Introduction and Business Problem

- Company XYZ, headquartered in India is planning to set up a conference in North America. It has offices in both Toronto and New York. Janice, the senior manager has never been to both cities and is currently deciding between Toronto, Canada and New York, USA. The two cities are located in the east coast, are financial capitals and have a diverse culture. She wants to explore the two cities and compare them to find out how similar or dissimilar they are.
- She wants to know the different neighborhoods of the cities. They want to particularly explore Downtown Toronto neighborhoods and Manhattan city to find out the most common venues in both neighborhoods, as they will most likely host the conference in downtown.
- She also wants to explore the demographics of the two cities such as population, average income and density of each neighborhood.

Data

We will need the following data:

- Toronto data that contains different borough, neighborhoods and their postal codes https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

]:

	Postalcode	Borough	Neighborhood
0	M1B	Scarborough	Rouge, Malvern
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union
2	M1E	Scarborough	Guildwood, Morningside, West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

- Coordinates of the Toronto neighborhoods http://cocl.us/Geospatial_data/Geospatial_Coordinates.csv

	Postalcode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711

- New York data containing different boroughs, neighborhoods and their coordinates https://cocl.us/new_york_dataset

2]:


```
{'type': 'Feature',
 'id': 'nyu_2451_34572.1',
 'geometry': {'type': 'Point',
  'coordinates': [-73.84720052054902, 40.89470517661]},
 'geometry_name': 'geom',
 'properties': {'name': 'Wakefield',
  'stacked': 1,
  'annoline1': 'Wakefield',
  'annoline2': None,
  'annoline3': None,
  'annoangle': 0.0,
  'borough': 'Bronx',
  'bbox': [-73.84720052054902,
   40.89470517661,
   -73.84720052054902,
   40.89470517661]}}
```

- The coordinates (latitude, longitude) of the neighborhoods of Downtown Toronto and Manhattan using geolocator

```
# Let's get the geographical coordinates of Manhattan.
address = 'Manhattan, NY'

geolocator = Nominatim()
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of Manhattan are {}, {}.'.format(latitud
```

```
/opt/conda/envs/Python36/lib/python3.6/site-packages/ipykernel/__main__
nt` is strongly discouraged, as it violates Nominatim's ToS https://ope
P errors. Please specify a custom `user_agent` with `Nominatim(user_age
tions.default_user_agent = "my-application"`. In geopy 2.0 this will be
```

```
The geograpical coordinate of Manhattan are 40.7900869, -73.9598295.
```

From Foursquare API, we will need the following data:

- Nearby venue names with their latitudes and longitudes

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Rosedale	43.679563	-79.377529	Mooredale House	43.678631	-79.380091	Building
1	Rosedale	43.679563	-79.377529	Rosedale Park	43.682328	-79.378934	Playground
2	Rosedale	43.679563	-79.377529	Whitney Park	43.682036	-79.373788	Park
3	Rosedale	43.679563	-79.377529	Alex Murray Parkette	43.678300	-79.382773	Park
4	Rosedale	43.679563	-79.377529	Milkman's Lane	43.676352	-79.373842	Trail

- Top 5 most common venues for each neighborhood
- Top 10 most common venues for each neighborhood

