

# Choosing Restaurant Location

Saad Elshazly

July 12, 2020

## 1. Introduction

### 1.1 Background

Riyadh city has many neighborhoods, each neighborhood has multiple venues, the venues are categorized where multiple venues falls into the same category. Venue Category may be a Coffee shop, Gym or Restaurant and so on.

### 1.2 Problem

If you need to start your business and you don't know the type of business that is favorite to people and what is the location that will make it successful, so it is not easy mission to choose the previous criteria, in this case you have to choose the best business type as well as the best location start it.

### 1.3 Interest

Our analysis focuses in restaurant Business, this services people who are interested in restaurant business and they want to know the most attractive restaurant categories based on neighborhoods. We will focus on about 19 neighborhoods of Riyadh city as follow:

- |   |   |
|---|---|
| 1. Zahrat al Badiyah District <span>ظهرة البديعة</span>   | 11. Ghirnatah District <span>حي غرناطة</span>       |
| 2. Al Aziziyah District <span>حي العزيزية</span>          | 12. King Fahd District <span>حي الملك فهد</span>    |
| 3. Al Nahdhah District <span>حي النهضة</span>             | 13. Al Yasmin District                              |
| 4. Al Malaz District <span>حي الملز</span>                | 14. Al Malga District <span>حي الملغا</span>        |
| 5. Al Aqeeq District <span>حي العقيق</span>               | 15. King Faisal District <span>حي الملك فيصل</span> |
| 6. Al Muhammadiyah District <span>حي المحمدية</span>      | 16. Al Olaya District <span>حي العليا</span>        |
| 7. Al Rabiea District <span>حي الربيع</span>              | 17. Al Ghadir District <span>حي الغدير</span>       |
| 8. Al-Ma'athar Al-Shimali District <span>حي المعذر</span> | 18. Hittin District <span>حي حطين</span>            |
| 9. As Sulaymaniyah District <span>حي السليمانية</span>    | 19. Al Murabba District <span>حي المربع</span>      |
| 10. Almuruj District <span>حي المروج</span>               |   |

## 2. Data Gathering and Cleaning

### 2.1 Data Source

We collected location data from foursquare API [here](#) . First we got all categories [here](#), and then we filtered those categories into just food categories. In The next step we collected the all venues for each category [here](#) . To specify each venue with its neighborhood we used [Nominatim](#). Finally we collected more data about each venue [here](#)

## **2.2 Data Cleaning**

Data cleaning started with venue data collected for each category where JSON Data converted into dataset that was having many duplicates, those duplicates has been removed and columns has been reordered where venue id, name, category and has park features comes first then location data such as latitude, longitude, etc.

Features names were not clear, so they renamed. Because we interested in just restaurants we have removed venues data that is not restaurants.

The second stage on cleaning data was data about neighborhoods where we collected them to assign each venue with its neighborhood, so in this dataset we assigned each venues id with a neighborhood name; we removed neighborhoods with less than 13 restaurants, then venue dataset and address or neighborhood dataset have been merged to produce one dataset having venues data plus neighborhoods.

The third step in cleaning our data was venues details, many features removed and just 4 features has been excepted (Id, Verified, Likes, Rating) , null values in rating columns filled with the mean of all ratings, the final details dataset merged with venue dataset by venue Id.

The final step was clean the merged dataset that contain all data we need for analysis, firstly duplicated removed, Verified columns null values filled with mode, rating data rounded up to one decimal value, likes column data type converted from float into integer.

## **3. Explanatory Data Analysis**

### **3.1 analysis overview**

The target from the analysis, as said before, is to help people who intend to open their business as restaurant and after cleaning data we have about 19 Neighborhoods, so ease analyzing data and get our target we will cluster or group neighborhood into 3 groups. Each group will be analyzed to find the most frequent ten venue categories, the best rated and the most liked ones.

We have grouped the neighborhoods into 3 groups by using K-MEAN Clustering Algorithm, Figure 1; each group has set of neighborhoods where groups are colored in the map as green, blue and red for groups 1, 2 and 3 consecutively.

