**FINAL REPORT**

**CAPSTONE PROJECT OF BATTLE OF NEIGHBORHOOD**

**INTRODUCTION**

This capstone project is for IBM Data Science course. In this project I am assuming myself as an entrepreneur to start a chain for ‘Thai Restaurant’ as people in Toronto like Thai food. Which place is suitable for the start-up restaurant in Toronto?

**BUSSINESS PROBLEM**

The aim of this problem is to find the most suitable place to open the Thai restaurant in Toronto where I should setup my first restaurant. By using Data Science tools I will be solving this problem by clustering the restaurants according to the neighbourhood, the areas having less number of restaurants or opening at a new place with no Thai restaurants. Which place is profitable?

**DATA SECTION**

For this project we need the following data:

1. Toronto city data that contains neighbourhoods and boroughs along with latitude and longitude.

* Data source for neighbourhoods and borough <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>
* Data source for latitude and longitude 'http://cocl.us/Geospatial\_data'

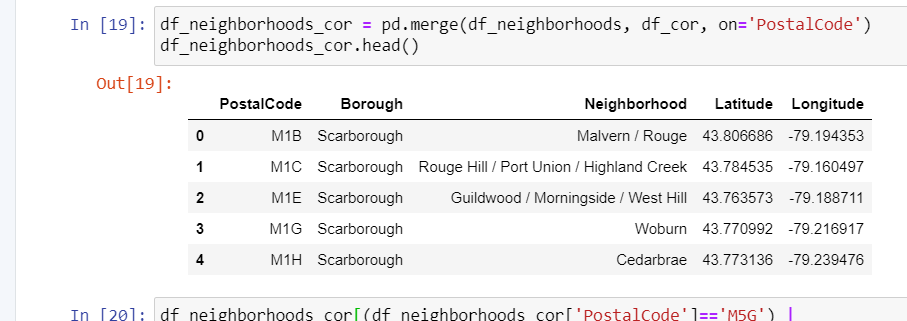
1. Thai Restaurant in each neighbourhood of Toronto

* Data source: Foursquare API
* This will provide all the venues in each neighbourhood. We can use only Thai restaurant venues of our choice.

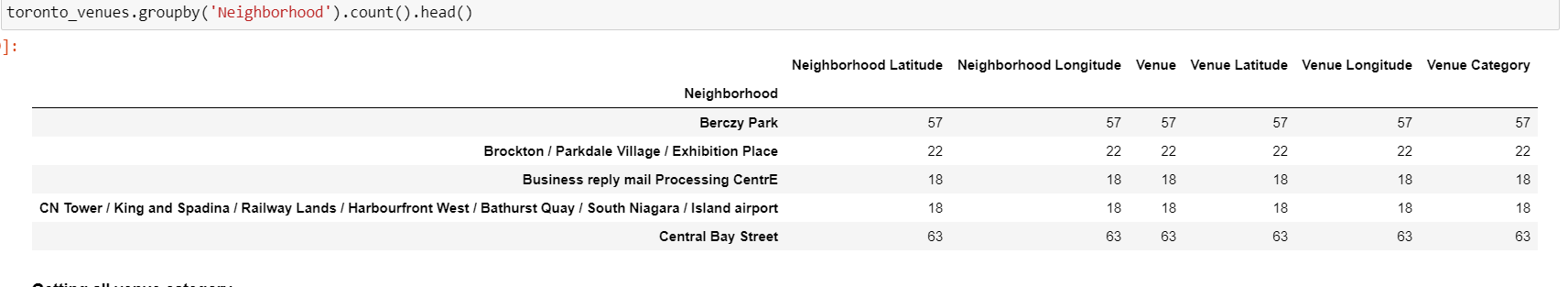
**METHODOLOGY**

1. We begin by collecting the Toronto city data from "[https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M"](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M%22)

