

# **Battle of Neighborhoods**

**Travel Recommendations for short-trip travelers to  
Seoul, South Korea**

**Shapna Muralidharan**

## **Contents:**

- Introduction
- Project Description
- Data Sources
- Methodology for the project
- Results and Discussion
- Conclusion

## **Introduction**

Seoul, is the capital of South Korea, is a huge metropolis where modern skyscrapers, many modern buildings and high-tech subways and famous k-pop culture. There are many Buddhist temples, palaces of old times and many street markets like Namdaemun, Dongdaemun and many more. There are many Notable attractions like Dongdaemun Design Plaza, Gyeongbokgung Palace, and many more tourist spots. There are many people who visit Seoul on a short trip and cannot visit all the neighborhood places they wish. Seoul is the principal tourist destination for visitors. The number of international tourists to Korea for this year is expected to reach a record 17.5 million amid efforts to draw more independent Chinese travelers and diversify markets. According to the Korea Tourism Organization (KTO), 16.05 million foreign tourists visited Korea between January and November, and an additional 1.45 million are expected to travel here by the end of the year. The previous annual record for inbound tourists is 17.24 million set in 2016. Taking a closer look,

the record number means one international traveler visits Korea per 1.8 seconds and 118 fully occupied 407-seater airplanes land daily over the course of a year. There are a number of short trip travelers to Seoul who are attending conferences and taking business trips. This project is to make an effort to help these travelers and help them get the best of experience in Seoul.

## **Project Description**

This project demonstrates an analysis of venues in Seoul, South Korea using heterogeneous data sources including data science methods. The process includes extraction, load, transformation and analysis of all data sources from Foursquare. Since Seoul is a busy city with high density of population there are lots of venues in Seoul which has a high probability of places to visit. In this project we will try to segment venues into Clusters of the Seoul city based on geographical location. I am particularly interested in helping people who will be visiting Seoul and want to explore the city without any chaotic situations. Furthermore, I would also prefer popular locations as there will also be people on a short-

trip or visiting for conferences, or visitors from nearby countries. Seoul has a high inflow of visitors from neighboring countries like China. I will try to use our data science capabilities to generate a few most promising neighborhoods based on these criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by visitors to Seoul.

In this project we will try to help individuals travelling to Seoul by recommending venues segmented in different clusters. Specifically, this report will be targeted to travelers who are going on a short trip interested in exploring different Venue in Seoul, South Korea.