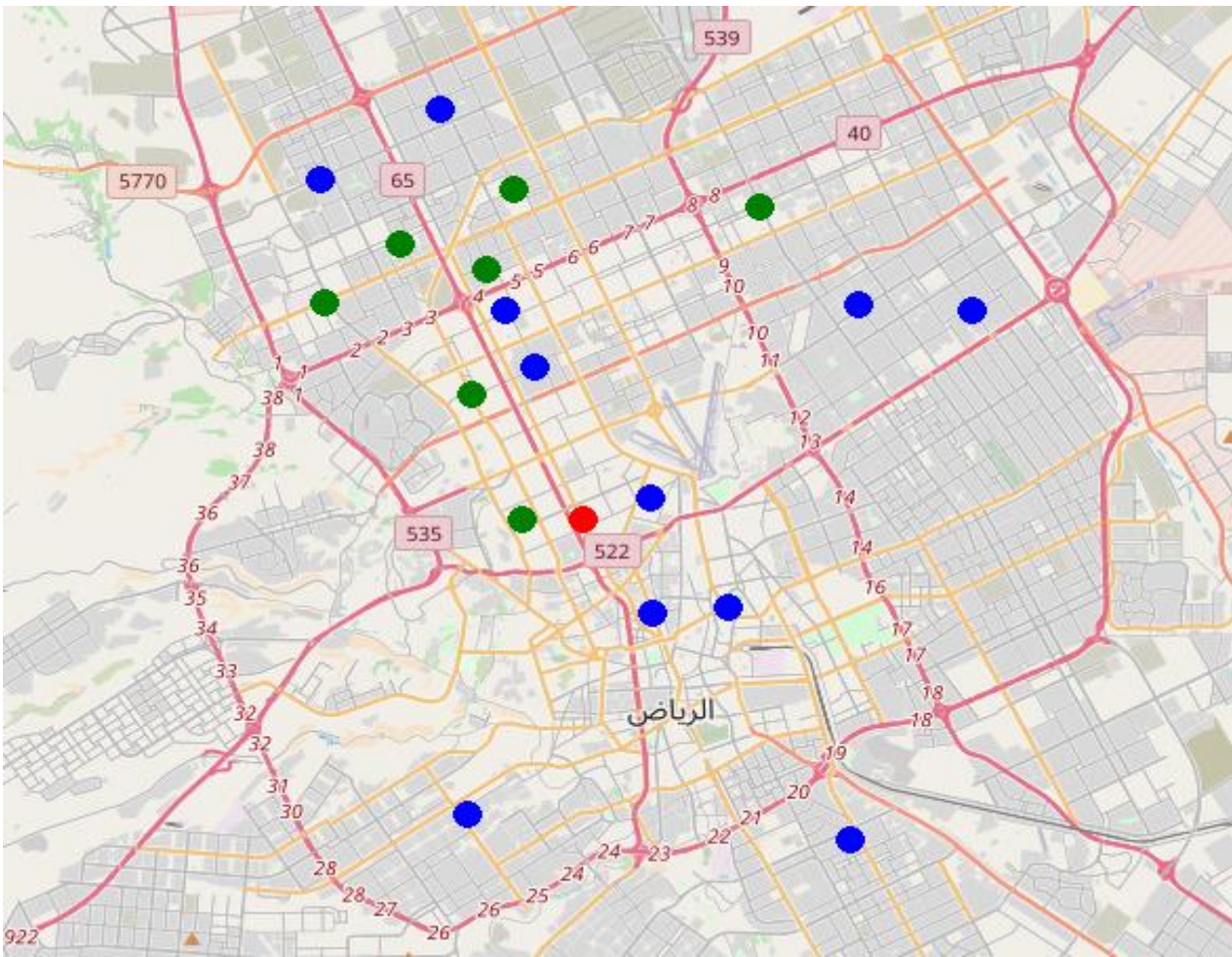


Figure 1. Map to show grouped neighborhoods in Riyadh City

## 3.2 First Group Analysis

### 3.2.1 Overview and Statistics

We have about 39 restaurant Categories with total 167 restaurants in 7 neighborhoods, the most common restaurant category is the French, and also Ghirnatah is the most frequent Neighborhood, most of restaurants in the first group are not verified.

#### 3.2.1 What are the Neighborhoods in the first group?

In figure 2, we see that there are 7 Neighborhoods and Ghirnatah and Al Aqeeq are the most Crowded with restaurants, Al Ma`athar Alshimali and Al Ghadir are less crowded with restaurants.

