

*LOCATING A SUITABLE BOROUGH AND
THEN NEIGHBOURHOOD FOR PERSONAL
TRAINER RELOCATION:
TORONTO VS NEW YORK*

INTRODUCTION

- The project involves the analysis of Boroughs within New York and Toronto.
- A personal trainer is interested in relocating to increase their client base.
- The trainer is interested in which area may have the greatest potential customers.
- It is likely the area with the most gym related venues represents an area with significant potential customer volume.
- A comparison of gym venues within the two cities will indicate which city, and specifically which boroughs and then neighbourhoods maximum customers are likely to occur.

OBJECTIVES

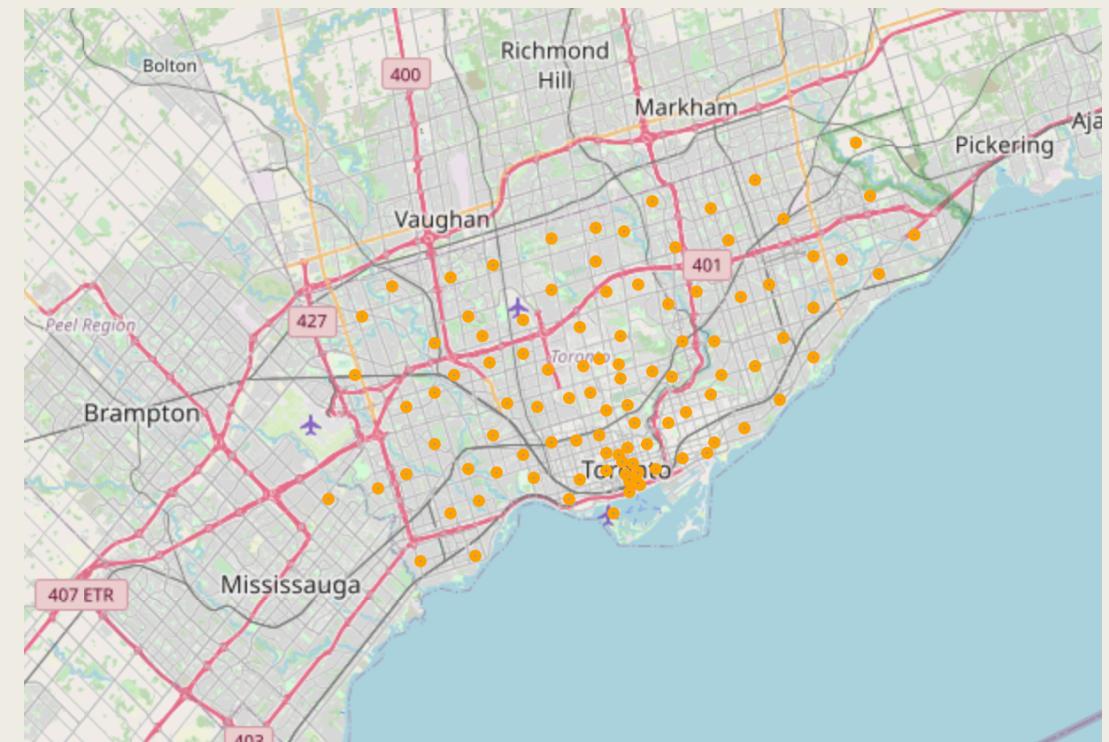
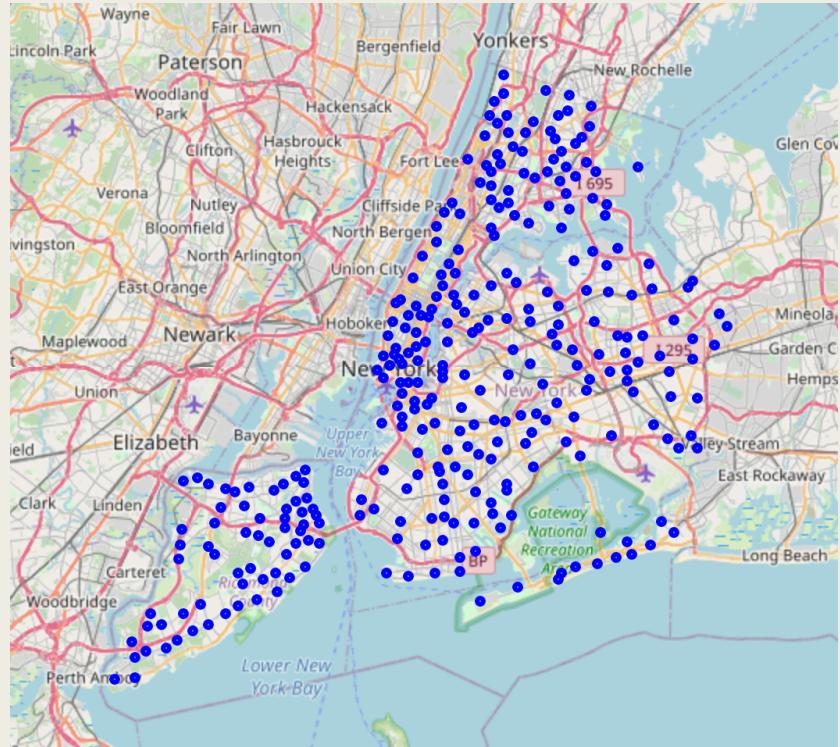
- Locate the boroughs within New York and Toronto which have the highest occurrence of venues of interest to a personal trainer.
- Locate the neighbourhoods within these boroughs which have the highest occurrence of venues of interest to a personal trainer.
- Cluster the above data and visualise where these areas of interest occur.
- Visualise the presence of venues of interest within the final selected neighbourhoods.