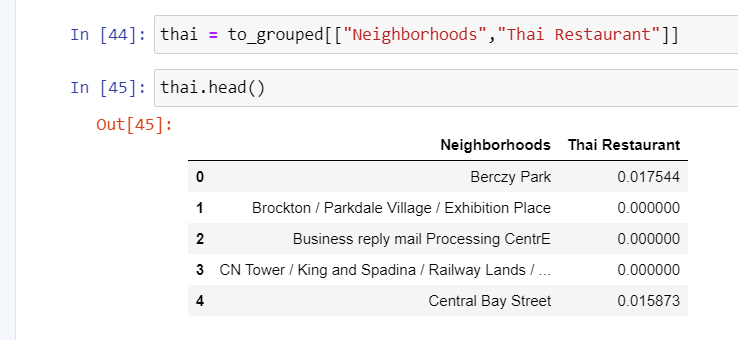
'http://cocl.us/Geospatial\_data



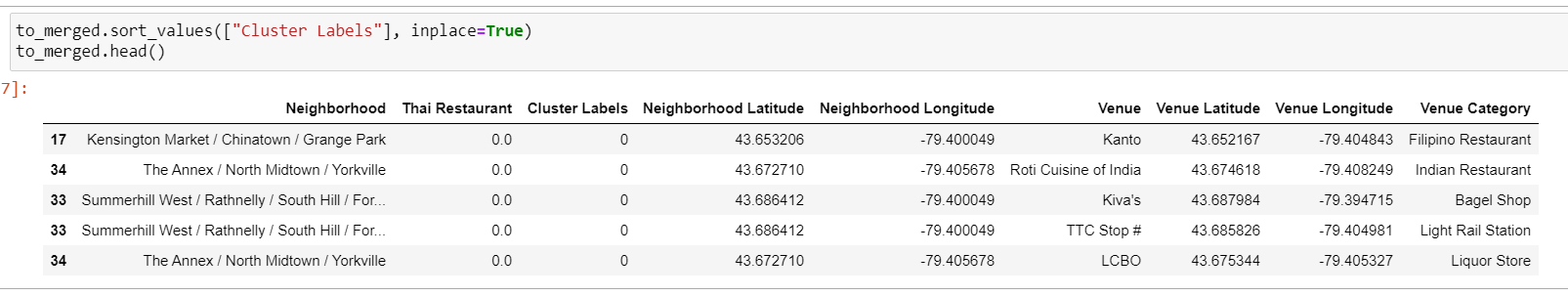
1. We will find all venues for each neighbourhood using FourSquare API.



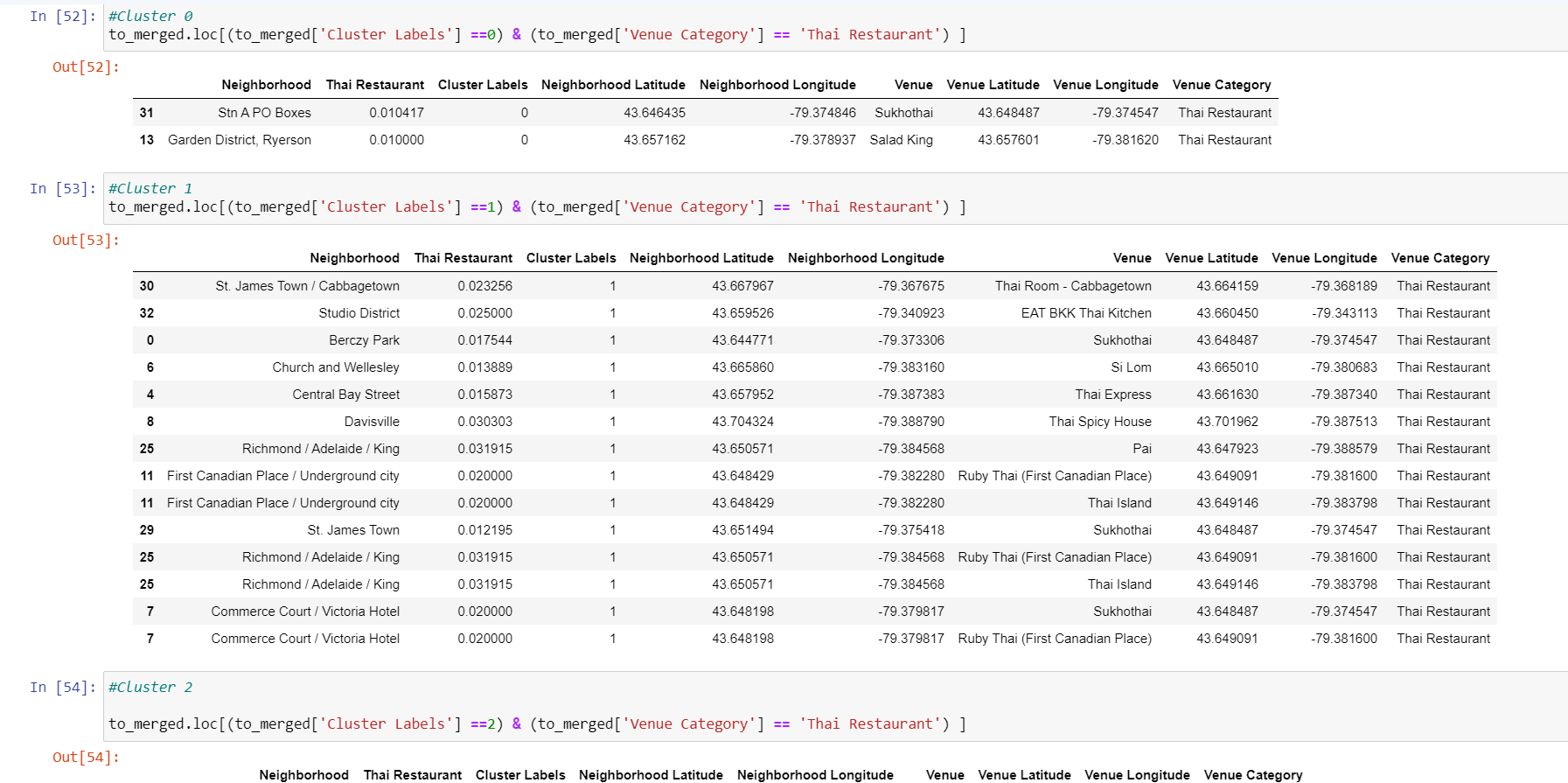
1. We will then filter out all Thai Restaurant venues.