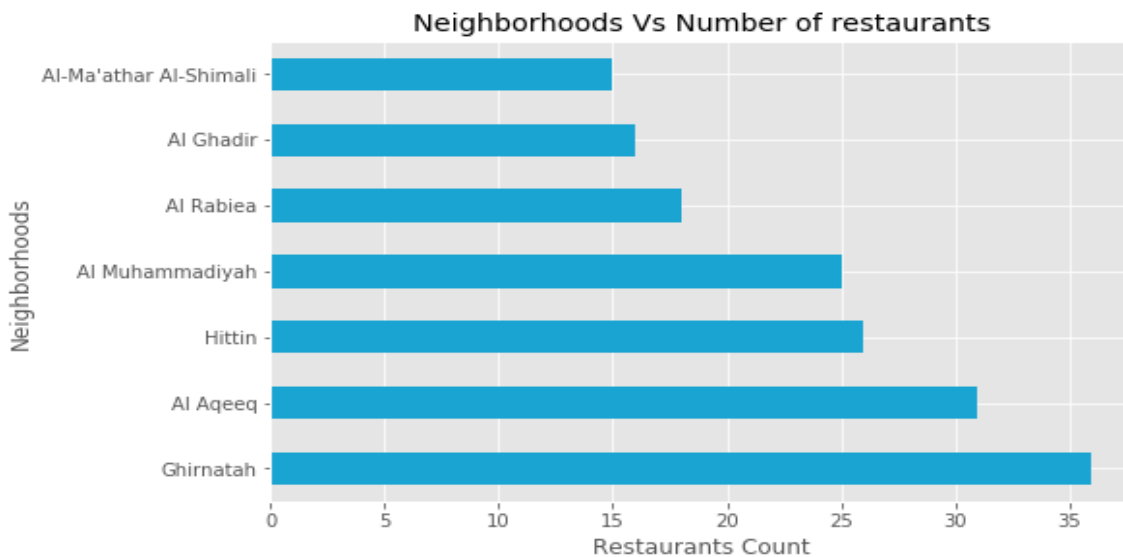


Figure 2. Number of restaurants for each Neighborhood in first group.

### 3.2.2 What are the common restaurant Categories in the first group?

In the word cloud image we see that the most common restaurant type has bigger font size, so we can observe that French, Mexican, American and Indian are the most common categories in the first group.



Figure 3. Show the most common Categories in the first group of neighborhoods

### 3.2.3 What is the most rated restaurant Categories in the first group?

Restaurant rating is a very significant feature in our analysis and it determines how people like and what is the degree of likeness of a specific category, so people in the first group of neighborhoods prefer Lebanese and Japanese restaurants and less preference for American food.