## **Data Sources**

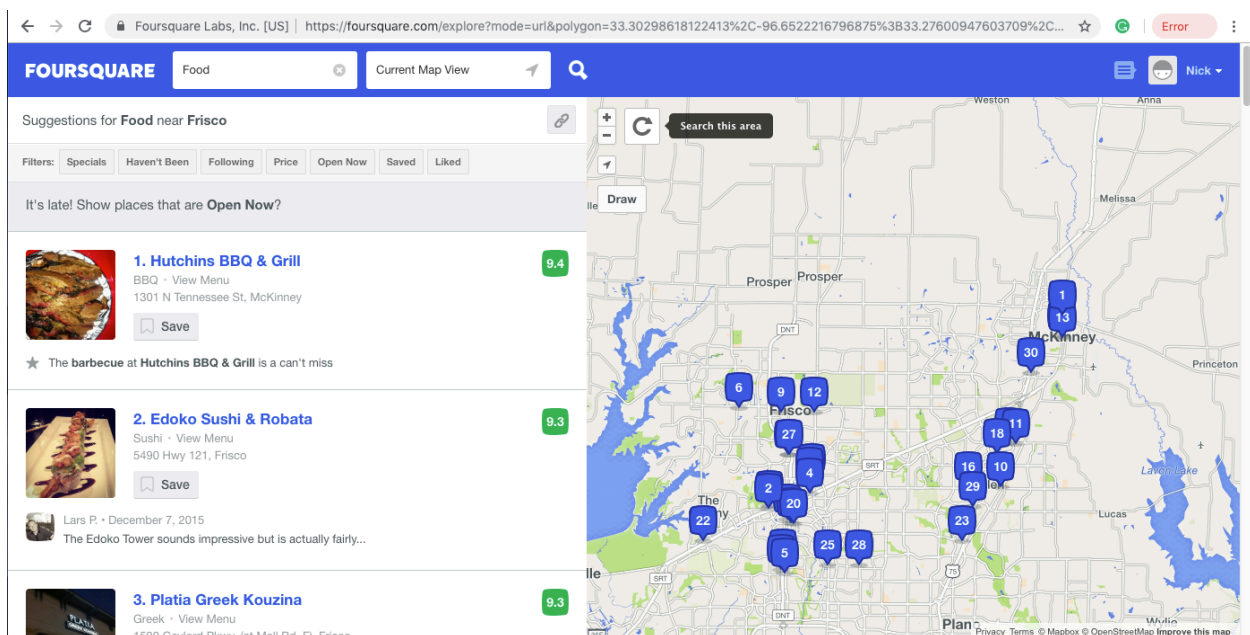
### **Data Sources for the project:**

Foursquare - a location technology platform that provides various API's under the parent 'Places API' to help us fetch data. There are numerous API's for venues, users, photos, check-ins', list. In this project I have used FourSquare two API's.

#### **1) Search for Venues**

## 2) Get Venue Recommendations

Now once we got the districts' latitude and longitude, let's use Foursquare Location to get the best venues from Seoul which will give us an idea of where the tourist can visit in Seoul.



## Data Cleaning:

Foursquare API's used in this project return data in JSON format. These API's mainly return the fields like Venue Name, Address, Latitude, Longitude and Subcategory. While exploring the project

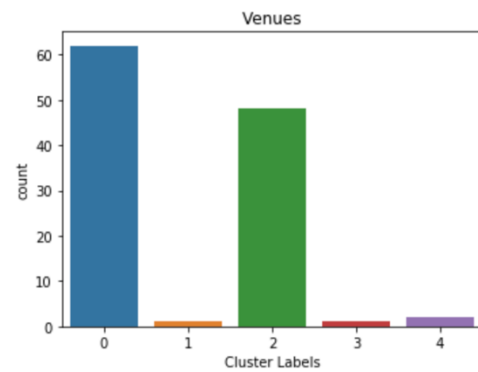
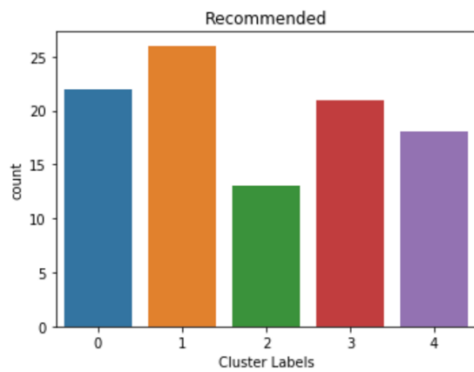
category function is used to add category column to the data frame. Two data frames are made using the above API's to form recommended and venues respectively. Also, none of the columns have any missing values in any of the two data frames.

## **Methodology for the project**

For the methodology of the project first we tried finding the most number of venues and recommended venues in cluster using the foursquare API. This is based on the previous reviews from the travelers who have visited the places earlier. Next we segmented all the venues in the cluster to find recommended restaurants, cafes, lounges and also hotels because these are the places frequently visited by short-trip travelers while sightseeing.