DATA

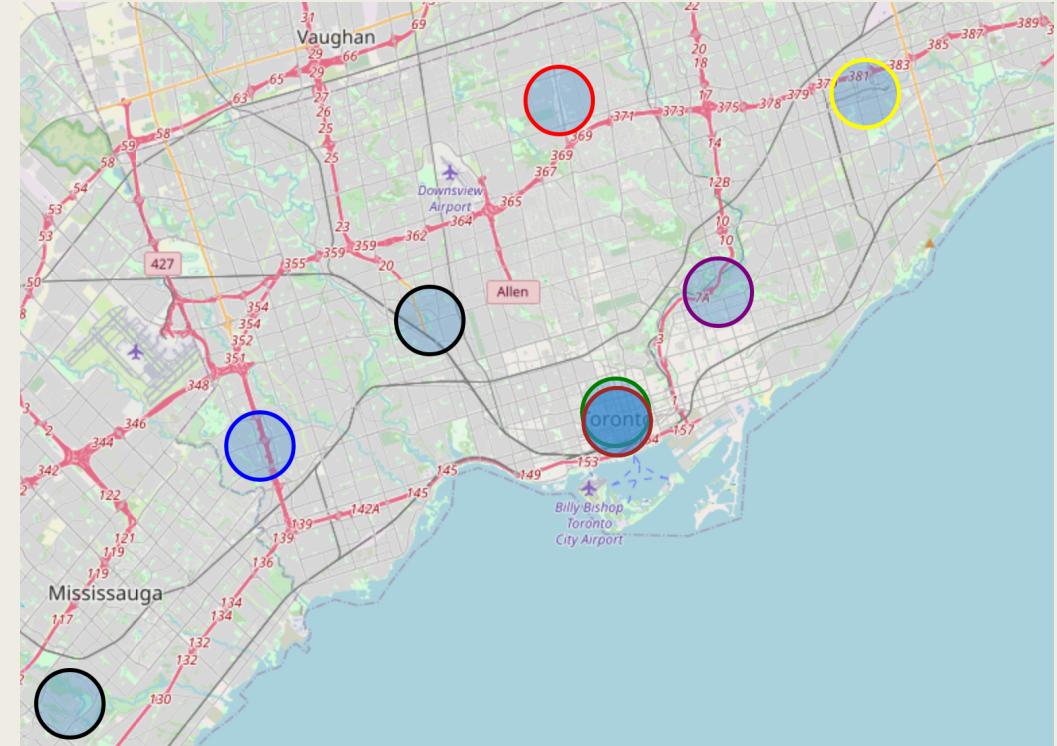
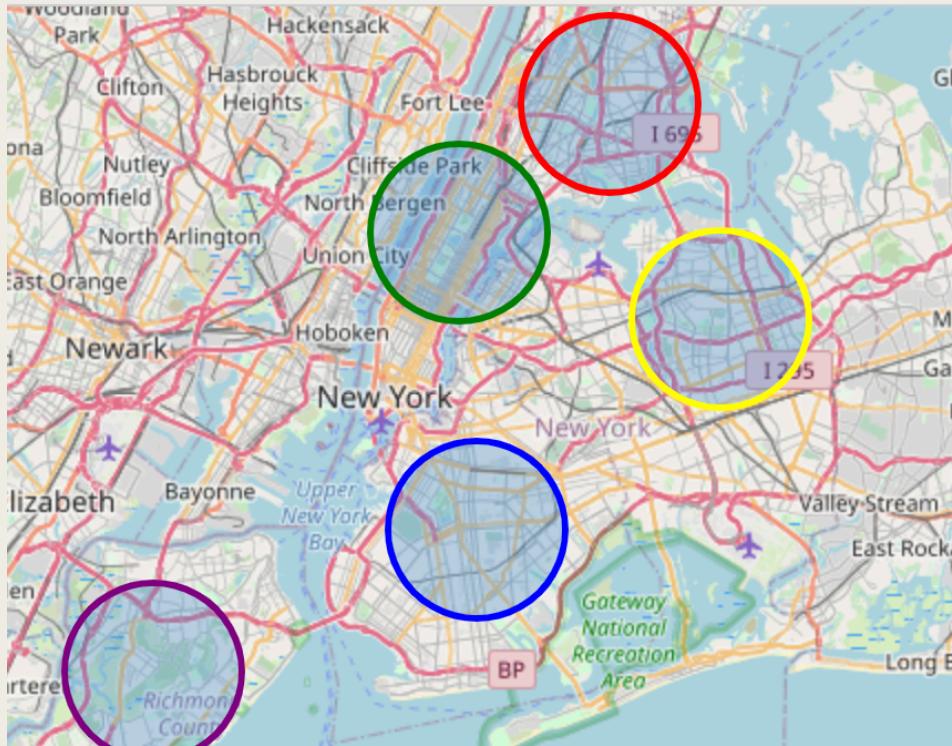
Data	Source
1) Longitude and latitudes of New York and Toronto Boroughs and Neighbourhoods.	Python GeoCoder package
2) Venue information; specifically gym related venues	foursquare API
3) Boroughs and Neighbourhoods in New York	https://cocl.us/new_york_dataset
4) Boroughs and Neighbourhoods in Toronto	https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

