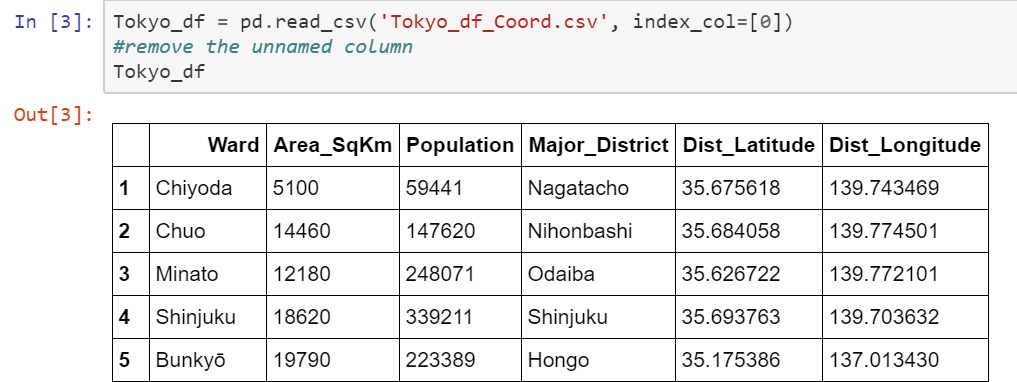
Introduction

This project is part of the Coursera Capstone project and will utilise Foursquare API for venue information of Tokyo for analysis on the restaurants located in Tokyo. This analysis can be used for people who are looking to explore Tokyo and would like to know which locations have similar restaurants. Restaurant owners can also learn which areas to open their branches.

Data used

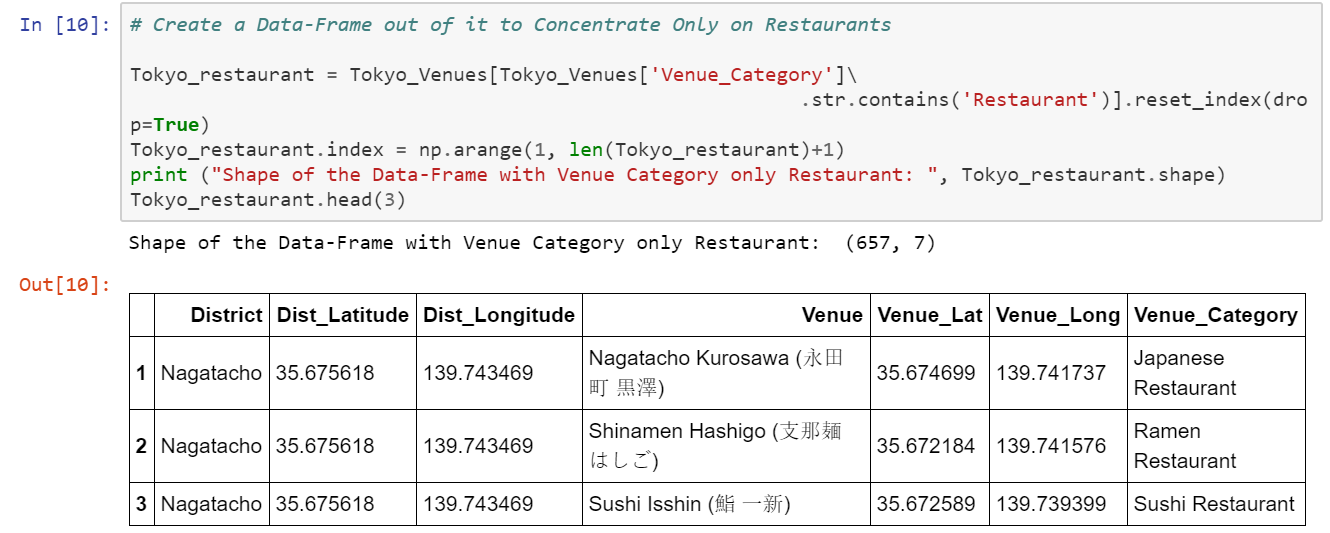
A table from https://en.wikipedia.org/wiki/Special\_wards\_of\_Tokyo will be used for the location data. It will be imported into a dataframe using BeautifulSoup.



Geopy will be used to get the coordinates of the locations.

