How to find the best restaurants in Berlin? A Data Science study

Problem Presentation

A data scientist who loves food will have a stop over of 16 hours in Berlin due to his business trip to Dublin.

During this time, he would like to stay close to Brandenburg Gate (a place he wants to visit very much) and also would like to know the best restaurants in the neighborhood to maximize his experience.

To solve this problem, he decided to evaluate the area, helping him to decide which is the best zone to stay and fulfill his objective, considering the restaurant's type

Data Collection and Transformation flow

1 Requirements

- Geolocator defined as Brandenburg Gate.

- Restaurants closest as possible to geolocator object.

2 Foursquare API

- Collect Berlin geographic locations (lat / lon)

- Collect candidates location geometry information
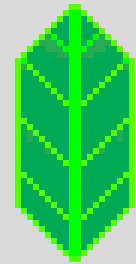
3 Data Analysis

- Crate dataset, removing unnecessary data.

- Define clusters according to desirable requirements.
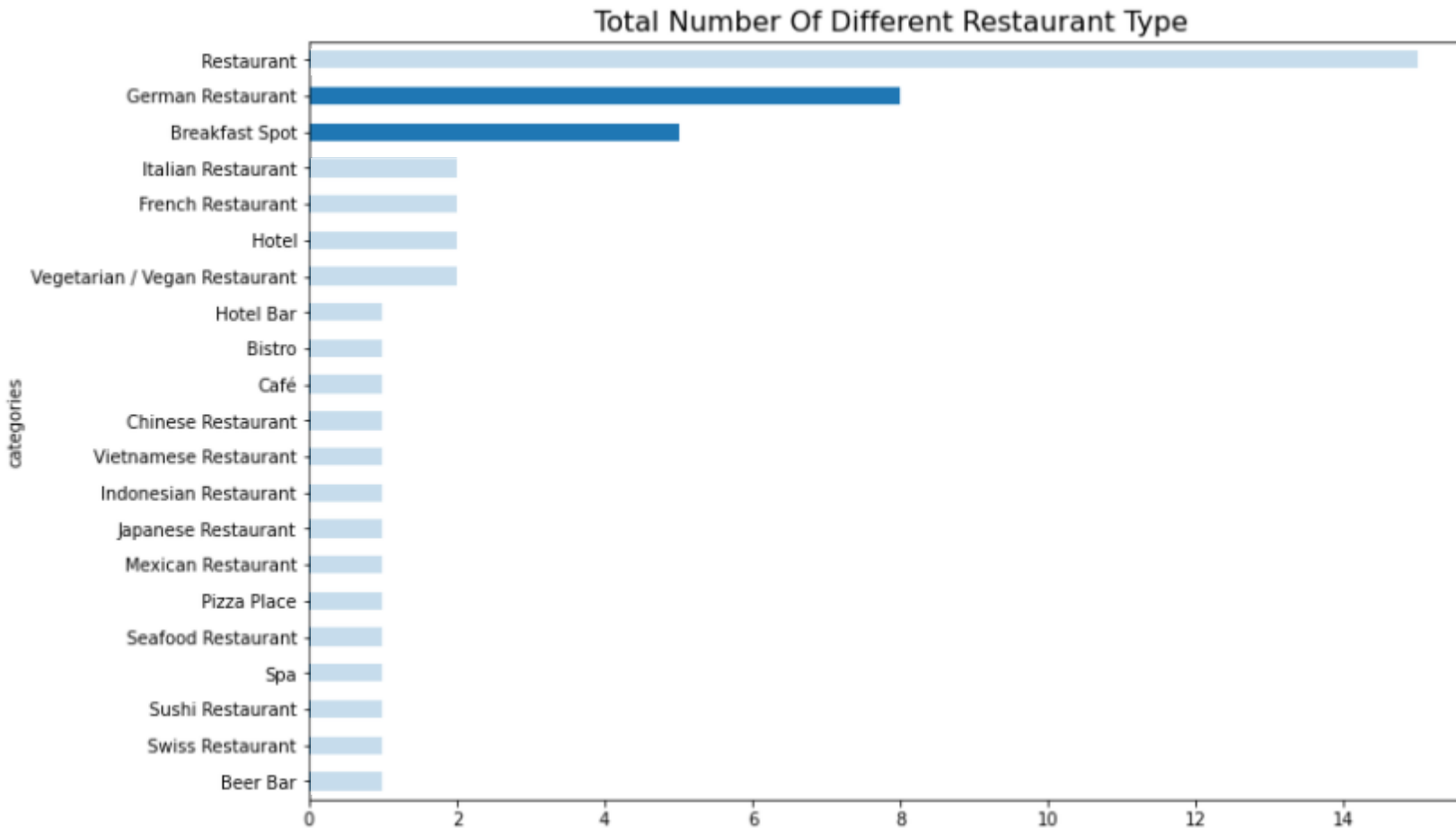
4 Final Results