METHOD OLOGY

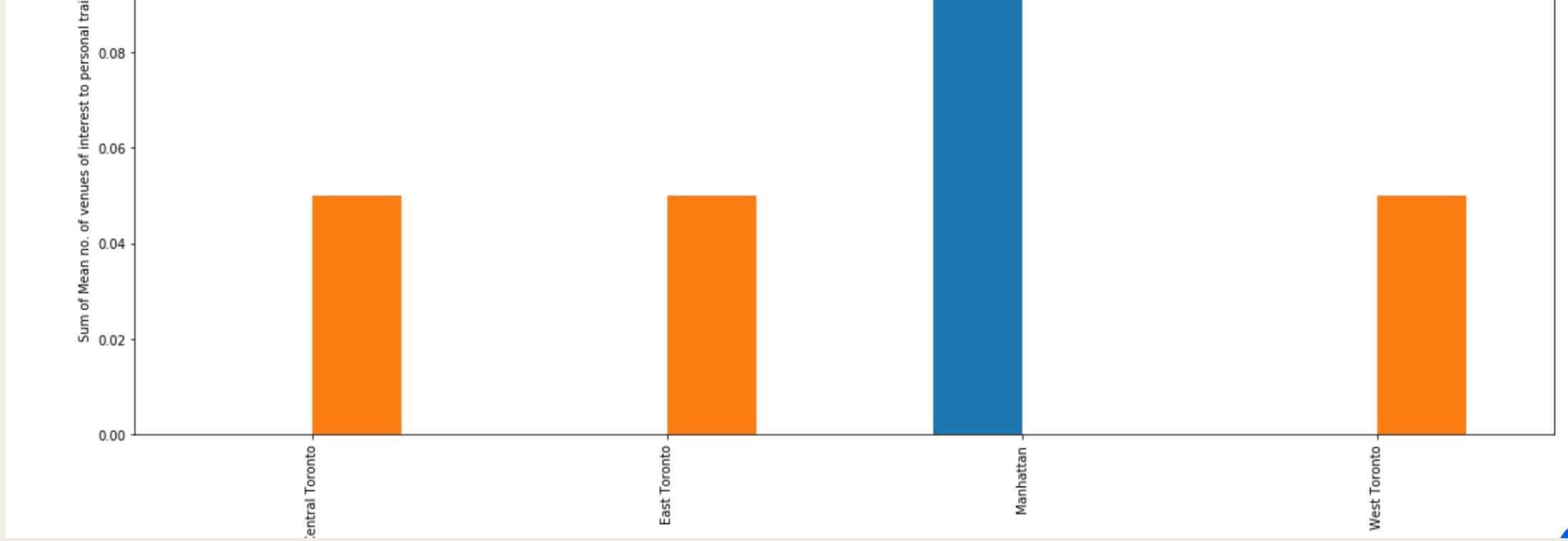
- 1) Obtain relevant boroughs and neighbourhoods for Toronto and New York; including longitude and latitudes.
- 2) Visualise the neighbourhoods and then boroughs within these cities.
- 3) Use the foursquare API to pull 100 venues for each borough. Look at the mean occurrence of venues of interest to a personal trainer; relative to ones that are not.
- 4) Compare the sum of mean occurrence for boroughs within the two cities.
- 5) Repeat steps 3 and 4 with the neighbourhoods which display the greatest no. of venues of interest.
- 6) Cluster these neighbourhoods k-means to visualise where neighbourhoods' of interest are situated.
- 7) Compare the sum of mean venues of interest to understand which neighbourhood has the highest occurrence and therefore most potential for relocation.