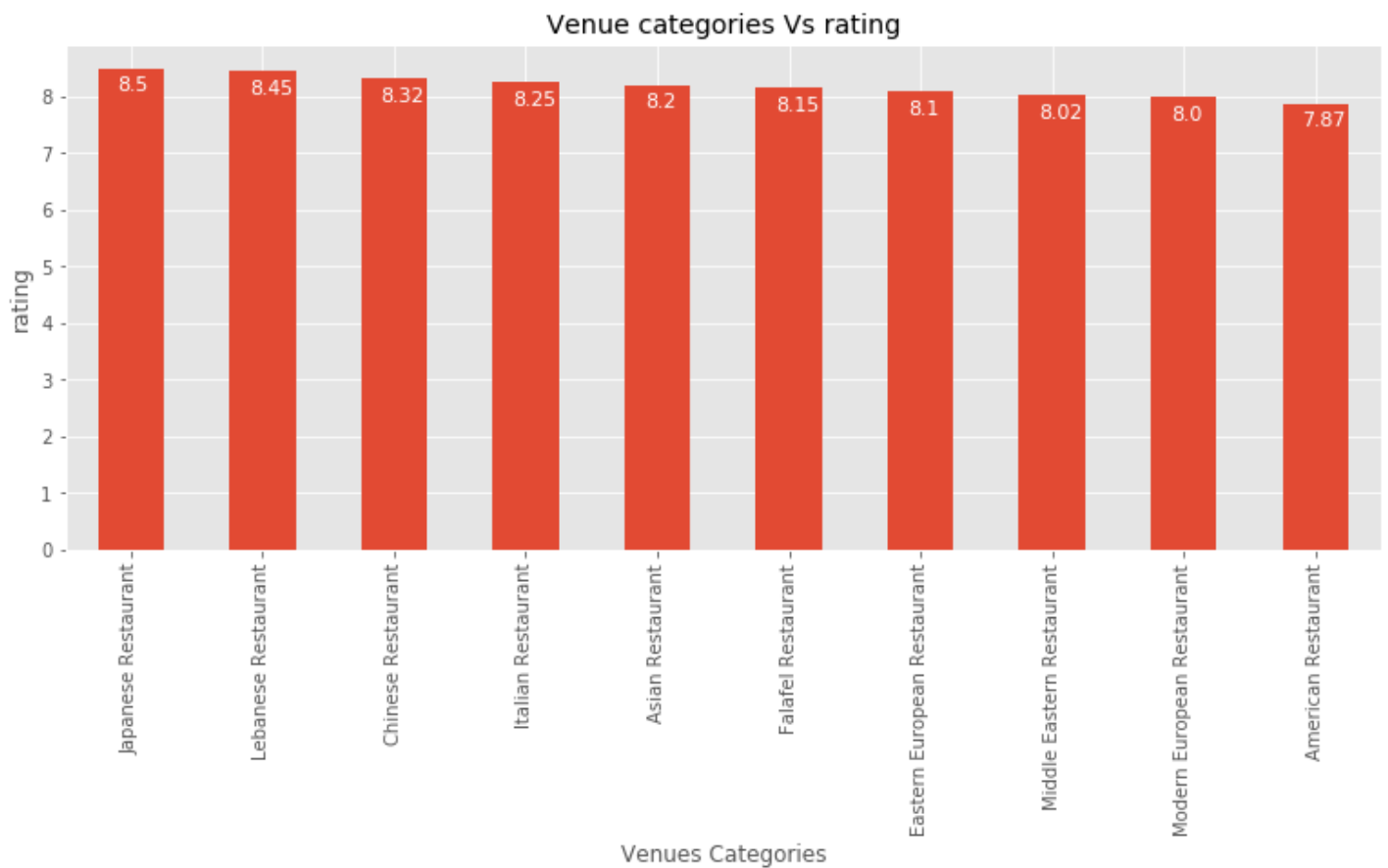


Figure 4. Show the most ten rated categories in the first group of neighborhoods

### 3.2.4 What about rating distribution?

The difference in rating is not big between restaurants in all neighborhoods but there are some outliers that didn't stick to normal rating distribution such as there is a restaurant with about 5.5 rating and in the other side there is 8.5 rating.

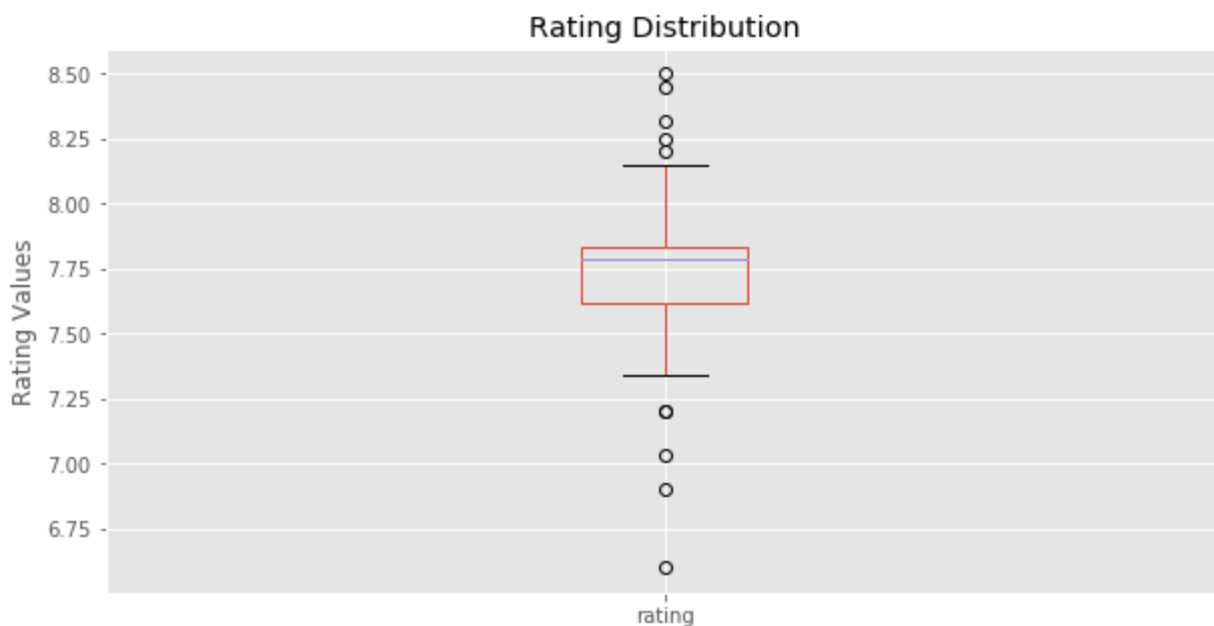


Figure 5. Show rating distribution between all restaurants in the first group of neighborhoods

### 3.2.5 What is the most liked categories in the first group?



Like is a good feature in the analysis but it may be affected by the popularity of the restaurant category, So French and American are the most liked categories and Japanese and Turkish have less likes.