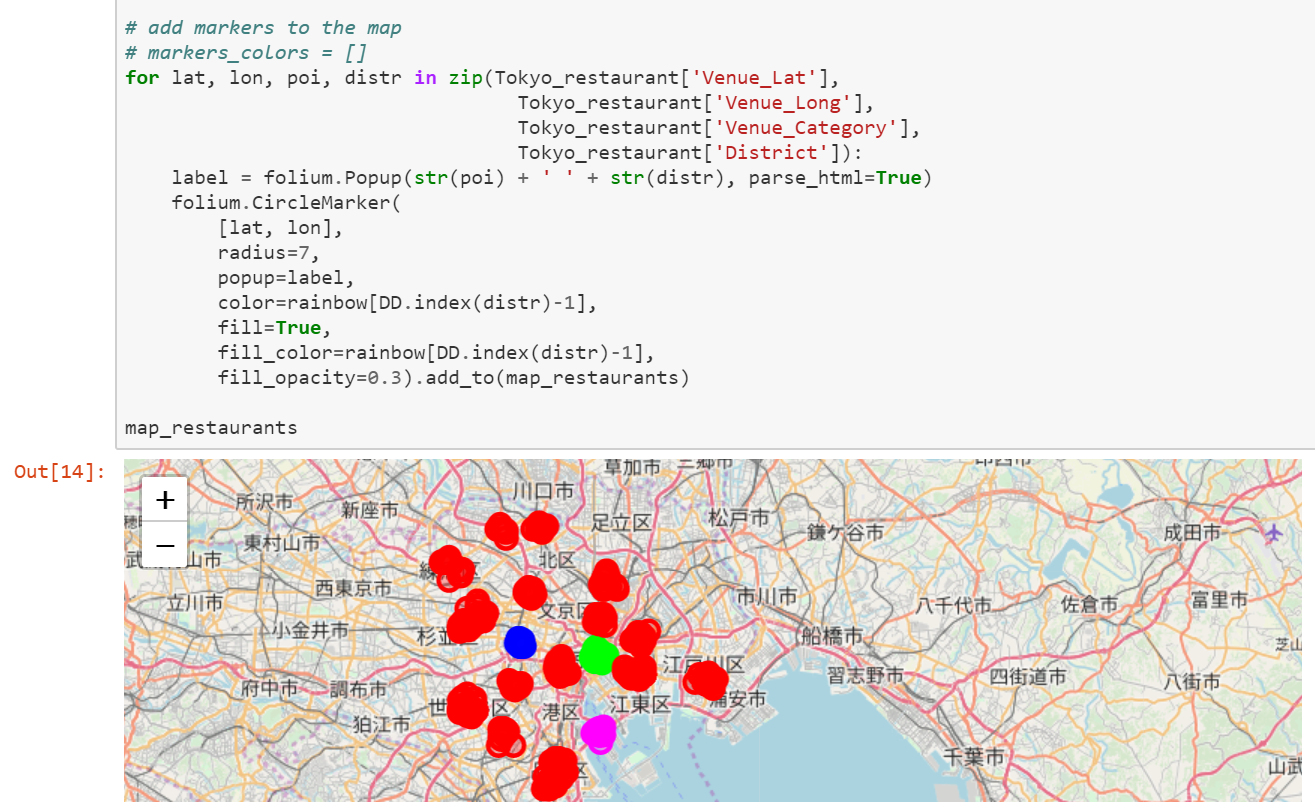


Foursquare will be used for venue data, I will query it for top 100 venues within 1km of the districts in Tokyo. The resultant JSON file will be imported into another dataframe.