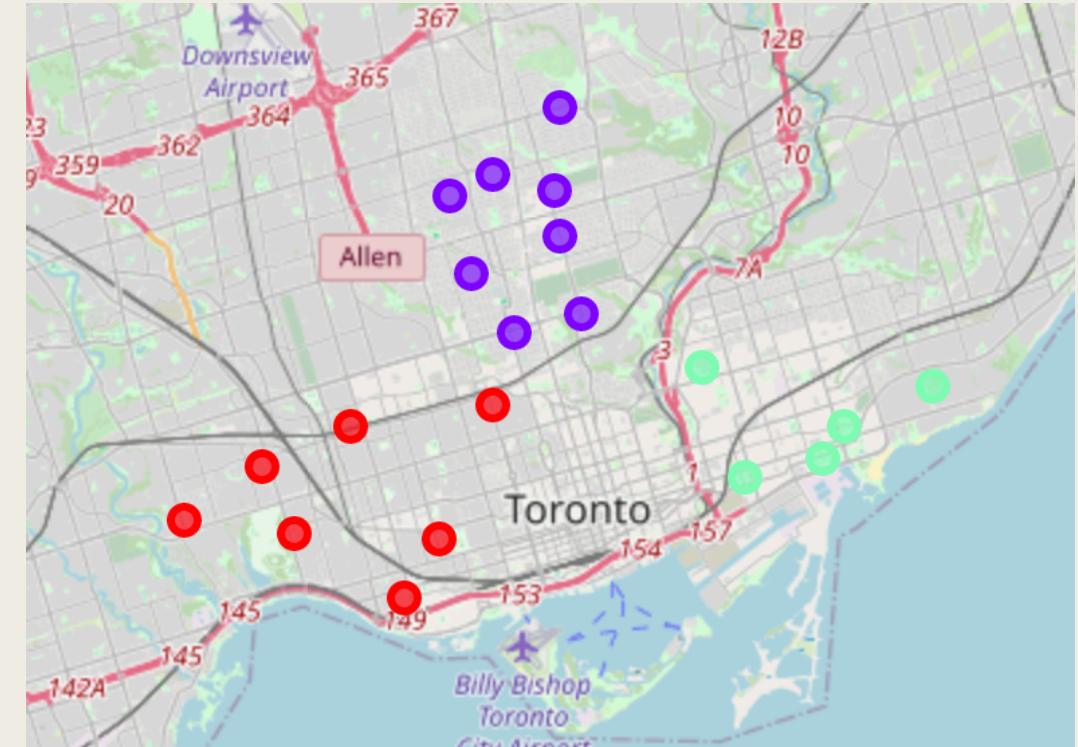
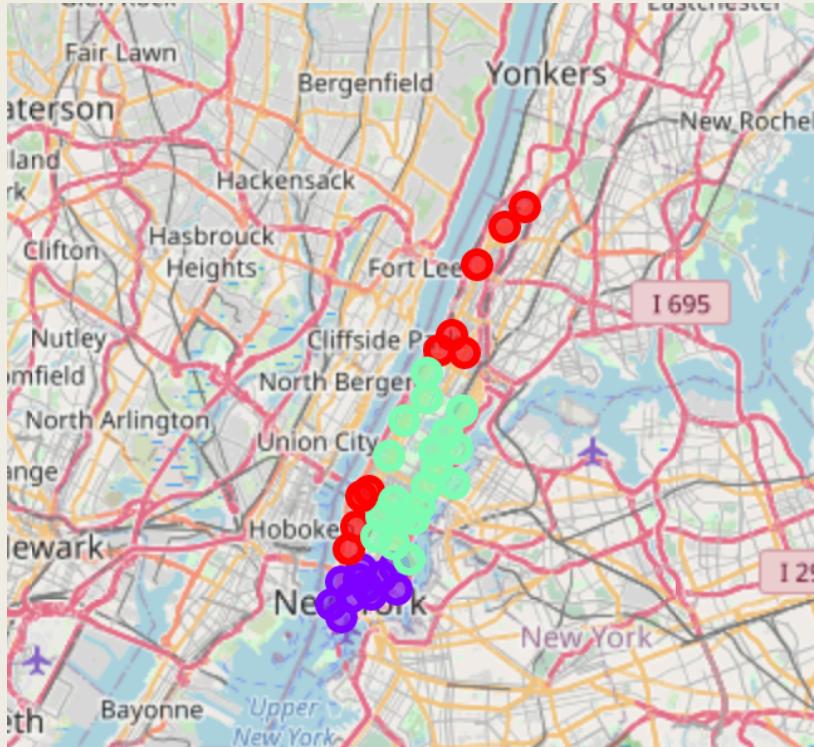
NEIGHBORHOODS WITHIN NEW YORK (BLUE) & TORONTO (ORANGE)