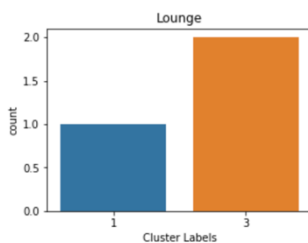
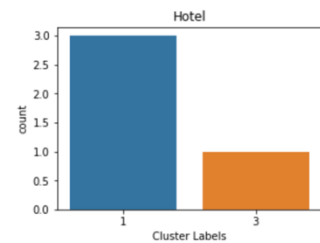
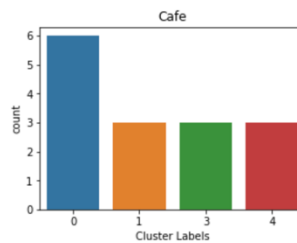
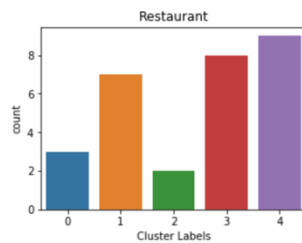
## **Results and Discussion**

**Plot to find the most number of venues and recommended venues in cluster:**



From the plot we can see that cluster 1 has the most number of recommended venues and cluster 0 has the most number of venues.

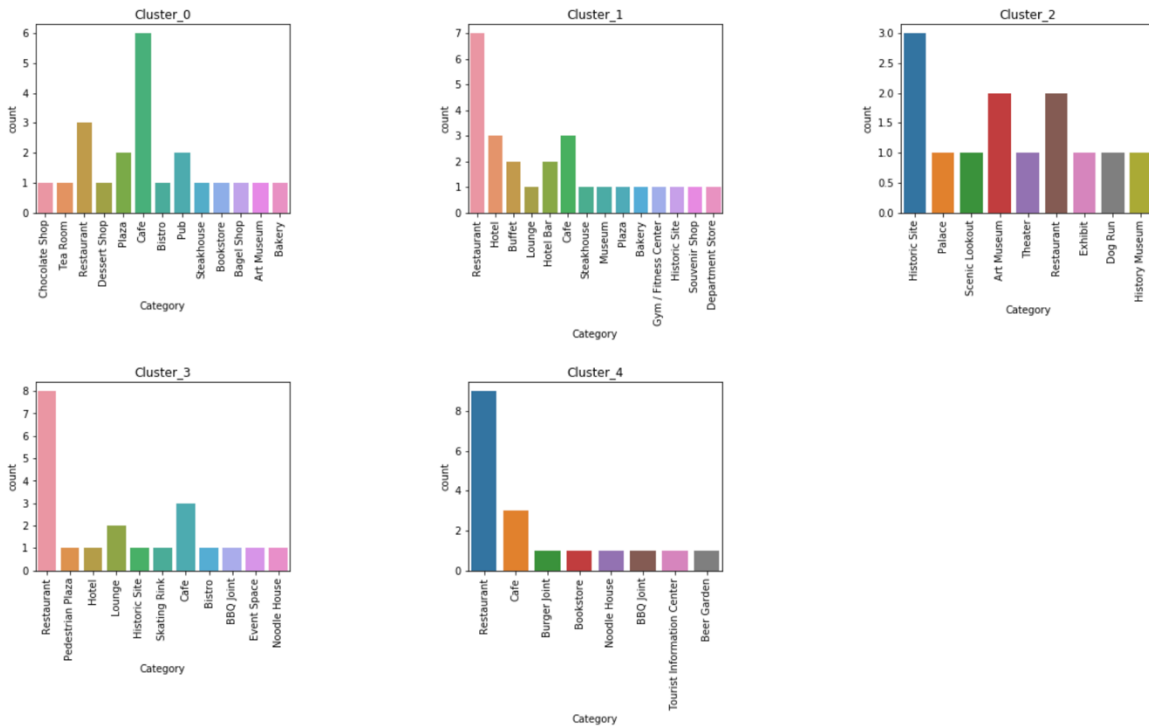
**Plotting to show the recommended venues across different Clusters:**



From the plots we can visualize there are more recommended restaurants in Cluster 4, more cafes in Cluster 0, more hotels in Cluster 1 and more lounges in Cluster 3.