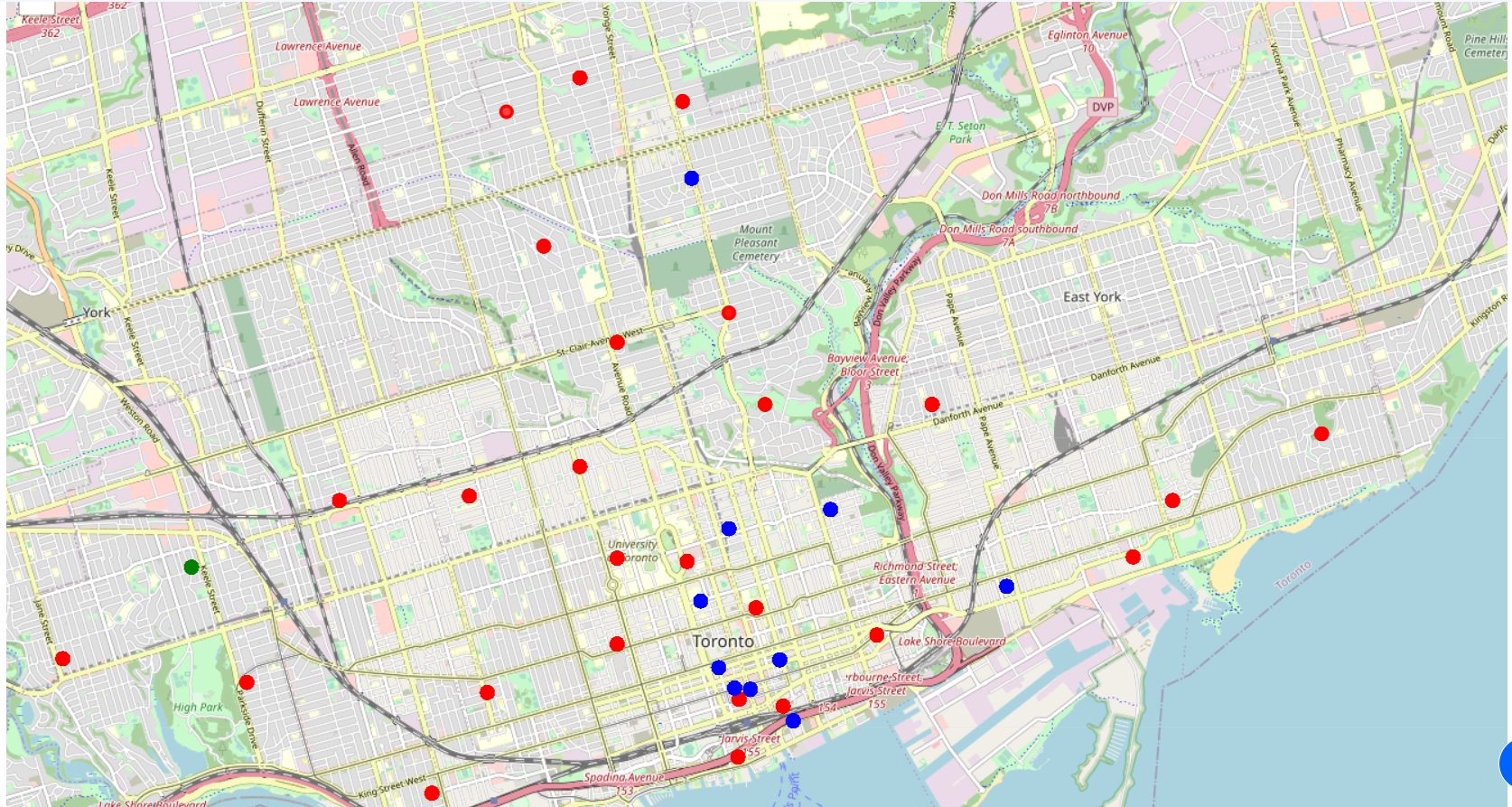
1. We will cluster the restaurants according to the neighbourhood.



1. We will analyse each cluster to find out which area is having less number of Thai restaurant.



1. We will visualize the neighbourhoods using python's Folium library.



**RESULT**

So we have answer to every question in methodology. From this analysis we come to know that in cluster 2 with green dots has less number of Thai restaurants we have great chance to establish the restaurant in that cluster near RUNNYMEDE ROAD, DUNDAS STREET WEST, HIGH PARK/ THE JUNCTION SOUTH etc.

**CONCLUSION**

We have analyse whole data of Toronto city according to the given data if there is any improvement in the data then the results will be different.