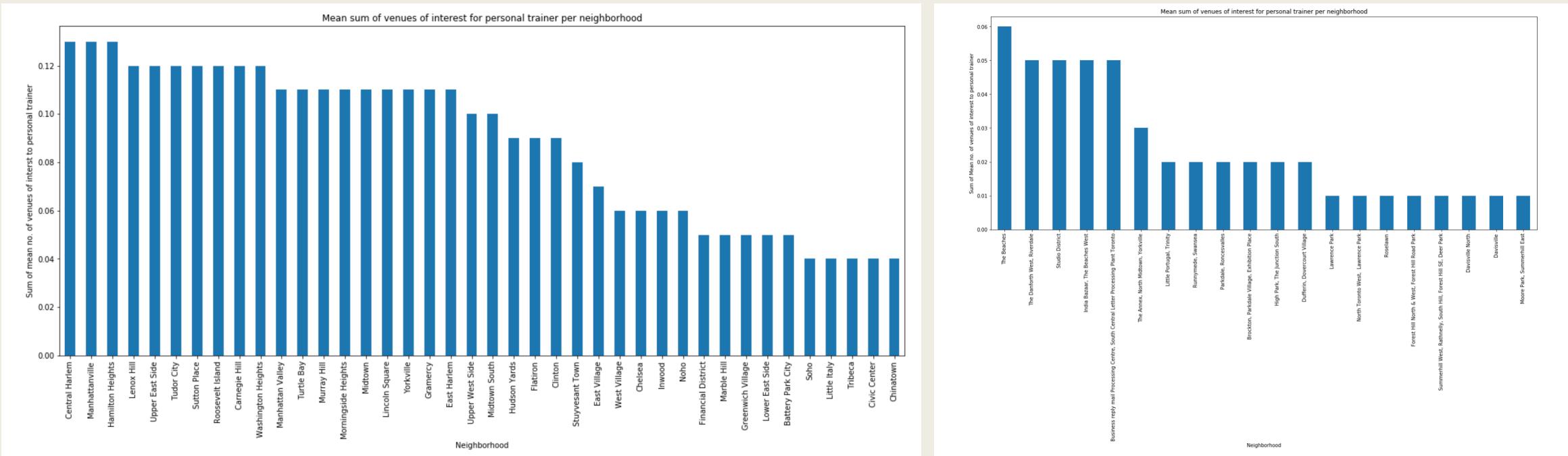


BOROUGHS WITHIN NEW YORK (LEFT) & TORONTO (RIGHT)