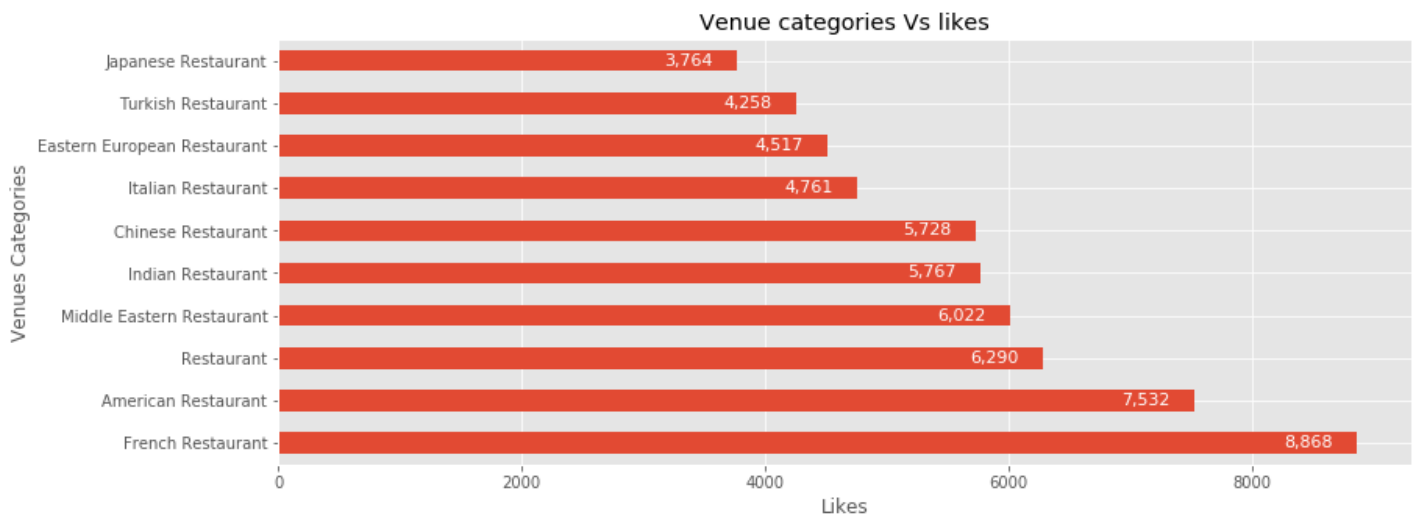


Figure 6. Show the best liked ten restaurant categories in the first group

### 3.3 Second Group

#### 3.2.1 Overview and Statistics

We have about 47 restaurant Categories with total 215 restaurants in 11 neighborhoods, the most common restaurant category is the Kebab Restaurant, and also As Sulaymaniyah is the most frequent Neighborhood, most of restaurants in the first group are not verified.

#### 3.3.1 What are the Neighborhoods in the Second group?

In figure 7, we see that there are 11 Neighborhoods and Yasmeen and Al Sulaymaniyah are the most Crowded with restaurants, Zahrat Al-Badiah and King Faisal are less crowded with restaurants.