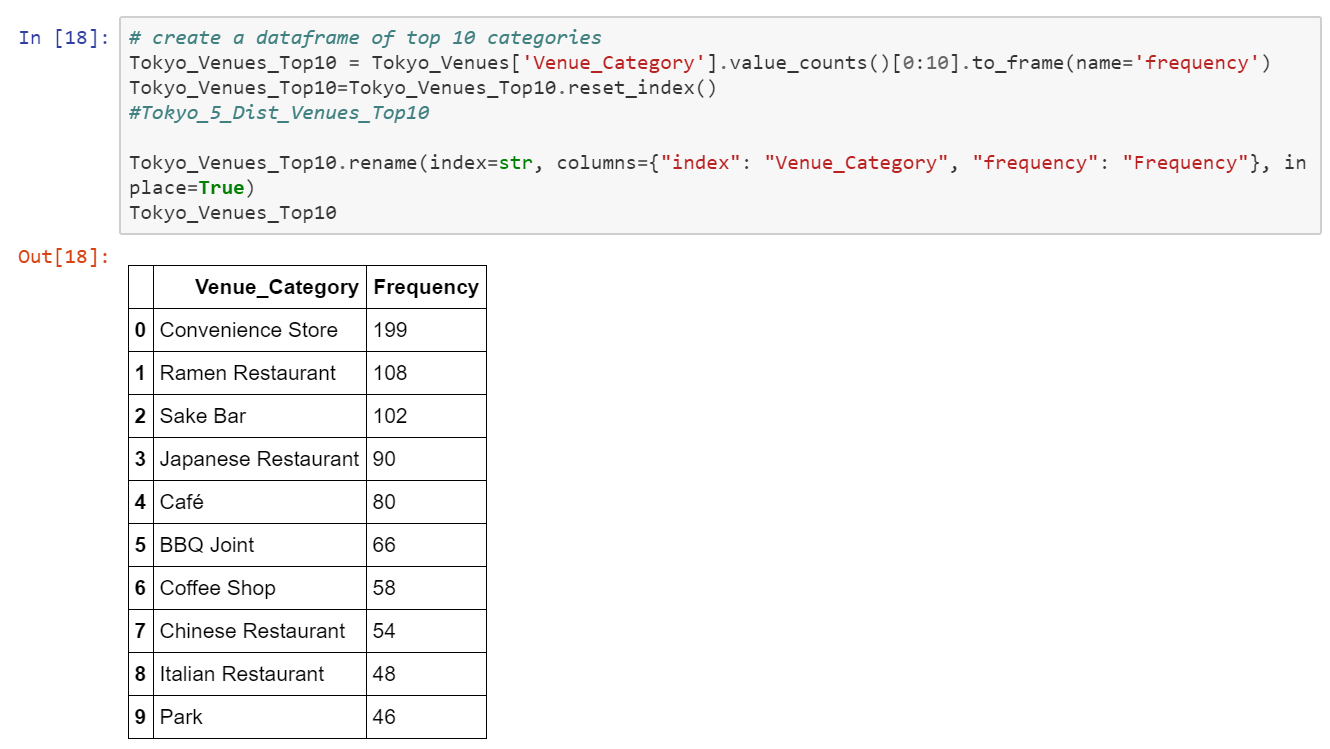


Methodology

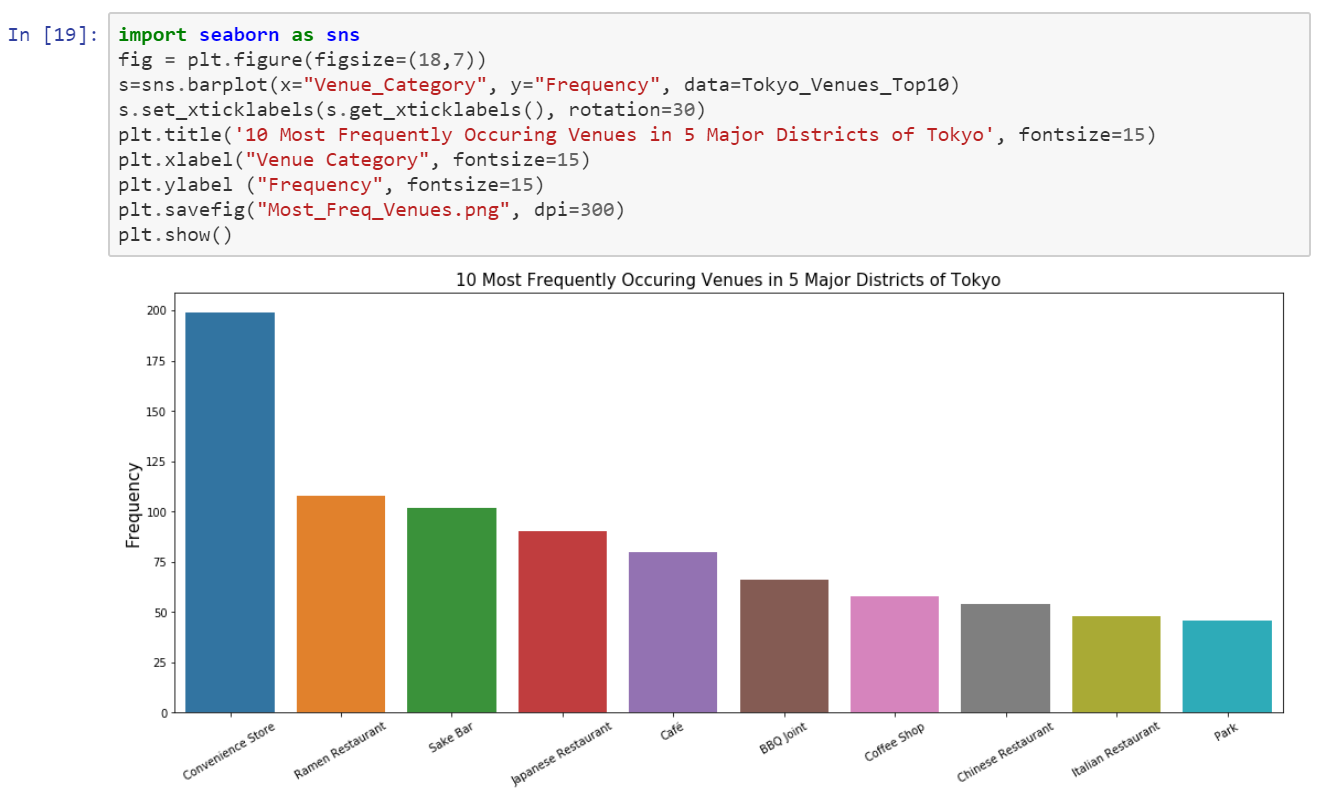
Using Folium, I show the venue locations on a map.