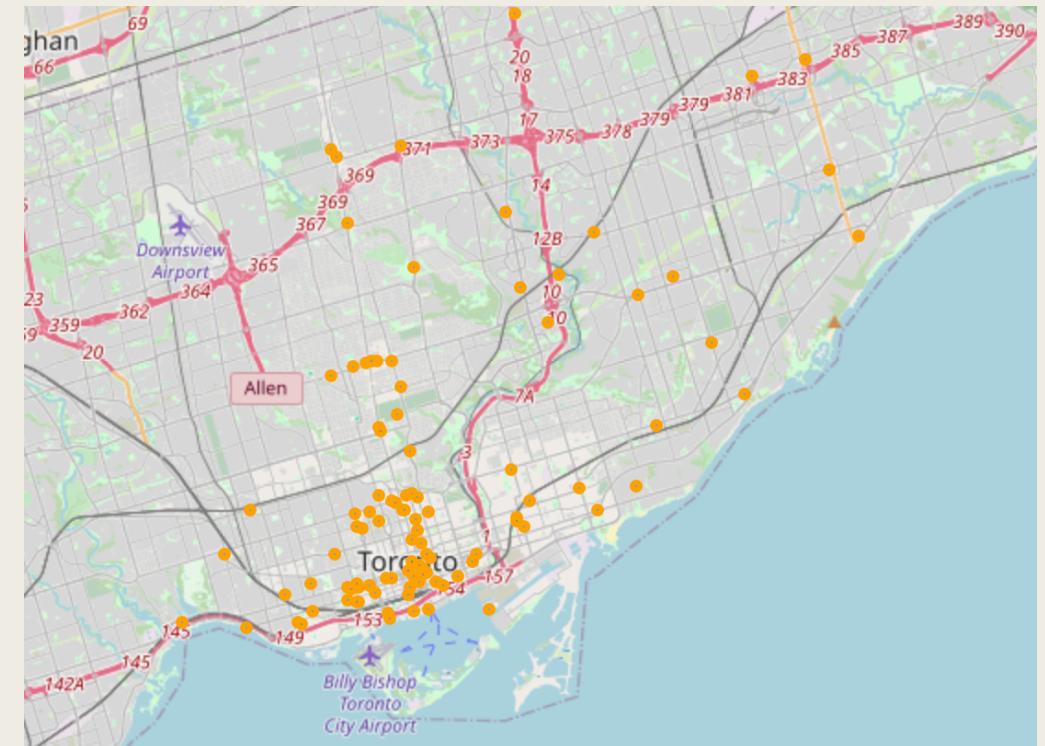
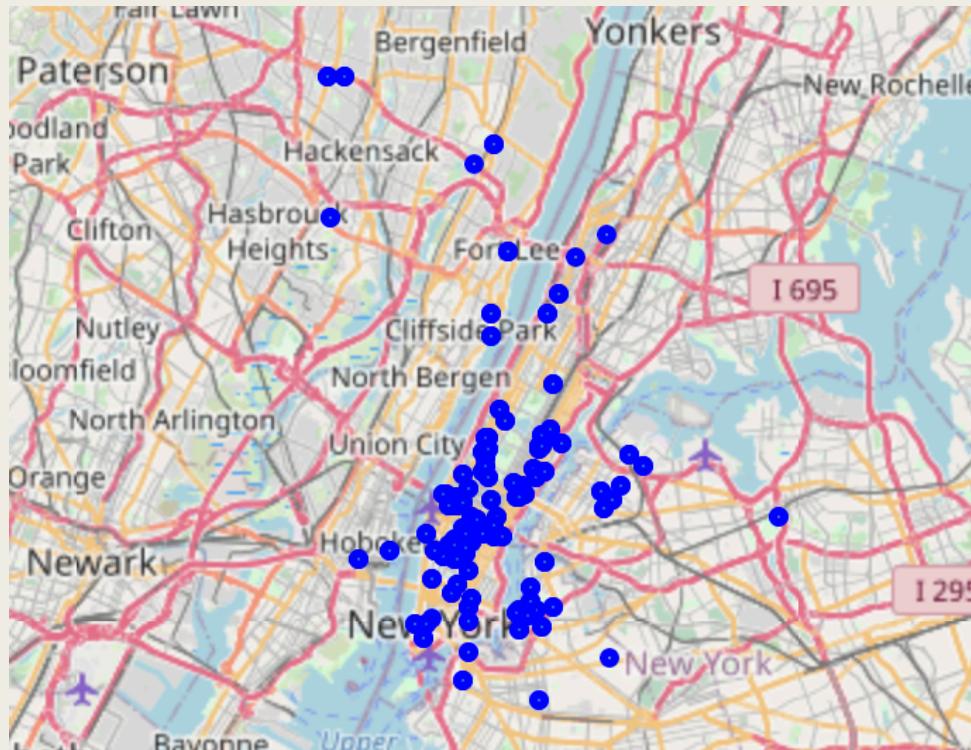


