Capstone Project
The Battle of Neighborhoods

Introduction

A Japanese firm is looking for a location to setup their office in Bangkok. They specifically want a location near by a BTS station and would like a recommendation list of BTS stations that would be interesting.

Data

On this project, location data of all BTS station is needed to get venues around each BTS station using Foursquare. This will be prepared as a csv file because it is not possible to get accurate location coordinates of the stations given its name.

Methodology

First, a list of top venue categories of each station was created and shown with the frequency of each venue category. BTS stations were clustered using K-means clustering algorithm into 7 clusters using the list of top venues. The reason behind using K-means clustering algorithm is that the data needed to be separated into groups without any expected result.

Results

As the results, there are 7 clusters of BTS stations shown in the notebook.

Discussion

It was quite obvious that many BTS stations fall into the same group as the diversity in the area is not very high. It also turned out cluster number 5 and 7 are quite interesting with leading numbers of coffee shops ,Japanese restaurants, and convenient stores.

Conclusion

Even though Japanese restaurants, coffee shops, shopping malls, and convenient stores are very common in the city of Bangkok. We were still able to cluster the neighborhoods near BTS stations into different groups with different characteristics. Cluster number 3 turns out to be most suitable for Japanese firms that are willing to setup a new office near BTS stations in Bangkok if they are looking for places with Japanese restaurants, coffee shops, and shopping malls.