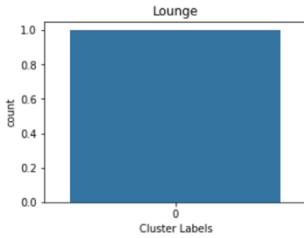
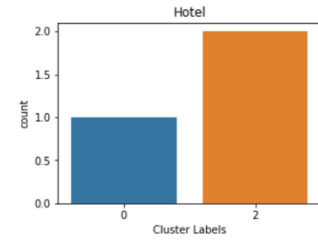
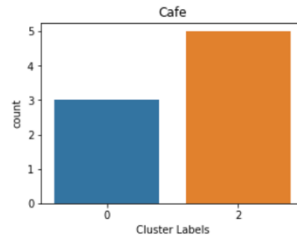
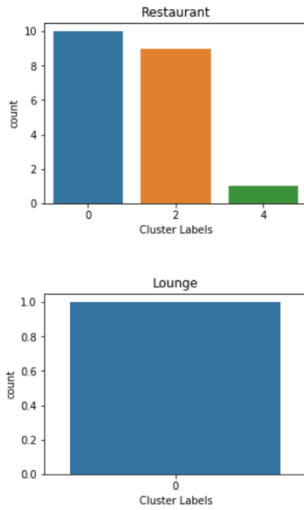


## Plots to visualize recommended places to visit in each cluster:



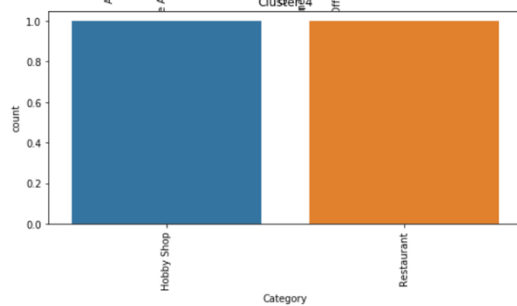
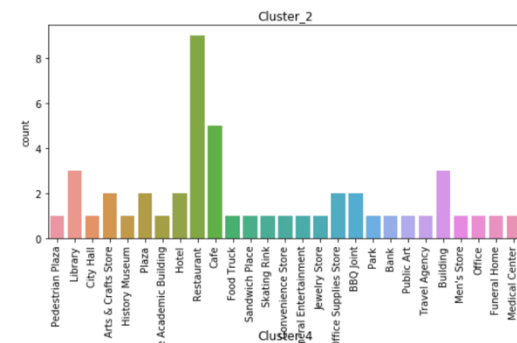
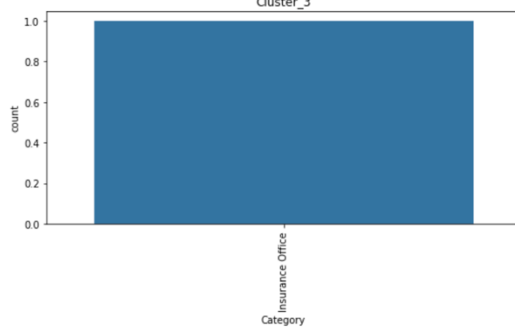
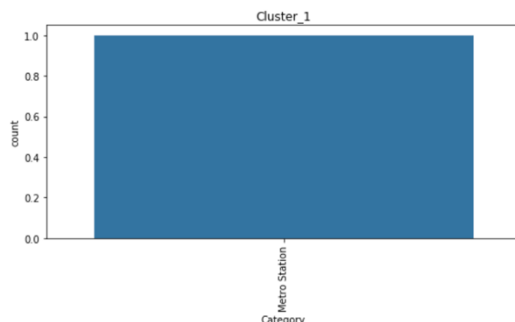
We can see Cluster 2 has most recommended venues followed by Cluster 0 and Cluster 1.