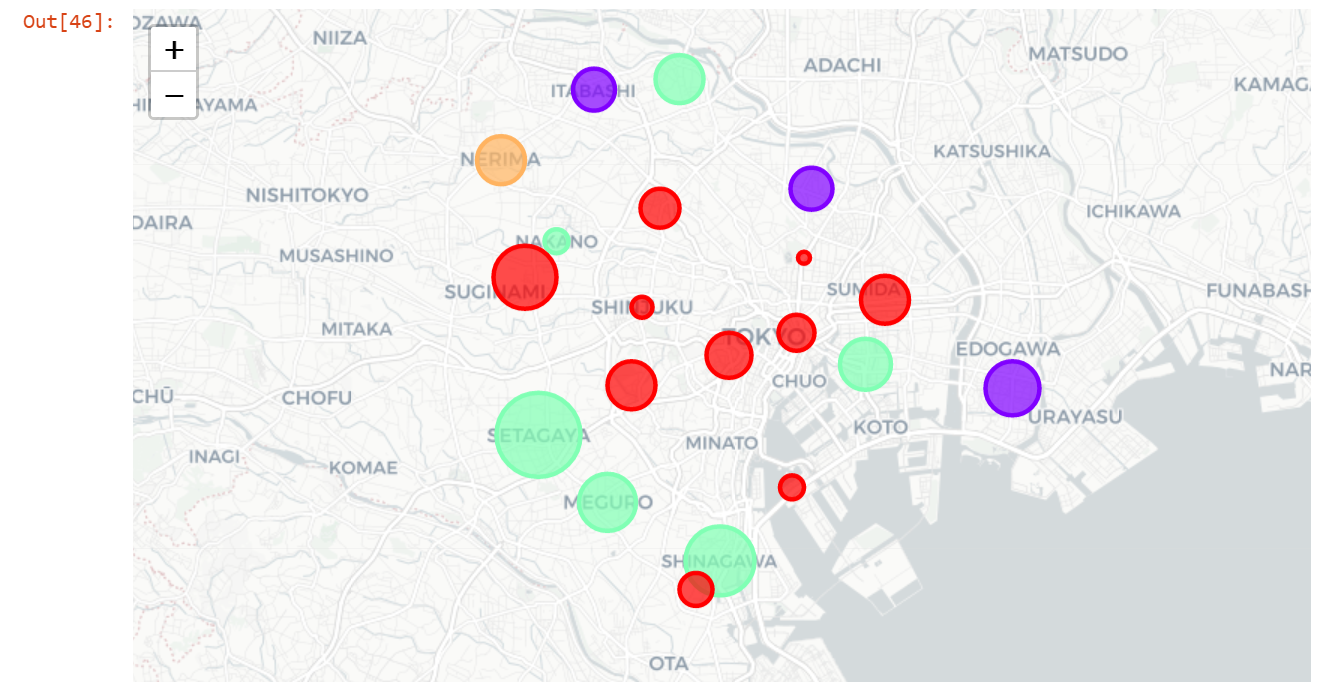
Then I get the top ten venues and their frequencies.