BOROUGHS WITH HIGHEST SUM OF MEAN OCCURRENCE OF VENUES OF
INTEREST. NEIGHBORHOODS WITHIN THESE BOROUGHS THEN
INVESTIGATED FURTHER
NEW YORK (BLUE) & TORONTO (ORANGE)





COMPARING THE SUM OF MEAN OCCURRENCE OF VENUES OF INTEREST. NEW YORK (LEFT) & TORONTO (RIGHT)



VISUALISING WHERE VENUES RELATED TO QUERY “GYM” ARE SITUATED IN AREAS OF INTEREST FOR NEW YORK (BLUE) & TORONTO (ORANGE)

FACTORS TO NOTE

- 100 venues are returned by foursquare per location analysed. Therefore, mean occurrence of venues of interest is only a approximation and may not be truly representative of the wider borough/neighbourhood.
- New York had a wider available range of venue categories, and this may play a crucial role in creating consistently higher venue occurrence values. This will be useful to investigate further.
- Foursquare venue pull can result in variance in highest performing neighbourhoods. This is likely due to the random nature of 100 venues pulled.

CONCLUSION

- On the basis of occurrence of venues of interest, **Central Harlem, Manhattanville and Hamilton Heights** haVE the most to offer with a sum of **0.13** of mean occurrence of venues of interest. Assuming a personal trainer has complete freedom to relocate.
- For **Toronto, The Beaches** is the highest neighbourhood with a sum of **0.05**. Although, less venues of interest appeared within the foursquare API request.
- Clearly other factors would have to be considered in the viability of any neighbourhood for potential relocation.

FURTHER STUDY

It will be interesting to look at others factors which may have influence, these may include:

- More venues per area to increase representativeness of study
- House price within each neighbourhood
- Average age of residents and their disposable income
- Presence of other personal trainers in the area
- Available accommodation for future house relocation
- Likelihood of gyms wanting to partner with personal trainer.