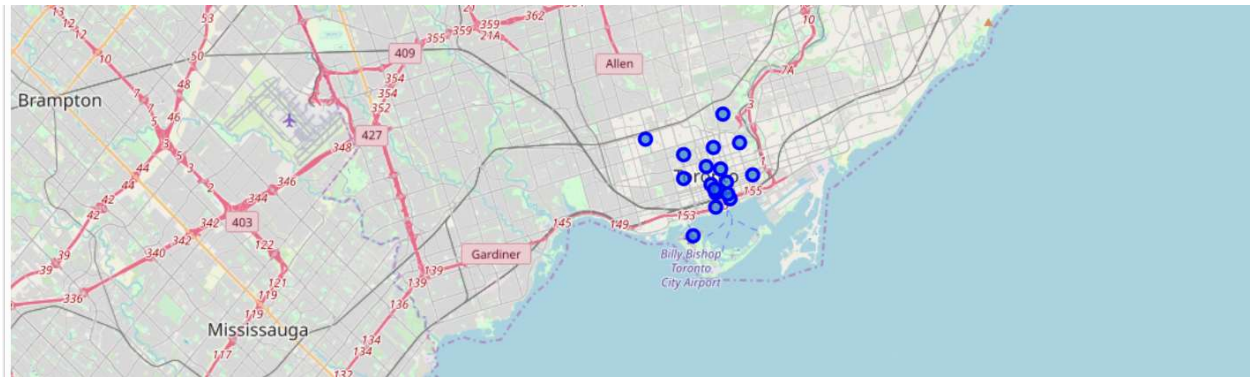
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Adelaide, King, Richmond	Coffee Shop	Café	Thai Restaurant	Bar	American Restaurant	Steakhouse	Gym	Burger Joint	Sushi Restaurant	Hotel
1	Berczy Park	Coffee Shop	Cocktail Bar	Café	Beer Bar	Bakery	Seafood Restaurant	Farmers Market	Steakhouse	Cheese Shop	Eastern European Restaurant
2	CN Tower, Bathurst Quay, Island airport, Harbo...	Airport Service	Airport Lounge	Airport Terminal	Harbor / Marina	Bar	Coffee Shop	Plane	Sculpture Garden	Boutique	Boat or Ferry
3	Cabbagetown, St. James Town	Coffee Shop	Café	Pub	Italian Restaurant	Park	Bakery	Pizza Place	Restaurant	Japanese Restaurant	Sandwich Place

- Average Income, Population and Density of New York
  -Wikipedia (Demographics of New York City)
- Average Income, Population and Density of Toronto
  -Wikipedia (Demographics of Toronto neighbourhoods)

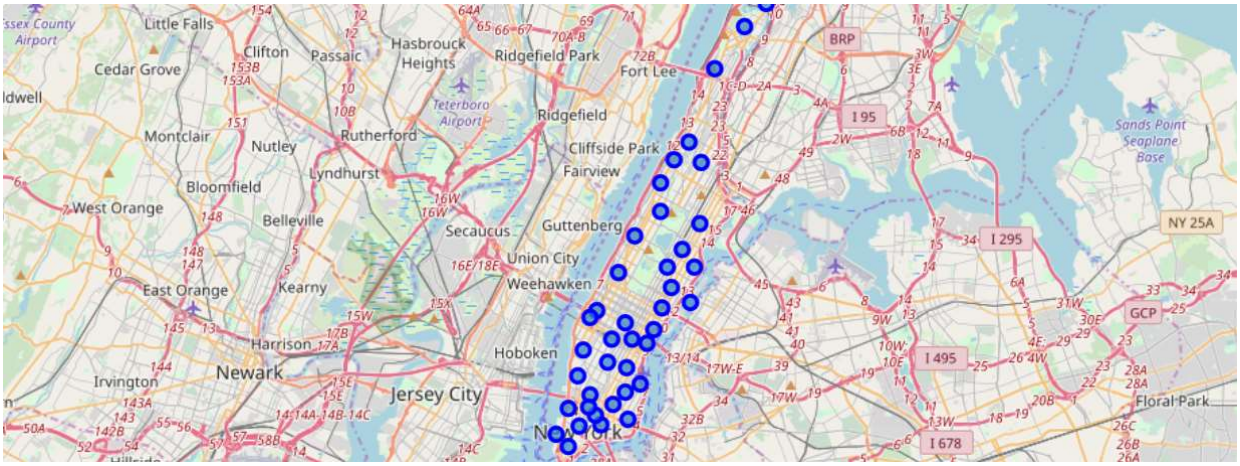
Methodology

- Extract Toronto and New York borough and neighborhood data
- Get the latitudes and longitudes of the neighborhoods
- Compare Downtown Toronto and Manhattan
  -Coordinates