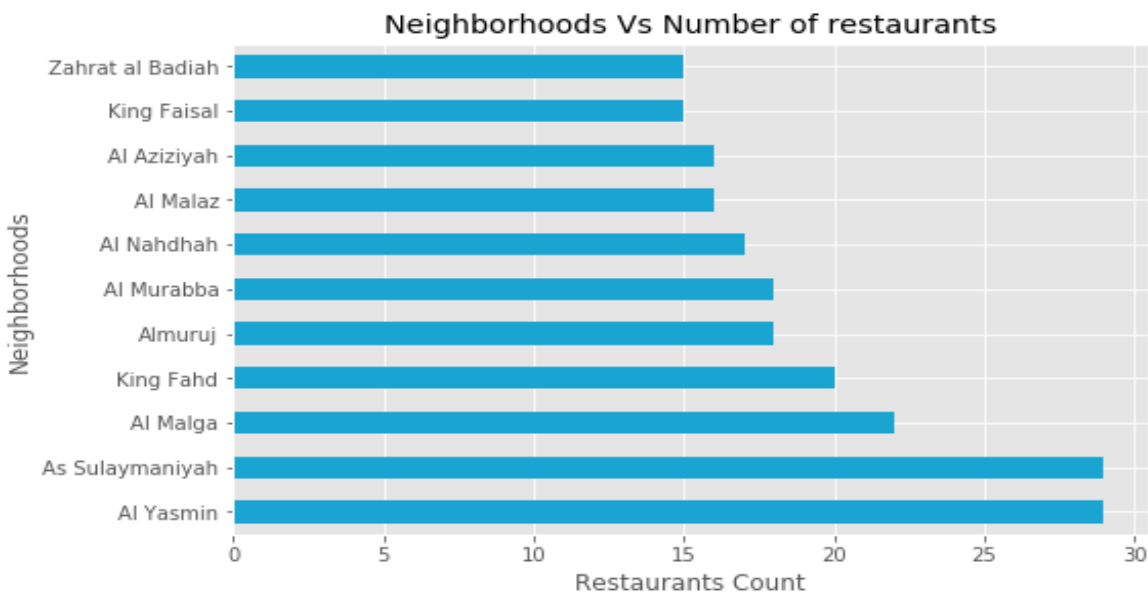


Figure 7. Number of restaurants for each Neighborhood in the second group

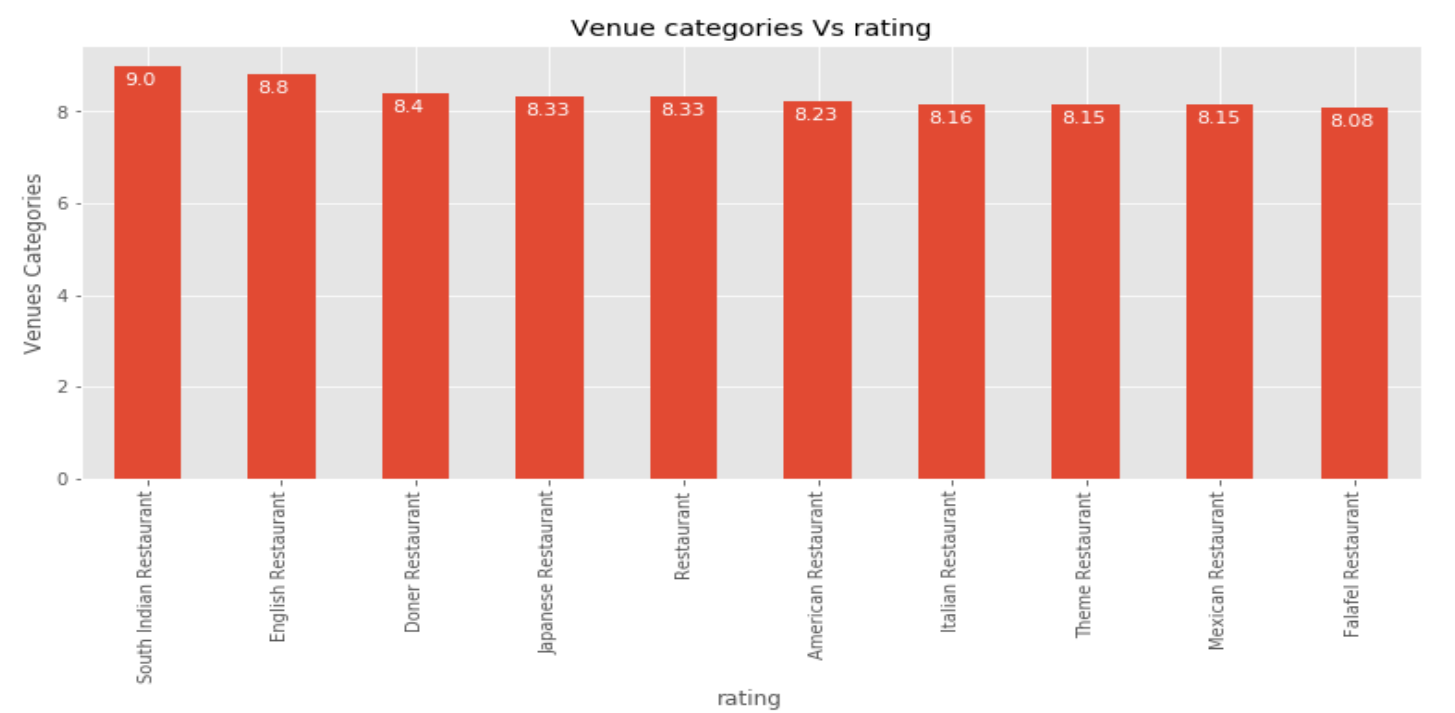
#### 3.3.2 What are the common restaurant types in the Second group of neighborhoods?

In the word cloud image we see that the most common restaurant type has bigger font size, so we can observe that Kebab, Seafood, Falafel and Italian are the most common restaurant categories in the second group.



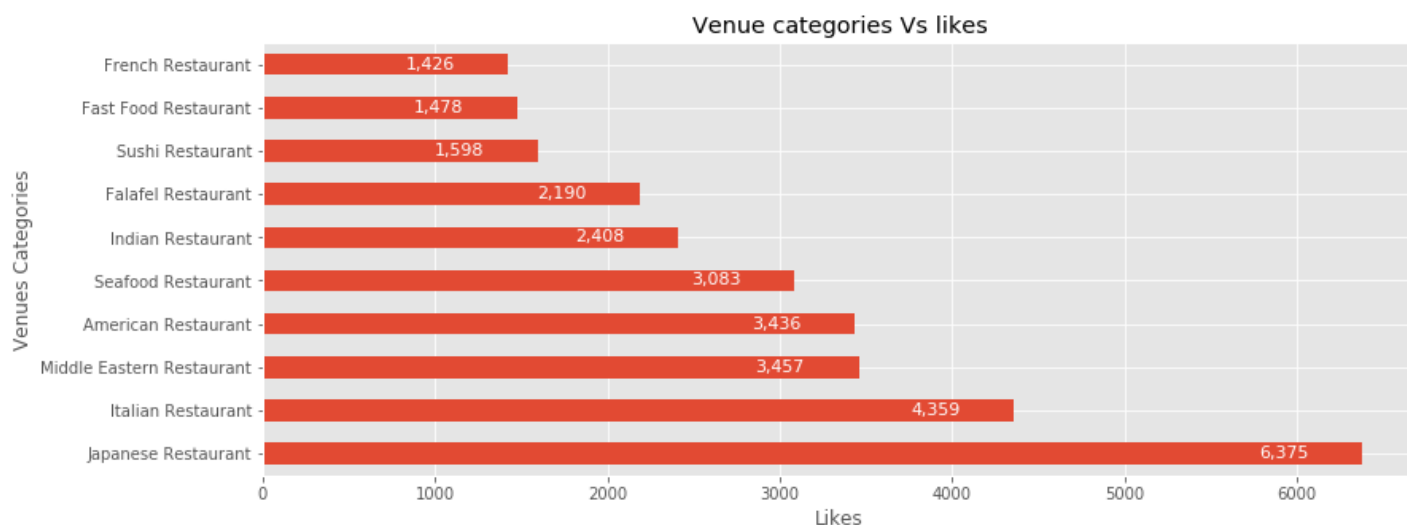
### 3.3.2 What are the best rated restaurants in the second group?