## Segmenting all venues in all the clusters:



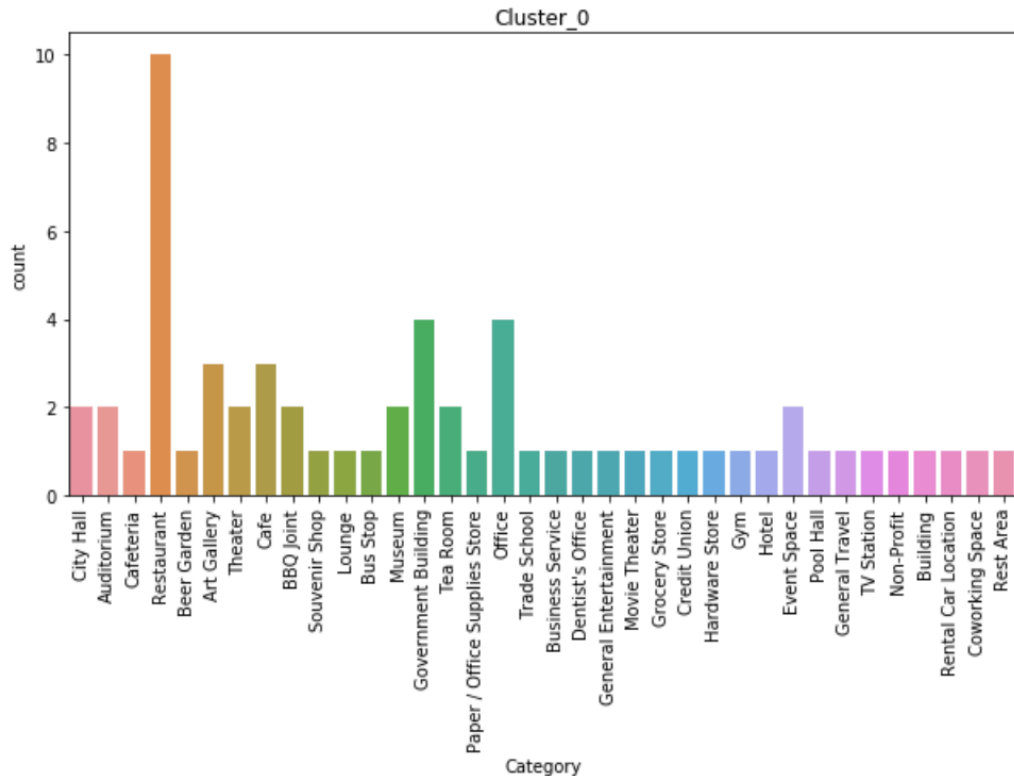
**Cluster 0 Has the most number of restaurants, cafes, hotels and bars, while cluster 4 follows it**

**Visualize various venues in all the Clusters:**



**Cluster 0 has the most number of venues in all the clusters while cluster 2 follows behind**

**Cluster 0 Detailed view:**



## Results and Discussion

Results and Discussion: Based on our analysis to suggest recommended venues for short-trip travelers to Seoul, South Korea we have the following findings.

- Cluster 1 has the most number of recommended venues and cluster 0 has the most number of venues to visit

- We can visualize there are more recommended restaurants in Cluster 4, more cafes in Cluster 0, more hotels in Cluster 1 and more lounges in Cluster 3.
- Cluster 0 Has the most number of restaurants, cafes, hotels and bars, while cluster 4 follows it
- Cluster 0 has the most number of venues in all the clusters while cluster 2 follows behind
- Cluster 0 has the most number and variety of venues with also a number of hotels, restaurants, cafes and lounges.
- So short trip travelers can experience the real Seoul in Cluster 0.

## **Conclusion**

The main objective and Purpose of this project was to identify good recommended venues for short-trip travelers to visit in Seoul. This can narrow down the search time for travelers and also give them the best experience. The Foursquare Data has helped us in identifying the recommended venues visited by the earlier travelers and customers. Creating clusters among these venues helped us to segment and segregate the whole of Seoul

area and provide a easy view for the travelers to decide. The final decision taken by the travelers can include the type of neighborhood and additional factors like commuting, prices and their lifestyle.