Downtown Toronto neighborhoods

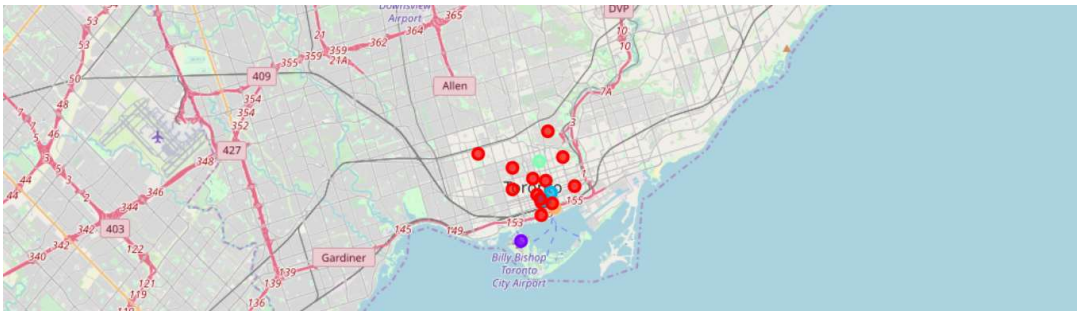


Manhattan neighborhoods

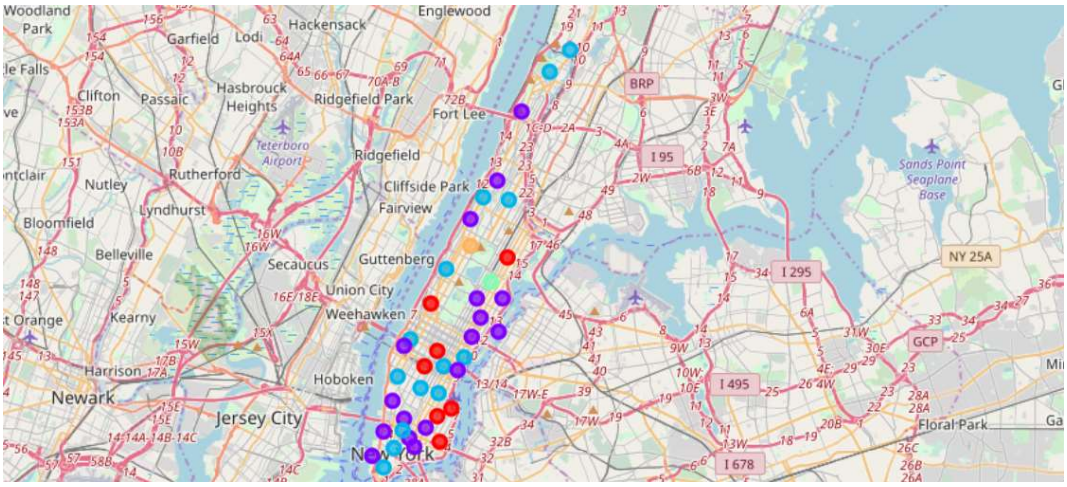


-use Foursquare API, to get nearby venues for each city
-Analyze each neighborhood to find the most common venues
-For each neighborhood, find the 10 most common venues
-Run k means clustering for each neighborhood by setting k=5 clusters
-Map 10 most common venues of each neighborhood with their clusters

Downtown Toronto neighborhood cluster



Manhattan neighborhood cluster:



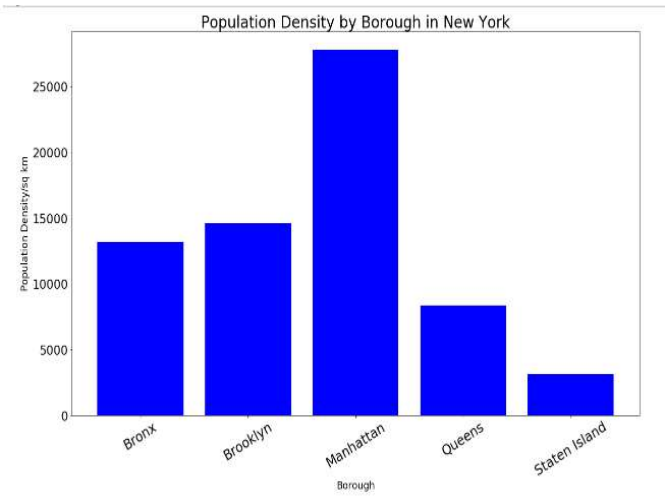
-Use Folium maps to compare the clusters of neighborhoods