- Plot final result using Folium library

Requirements and Foursquare API

- Since the man wants to visit Brandenburg Gate, we must consider it as a fixed variable in this evaluation;

- Since the time is limited, it is necessary to set a limit in meters from Brandenburg Gate until the most distant point of analysis. It could be up to 1000 meters from Brandenburg Gate, but it is strongly desirable to be up to 600 meters to reduce the time walking;

- The man would like to try German cuisine, but does not exclude another category if German restaurants are too far;

- It was used a free account of Foursquare API, which allows to capture only 50 rows, but since we wish a limited number of restaurants, this is not a problem.
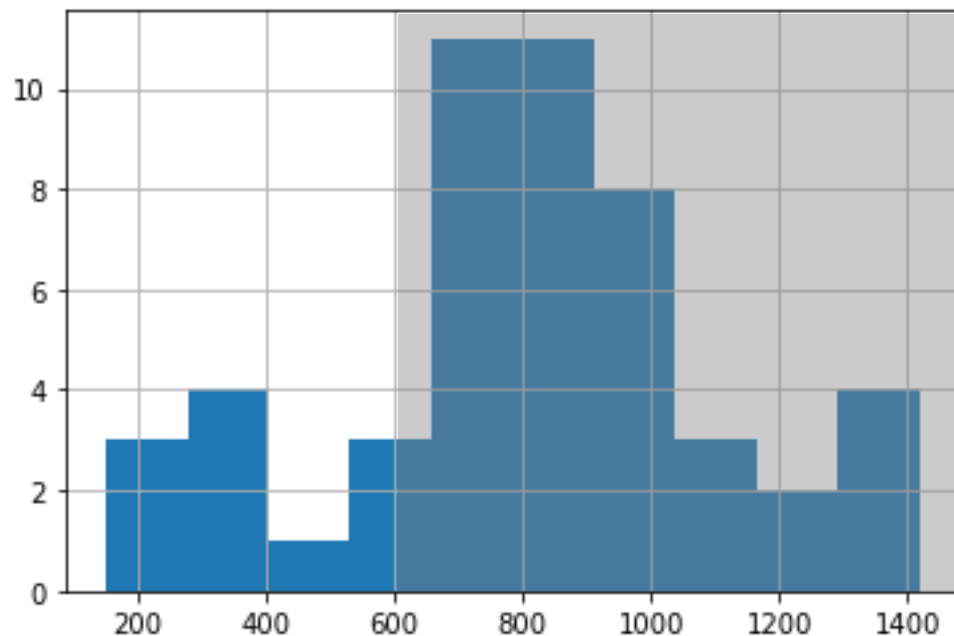
Data Analysis | Group by Categories

Total Number Of Different Restaurant Type

- The biggest group is named 'Restaurant' which does not add any value to the analysis;

- The other categories are pulverized with small number of aggregation (1 or 2). If the visitor would like a different type of restaurant, this could be decisive to the choice;

- Since the visitor wants to analyze the German *cuisine* (or the next one near to) we can focus the analysis in both of German Restaurants and Breakfast Spot.

Data Analysis | Group by Distance

- As the histogram below shows, most of places are between 600 and 1000 meters from Brandenburg Gate. However, for a first analysis, we will try to use a more strict client requirement (up to 600 meters). If does not work, we can try another distance range.

- Considering this new limit in the previous list of German and Breakfast places, we can exclude most of the option, choosing the **Dachgarten-Restaurant**

Restaurant Name	Category	Distance
Dachgarten-Restaurant	German Restaurant	282
Restaurant Coelln	Breakfast Spot	750
NH Collection Breakfast Restaurant	Breakfast Spot	880
Restaurant	Breakfast Spot	1099
Restaurant Nolle	German Restaurant	859
Restaurant im Deutschen Theater	German Restaurant	908
Restaurant Velis	German Restaurant	1165
Restaurant Aquin	German Restaurant	1323
Café Madrid Restaurant	Breakfast Spot	911
Eurostar Berlin breakfast restaurant	Breakfast Spot	948
Gendarmerie	German Restaurant	898
Josty	German Restaurant	752
Mark Brandenburg	German Restaurant	1074

Final Results (powered by Folium)

1) Discarded options are highlighted in blue

2) Chosen restaurant, following original requirements, highlighted in yellow

3) Brandenburg Gate is highlighted in red

4) Since the original requirements are achieved, new iterations were not necessary. However, in a scenario with more time, different analysis could be done considering second and thirdy options.