Introduction:

'Hypothetical analytics company' is focusing on providing consulting services for people who intend to buy a property. Due to large market competition, 'Hypothetical analytics company' has adopted data science techniques to provide the best recommendation system for its clients.

Business Problem:

A couple wants to relocate from Forest Hills, Queens, NY to Staten Island, NY due to the job requirement and they would like to keep similar lifestyle as they had at Forest Hills, Queens, NY. In addition, 'Hypothetical analytics company' should recommend similar neighborhood that is similar to the Forest Hills neighborhood. Moreover, a neighborhood recommendation is based on the two criteria that were specified by a couple. The following two criteria outlined by a couple are listed below. A couple would like their new neighborhood in Staten Island to have lots of Japanese Restaurants. They would also like their new neighborhood in Staten Island to have lots of parks.

Data:

To solve the task of relocation from Forest Hills, Queens, NY to Staten Island, NY and to provide the best recommendation of the neighbourhood at Staten Island, NY as per the specified criteria:

A couple would like their new neighborhood in Staten Island to have lots of Japanese restaurants.

They would also like their new neighborhood in Staten Island to have lots of parks.

We will use 2014 New York City Neighborhood data and Foursquare information on venues to tackle this task. New York City Neighborhood data can be accessed through: https://geo.nyu.edu/catalog/nyu_2451_34572

Methodology:

Summary:

To approach a solution for our task, we will use the Foursquare API to explore neighborhoods in Staten Island. In addition, we will use the 'explore' function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters. Moreover, we will use the k-means clustering algorithm to complete this task. Finally, we will use the folium library to visualize the neighborhoods in Staten Island and their emerging clusters.

Procedure:

Segmenting and clustering neighborhoods in Staten Island.

Explore neighborhoods in Staten Island by accessing Foursquare API data to explore venues of each neighbourhood.

Analyze Each neighborhood in Staten Island by using one hot encoding technique to get top 5 most common venues of each neighbourhood.

Apply k-means clustering algorithm to cluster all neighbourhoods into 5 clusters.

Filter and check each cluster for Japanese restaurants and parks. It turned out that only one of all of the neighbourhoods in Staten Island has Japanese restaurants and parks being one of their top venues.

Analyze and double verifying each neighbourhood in Staten Island for Japanese restaurants and parks. Again it has been confirmed that only one of all of the neighbourhoods in Staten Island has Japanese restaurants and parks being one of their top venues.

Results and Discussion:

Analyzing and exploring our client's requirements that are listed below:
New neighborhood in Staten Island needs to have lots of Japanese restaurants.
New neighborhood in Staten Island also is reqiered to have lots of parks.
We can clearly observe the following:
Only Castleton Corners neighborhood of all of the neighbourhoods in Staten Island has Japanese restaurants and parks being one of their top venues. (Japanese restaurant being 2nd Most Common Venue at Castleton Corners neighborhood and Park being 4th Most Common Venue at Castleton Corners neighborhood)
After careful analysis of the neighborhoods in Staten Island based on the client's requirements, Castleton Corners neighborhood should be recommended to our client.

Conclusion:

The result of our findings has some limitations. The venues used in this project are from the top 5 most common venues of each neighbourhood in Staten Island, which may neglect the neighbourhoods with more venues but less frequency of Japanese restaurants and Parks.
Moreover, we have used only 500 meters radius for searching the venues. There might me more venues available if our radius for searching the venues was larger but then that would affect the fact whether or not our client has a car or if there is a public transportation available to those venues.
In conclusion, Castleton Corners neighborhood in Staten Island, Queens, NY meets all of our client's requirements but better choices may exist as well.