Restaurant rating is a very significant feature in our analysis and it determines the degree of likeness of a specific category, so people in the second group of neighborhoods prefer South Indian and English restaurants and less preference for Falafel restaurants.



### 3.3.3 What are the most liked restaurants?

Like is a good feature in the analysis but it may be affected by the popularity of the restaurant category, So Japanese and Italian are the most liked categories and French and fast-food have less likes.



### 3.3 Third Group

#### 3.2.1 Overview and Statistics

We have about 53 restaurant Categories with total 99 restaurants in 1 neighborhood, Al Olaya Neighborhood, the most common restaurant category is the Italian Restaurant, and most of restaurants in the first group are not verified.

#### 3.4.1 What are the Neighborhoods in the Third group?

There is only one Neighborhood in third group, **Al Olaya**

#### 3.4.2 What is the common restaurant type in the third group?

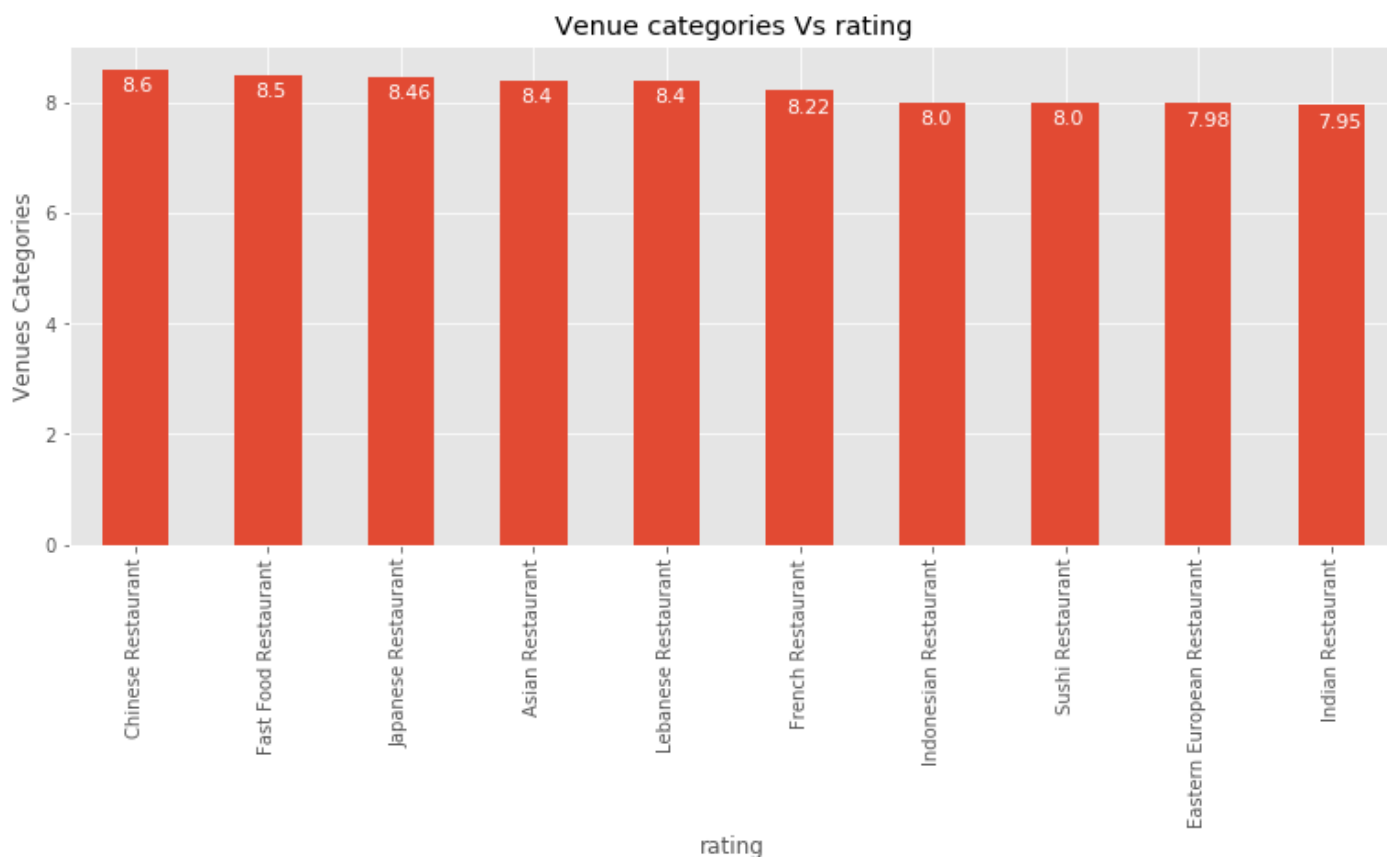
In the word cloud image we see that the most common restaurant type has bigger font size, so we can observe that French, Japanese, Middle-Eastern, Italian and Eastern-European are the most common restaurant categories in the second group.





