum high technology sphere, continues to expand.

Tourism is a vital industry for New York City, which has witnessed a growing combined volume of international and domestic tourists. Approximately 12 million visitors to New York City were from outside the United States, with the highest numbers from the United Kingdom, Canada, Brazil, and China.

Major tourist destinations include Times Square; Broadway theater productions; the Empire State Building; the Statue of Liberty; Ellis Island; the United Nations Headquarters; museums such as the Metropolitan Museum of Art; greenspaces such as Central Park and Washington Square Park; etc.

New York City's food culture includes an array of international cuisines influenced by the city's immigrant history. Central and Eastern European immigrants, especially Jewish immigrants from those regions, brought bagels, cheesecake, hot dogs, knishes, and delicatessens (or delis) to the city.

Italian immigrants brought New York-style pizza and Italian cuisine into the city, while Jewish immigrants and Irish immigrants brought pastrami and corned beef, respectively.

Chinese and other Asian restaurants, sandwich joints, trattorias, diners, and coffeehouses are ubiquitous throughout the city. Some 4,000 mobile food vendors licensed by the city, many immigrant-owned, have made Middle Eastern foods such as falafel and kebabs examples of modern New York street food.

The city is home to "nearly one thousand of the finest and most diverse haute cuisine restaurants in the world", according to Michelin. The New York City Department of Health and Mental Hygiene assigns letter grades to the city's restaurants based upon their inspection results. As of 2019, there were 27,043 restaurants in the city, The Queens Night Market in Flushing Meadows–Corona Park attracts over 10,000 people nightly to sample food from over 85 countries.

# Problem Definition

This project is to solve the below business cases:

1. First Business case is to target tourists, who are visiting the neighbourhood first time can get an understanding of the top venues surrounding in New York and Toronto. Also they will get a simple comparison between the top venues in different neighbourhood.
2. 2nd business case is, this analysis will help the entrepreneur to give an idea about the which type of business is popular in which area and which type of cuisine popular in which area and will give and idea of where to invest.

# Problem Definition and Target audience

The purpose of this sample is to show how to instantiate and execute a project using the data science structure and tools.

The intended audience of this project is usually the tourists who are first time visiting the neighbourhood of Toronto or New York City and don’t have an idea of the top venues around a place.

So they are sometime lost and didn’t get chance to visit the top places. Using this analysis we can help the tourists to self-guide to where to visit according to their choice.

Target audience of the 2nd use case is upcoming entrepreneurs who want to establish some business in these neighbourhoods.

The issue is most of the times the less experience budding entrepreneurs have less or no idea about the top running business in certain localities, which cuisine is most popular in which area, which shop is most trending in which area, and which area has what popular venues. So lack of this data often misguides them to enter in wrong business and can be a trouble to them.