Using seaborn to show the results in a chart



Using Kmeans 5 clusters



Results

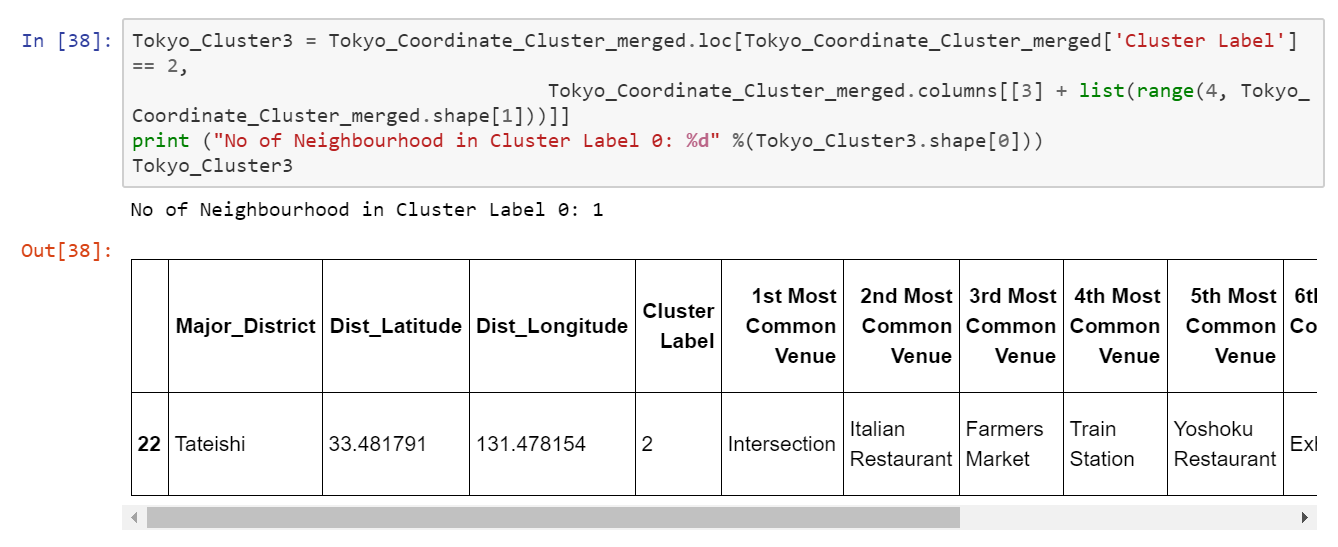
Cluster 1 is popular for Japanese restaurants



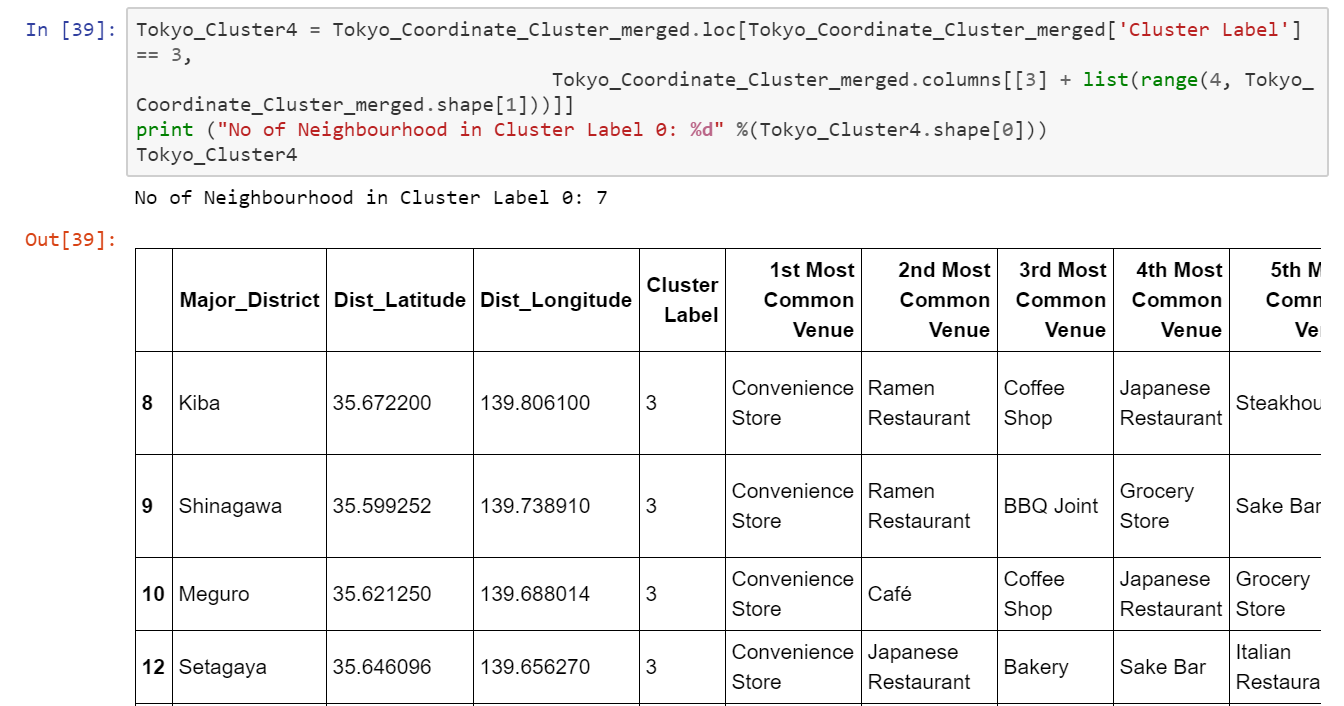
Cluster 2 has convenience stores and grocery stores