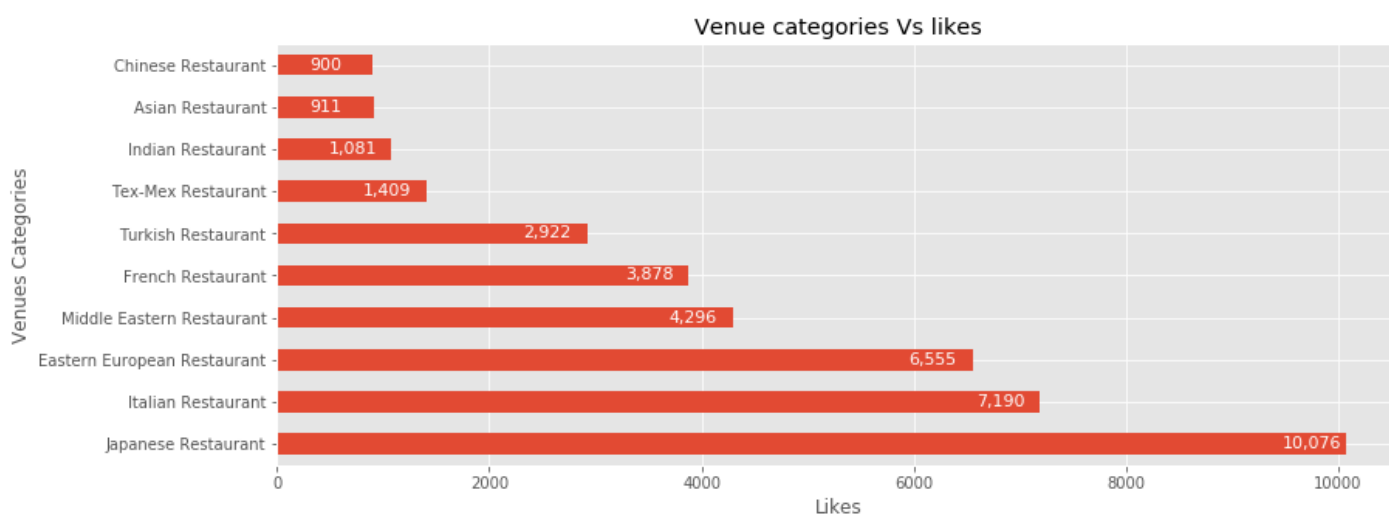
### 3.4.2 What are the best rated restaurants in the third group?

Restaurant rating is a very significant feature in our analysis and it determines the degree of likeness of a specific category, so people in the second group of neighborhoods prefer Chinese and Fast-food restaurants and less preference for Indian and Eastern-European restaurants.



### 3.4.4 What are the most liked restaurants?

Like is a good feature in the analysis but it may be affected by the popularity of the restaurant category, So Japanese and Italian are the most liked categories and Asian and Chinese have less likes, and there is a big diverse in likes between categories in the third group.



## 4. Result

As the above analysis for each group of neighborhoods we can recommend some restaurant categories for each group.

For the first group, there are some restaurants are good rated and liked despite they are not common such as Japanese, Lebanese, Chinese, Asian, Falafel, Eastern European and Modern European Restaurants. So it's good to open a restaurant of the previous categories in any neighborhood of the first group.

For the second group, the best rated and liked and not common restaurants are South Indian, English, Doner, Japanese, American, Theme and Mexican Restaurants. So they are the best choice restaurants to open in the second group of neighborhoods.

For the third group, the best rated and liked and not common restaurants are Chinese, Fast Food, Asian, Lebanese, Indonesian, Sushi and Indian restaurants. So they are the best choice restaurants to open in Al-Olaya neighborhood.

## **5. Conclusion**