

IBM Applied Data Science Capstone

Opening a New Movie Theatre in Tokyo

Dariush Safaee - May 2020

Business Problem

The objective of this research is to find the best location to open a new movie theatre in Tokyo, Japan.

In 2019, the global box office reached a new high of over \$42bn. Although the number of tickets purchased is dropping in areas like North America, sales are still rising overseas, especially in Asia.

This creates new opportunities for opening a successful movie theatre in one of the most vibrant cities of the world, Tokyo.

The target audience for this research is primarily entrepreneurs looking to open a new business, and existing cinema companies looking to expand in a new area.

Data

From this Wikipedia page, I have gathered a list of 84 distinct neighbourhoods in Tokyo.

Using the Python Geocoder package, I have identified latitude and longitude coordinates for each of the neighbourhoods.

Using the FourSquare API, I have retrieved data of the locations of movie theatres and Italian restaurants.

Methodology

Main steps:

1. Build neighbourhood list retrieving data from Wikipedia
2. Get coordinates of each neighbourhood using Geocoder
3. Get data of movie theatres and Italian restaurants using FourSquare API
4. Group data by neighbourhood and take the mean of the frequency of occurrence of each venue category
5. Perform machine learning K-means clustering technique
6. Visualise clusters in a map using Folium

Results

Neighbourhoods grouped into three clusters:

1. Red: Little to no venues
2. Purple: Highest concentration of venues
3. Green: Relatively low number of venues



Discussion

The green cluster contains areas with a low concentration of venues. There are various neighbourhoods with some Italian restaurants, with little to no movie theatres.

These represent possible places to consider where opening a new movie theatre could be successful.

We can see that in the cluster with the highest concentration of venues, the purple one, there is one neighbourhood, Hatsudai, with a very strong concentration of Italian restaurants but no movie theatres. This represents the best area to open a new cinema.

Conclusion

This research consisted of:

1. Identifying a business problem
2. Gathering the data required to perform the analysis
3. Preparing the data to be ready for modelling
4. Performed K-means clustering to spit data into clusters
5. Analysing the results to find a solution to the business problem

We have come to the conclusion that the neighbourhoods in the green cluster are the best areas to open a new movie theatre. However, Hatsudai, in the purple cluster, is likely the best area to open a new one.