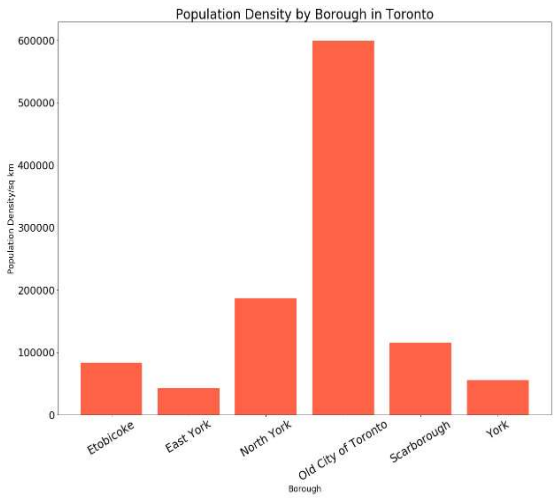
- Extract Average Income, Population and Density data of the two cities from Wikipedia.
- Compare the two cities for each of the categories using bar charts.

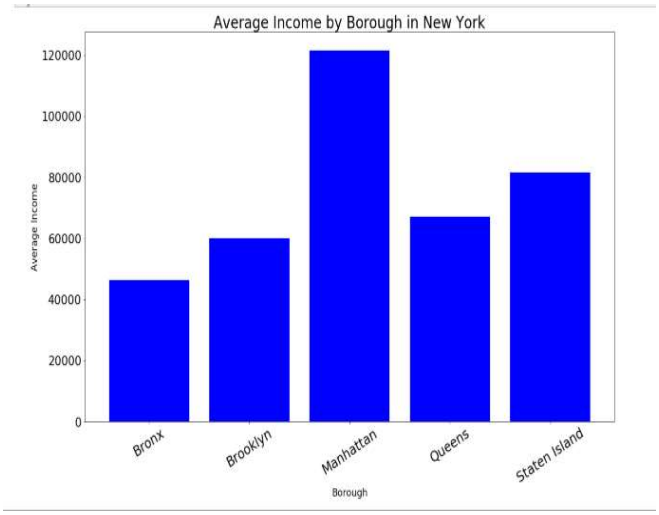
Results

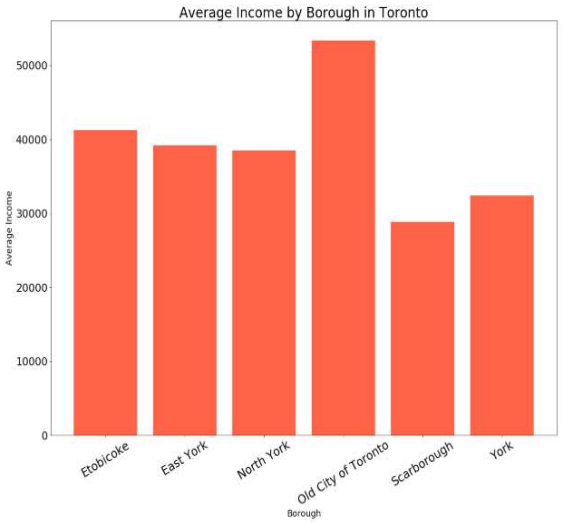
Manhattan is a bigger city, has more neighborhoods than Downtown Toronto. Manhattan and Downtown Toronto both are financial districts with World trade centre, Bay Street located in the heart of the city. Coffee shops, followed by restaurants seem to be the most common venue for Toronto downtown neighborhoods while restaurants are the most common venues for Manhattan. Toronto downtown as seen from the map is smaller than Manhattan city.

In terms of Population and Average income of both cities neighborhoods, they also display a similar trend. Manhattan and Old city of Toronto have the highest average income and are the highest population density. However, New York seems to have a higher average income compared to Toronto.









Discussion/Recommendation

We did a general comparison of the two cities. The model can be further improved by comparing:

- Hotel prices for the conference in both Manhattan and Downtown Toronto.
- Taking into account the number of conference attendees in both cities to reduce travel time for the employees.
- The distance between the chosen city office and the hotel for the conference.
- The data accuracy of venues depends on Four square API.

Conclusion

Toronto and New York are both similar in terms of Population density, with each of them having one dense borough and lower density for other boroughs. Average income in New York is higher, with Manhattan and old Toronto city having the highest Average Income.

Both cities have similar attractions and venue categories, with Toronto downtown being smaller than Manhattan.

-