
Searching best place for next shop for coffee shop owner
Yoshito Nagahama

Background

In this presentation, the person who has following characteristics.

- Successfully running one local coffee shop.
- The shop is located near a one of the station of train line in central Tokyo, Japan.
- The person is considering to open second shop to extend his/her business.

Issue and Objective

The owner need to find right place for business for the second shop.

Strategies for selecting the place are as follows.

- Similar environment to the first shop so as to utilize know-how obtained through the experience in the first shop.
- Along same train line to the first shop.

Data analysis and machine-learning were employed to find the solution to the issue.

This presentation describes the result of study.

Data

Set of latitude and longitude of stations around central of Tokyo are used for this study.
Targeted lines are following 2 lines which located around central of Tokyo.

- Keihin-Tohoku Line
- Negishi Line

Latitude and longitude are collected from https://www.odekakemap.com/ and merged into csv file('train_station_around_Tokyo.csv') for this study.

The nearest station from the first shop is **Akabane** station in Keihin-Tohoku Line

Methodology

The main processes for the study are as follows.

1. Obtaining the shop information around each station in the targeted train line via FourSquare API
2. According to the shop information, clustering of the stations are executed via machine-learning algorithm.
3. Based on the result of clustering, find similar stations to the nearest station from the first shop.

Clustering via machine-learning are conducted to find the place similar to the place where the 1st shop is located.

K-means algorithm are employeed for the clustering because it is very common for clustering via machine-leaning.

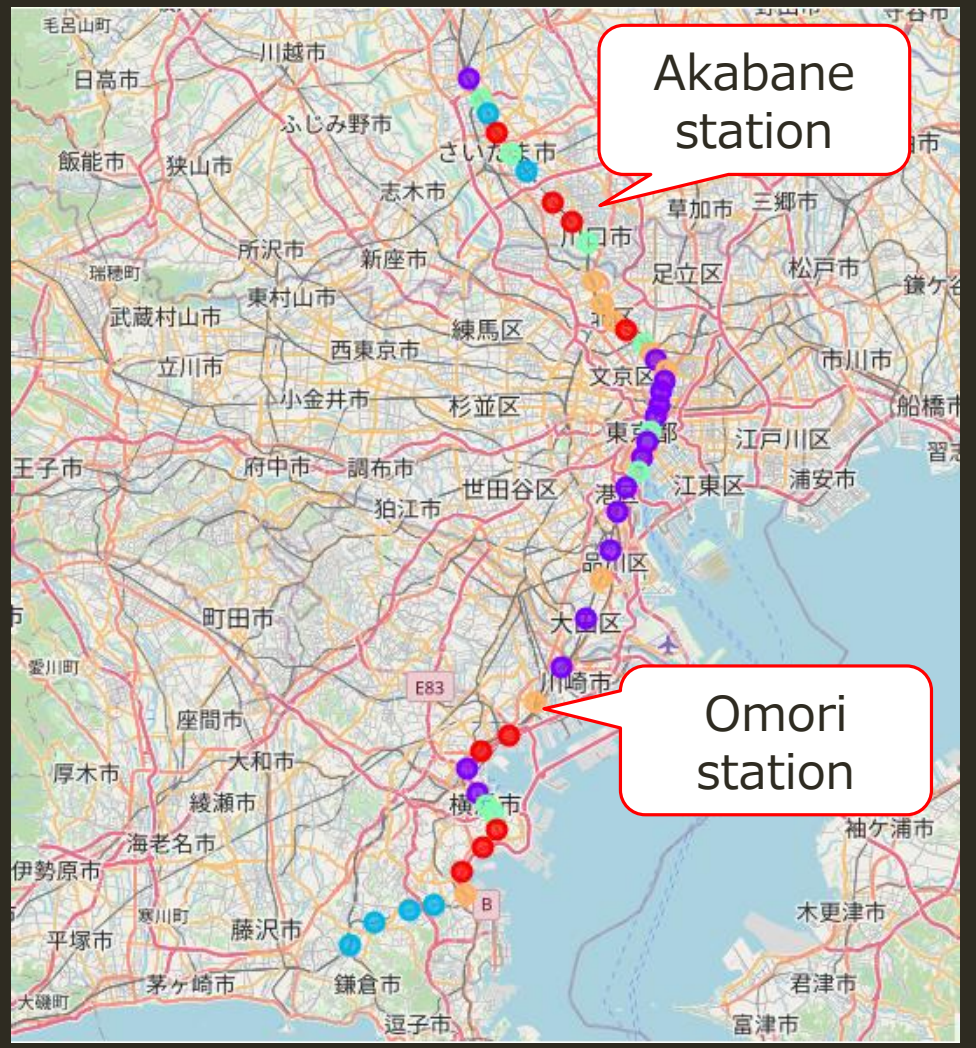
Result and Discussion

Stations which have similar character to Akabane station were suggested by k-means clustering.(Stations marked as orange)

These stations are possible candidate for the second shop.

For further study, similarity from Akabane station is calculated by Euclidean distance.

Omori station has the highest similarity to Akabane station, which implied that Omori station would be the best place.



Conclusion

To find the best place for the next coffee shop to be opened, data analysis were conducted employing machine-learning.

According to the study, following thing are suggested.

1. Stations which have similar characteristic were suggested by K-means clustering. They will be possible candidate for the second shop.
2. Especially, Omori station would be the best place for the second shop according to similarity evaluation based on Euclidean distance.

Link

- Jupyter Notebook for this study.

https://github.com/Yoshito-Nagahama/Coursera_Capstone/blob/ccbea4aeae1bcab4ea32a5c50aa2ce3d86657a76/Final_assignment.ipynb

Thank you!