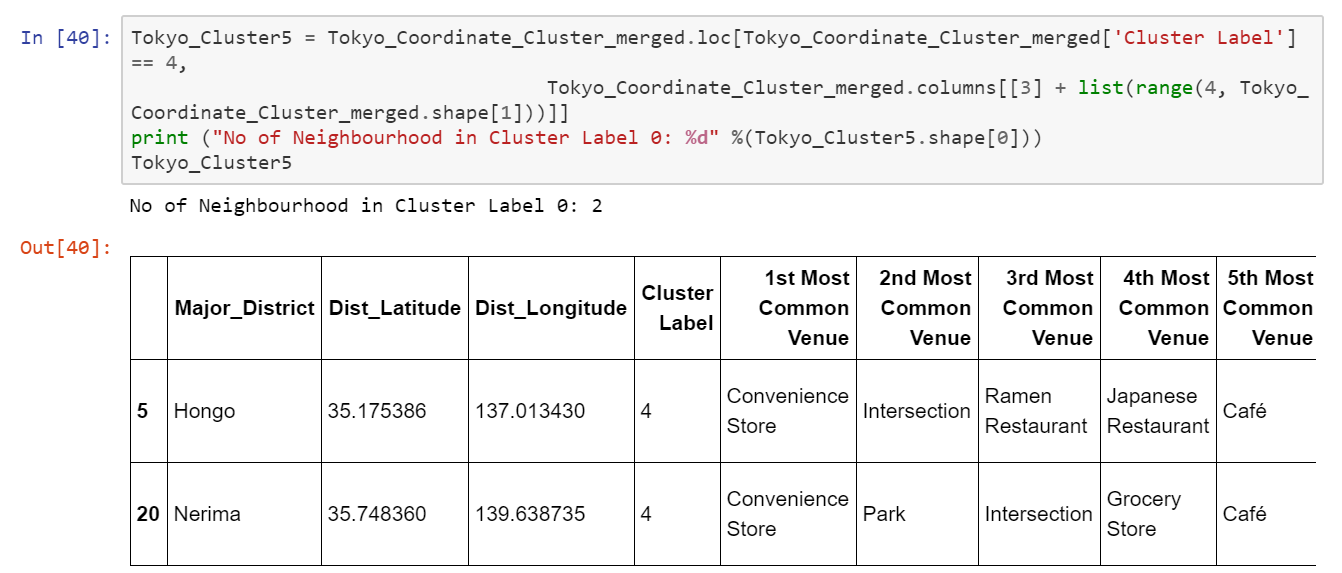
Cluster 3 is outlier with intersection and train station



Cluster 4 has convenience stores and ramen restaurants



Cluster 5 has convenience stores and cafes



Conclusion

Tokyo is a big city with wide variety of locations and each one is different with some similarities. It is great to be able to use technology to analyze the location data using different techniques to show the similarities and differences of locations which would be otherwise difficult without the use of these technology.