**Opening a Sake bar in Lisbon**

# **Introduction**

The business problem consists of finding the right location in the city of Lisbon, Portugal, for opening a Sake bar. According to many available online resources (such as the well-known Asia-based financial website: [https://valueinvestasia.com/restaurants-cluster](https://valueinvestasia.com/restaurants-cluster/)), it makes sense for restaurants to cluster, e.g., being closer to similar restaurants.

In the particular case of Sake bar, we believe that it makes sense to be close to Japanese and Sushi restaurants, which already attract Japanese or Japanese culture loving clientele. We expect that the Sake bar will benefit by providing before and after dinner drinks to the patrons going to those restaurants, and, in some cases, even provide light meals.

After the successful opening of five Sake bars in Madrid and Salamanca, Spain, our friend Hiro S. Ake contacted to help him find the right location for his first Sake bar in Portugal.

# **Data**

On one hand, very limited data is available concerning the taste of Portuguese residents and visitors regarding Japanese products. We found a venue called Sake bar in Lisbon. However, after visiting the place, we concluded that it was not a proper Sale bar, at least, in the sense defined by our client, in terms of decoration, serving procedures and range of products available.

On the other hand, foursquare provides detailed information on the more relevant venues for the different neighborhoods of Lisbon, and we have access to the average rental price per square meter for properties in the different neighborhoods of Lisbon. Rental costs was also a relevant factor in our client’s decision.

# **Methodology**

This project involved using several tools and collecting information provided by third parties, and it was developed in python in the context of a jupyter notebook, with the support of several libraries, such as Pandas and Numpy.

After importing the supporting libraries, the list of neighborhoods in Lisbon was collected from an external website ([www.visitarportugal.pt](http://www.visitarportugal.pt)). Relevant data was extracted, cleaned and passed to a dataframe with Lisbon neighborhoods, using the BeautifulSoup tool, which provides a high level of convenience when extracting data from an html page.

In the following step, latitude and longitude data was collected for each of the neighborhoods. The coordinates data was obtained using Geopy, a Python 2 and 3 client for several popular geocoding web services. As an input, the name of each city was combined with the name of the city to ensure that the coordinates received corresponded with the right location. The neighborhoods location was plotted in a map using the Folium tool.

With the neighborhoods information available, venues data was obtained from Foursquare. Data received from Foursquare was tested to confirm that venues were being provided to all the relevant neighborhoods (neighborhoods of Alcântara and Benfica were excluded as these are not currently issuing any bar licenses). The list of venue category was also studied to identify relevant categories for this project.

As stated in the introduction section, we were only interested in restaurant related venues. Therefore, venues belonging to a significantly different category were excluded from the analysis. The three most common restaurant categories were identified for each neighborhood and a clustering algorithm was applied to the data. The clustering algorithm is k-means and the number of clusters was set to 3.

Unfortunately, the clustering exercise did not provide any significant conclusion and a more direct approach, applying a dataframe groupby function to determine the neighborhoods with the highest number of Japanese and Sushi restaurant.

Finally, as the rental price was also a relevant factor for our client, we extracted rental price per sqm data for the Lisbon neighborhoods under analysis. The public source we found online included rental prices for residential and not commercial property, however, we assumed, that the price relation between neighborhoods would be the same for commercial property.

# **Results**

As previously mentioned, the clustering exercise was inconclusive, as none of the clusters showed a closer proximity to the characteristics we were looking for. The resulting 3 clusters could be described in the following way:

* Cluster 0 - Mostly meatless restaurants;
* Cluster 1 - Portuguese/Classic restaurants;
* Cluster 2 - Informal dining.

As we were not able to derive a definite conclusion from the clustering exercise, we were forced to directly check which were the neighborhood(s) with the highest number of japanese/sushi restaurant:

| **Neighborhood** | **Jap/Sushi counts** |
| --- | --- |
| Santo António, Lisbon | 4 |
| Avenidas Novas, Lisbon | 4 |

Two neighborhoods lead the list with 4 counts each: Santo António and Avenidas Novas. We then checked the rental price per sqm, and Santo Antonio showed the highest price (5153 euros/sqm), while Avenidas Novas had the third highest price (4040 euros/sqm).

| **Neighborhood** | **Price\_sqm** |
| --- | --- |
| Freguesia de Santo António: | 5153 euros/m2. |
| Freguesia da Misericórdia: | 4917 euros/m2. |
| Freguesia das Avenidas Novas: | 4040 euros/m2. |

# Discussion

Based on the results presented in the previous seciotn, and taking all factors into consideration, our recommendation is that our client opens his Sake bar at Lisbon’s neighborhood Avenidas Novas.

# Conclusion

We are satisfied with the results and the recommendation provided, taking into account the data and time constrains available for executing this project. With more time and resources available, we would to conduct further market analysis and research additional online resources.