Exploring European Restaurant Market

Part of the Capstone Project - The Battle of Neighborhoods

Simon Gross

June 1, 2020

INTRODUCTION

Background

Investing in a new restaurant is often not an easy decision. The decision made is a long-term project and involves a great deal of resources to find the right location, to renovate it, but also to find out what kind of cuisine to offer. Whether or not an investment will be successful is often not visible for a long time, as a restaurant must first build up a good reputation and a constant flow of customers, which can take up to several years. It is therefore always very important to reduce the risk of investing in a restaurant of a particular category in a geographical area. But how can you find out if the dishes offered in a restaurant are appreciated by the locals or if the location of the restaurant is good?

Problem

The first problem, whether people want the provided dishes, is often related to cultural preferences. In southern Italy, for example, people like fish and seafood, while in southern Germany they prefer rustic cuisine. Therefore, an investor must find the right cuisine for the region in which he wants to invest. The second problem of whether a location is a good one is often related to how many competitors there are in that particular neighbourhood or town. Therefore, an investor needs to understand the city's surroundings, where there might be room for another Asian restaurant and where there isn't.

Interest

This market analysis is particularly interesting for large investment companies that want to explore the restaurant market and are not sure where and in what kind of restaurants to invest. This could be a Chinese investor who is not familiar with the European market and does not know the regional preferences. They will obtain detailed information about the European restaurant market. In which region are some restaurants and dishes preferred and some not and would therefore be associated with a high risk.

Goal

The aim of this project is to get a visual impression of what kind of restaurants are preferred in each region of Europe and therefore would have a better chance of success.

DATA

Data Sources

In order to cover the most relevant areas in Europe, the 500 largest cities in Europe will be investigated. A list of the cities together with their country and population can be found [here](). Based on this list of city and country we collect the coordinates through Geopy. These coordinates are then used to find 100 locations within a 1000 meter range through Foursquare.

	City	Country	Latitude	Longitude
0	MOSKVA (Moscow)	Russia	55.750446	37.617494
1	LONDON	UK	51.507322	-0.127647
2	St Petersburg	Russia	59.960674	30.158655
3	BERLIN	Germany	52.517037	13.388860
4	MADRID	Spain	40.416705	-3.703582

Fig. 1 : Head of table after receiving cities from webpage and coordinates from geopy.

Data cleaning

In order to use only relevant data, venues that do not include "restaurant" in their category name are deleted. If the category name only contains "Restaurant", it will also be deleted, as it does not contain any information about the dish served.

```
df2['Venue Category'].unique()
```

```
array(['Italian Restaurant', 'Spanish Restaurant', 'Thai Restaurant',
       'Japanese Restaurant', 'French Restaurant', 'Greek Restaurant',
       'Pakistani Restaurant', 'Ramen Restaurant', 'English Restaurant',
       'Lebanese Restaurant', 'Chinese Restaurant',
       'Modern European Restaurant', 'African Restaurant',
       'Vegetarian / Vegan Restaurant', 'German Restaurant',
       'Sushi Restaurant', 'Tapas Restaurant', 'Seafood Restaurant',
       'Mexican Restaurant', 'Argentinian Restaurant', 'Asian Restaurant',
       'Peruvian Restaurant', 'Paella Restaurant',
       'Mediterranean Restaurant', 'Falafel Restaurant',
       'Roman Restaurant', 'Caucasian Restaurant', 'Turkish Restaurant',
       'Scandinavian Restaurant', 'Swiss Restaurant',
       'Ukrainian Restaurant', 'Alsatian Restaurant',
       'Auvergne Restaurant', 'Portuguese Restaurant',
       'Burgundian Restaurant', 'Corsican Restaurant',
       'Romanian Restaurant', 'Eastern European Restaurant',
       'Hungarian Restaurant'], dtype=object)
```

Fig. 2: An example of a list of categories for restaurants after cleaning.

Data processing

Venue categories are ranked according to how often they occur in each city, and a ranking of the most frequent restaurants in each city is created. In addition, k-means clusters are used to cluster the cities and their preferences.

Cluster visualization

Since each city is clustered in its restaurant preferences, these are marked on a map and the marking is colored according to the cluster class. This identifies European areas where certain restaurants are more common and therefore preferred.