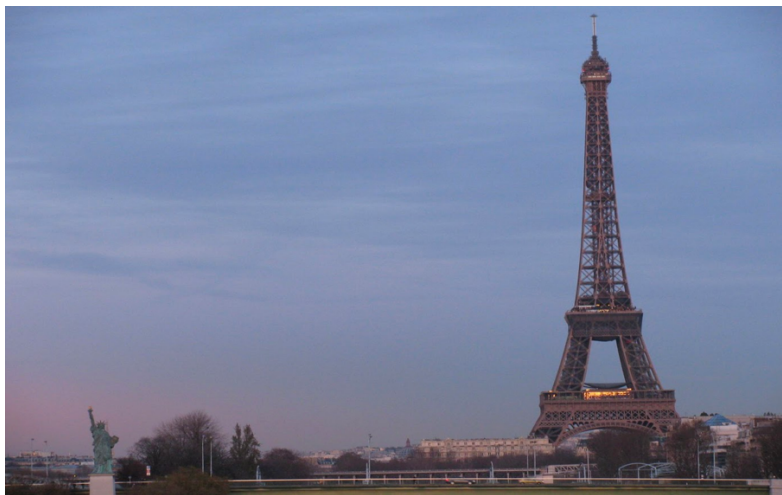Coursera - IBM Applied Data Science Capstone

Indian restaurant in Paris!

Seetharaman Chandramouleeswaran

Introduction

Paris is a beautiful multi-cultural city, and it is also a city of food lovers. The city has a very vibrant immigrant culture, especially from India and other south Asian countries where the cuisine is similar to that of Indian. Also, Indian food is popular worldwide. Hence, opening a restaurant in Paris specializing in Indian cuisine will be a good, profitable business venture.

Business Problem

Paris is a large city with various districts populated by people of different ethnicity and nationality. So, identifying the right location is key to growing the business sustainably. Setting up the restaurant in a district where there is already good number of Indian restaurants would be preferable due to the higher population of the Indian community in that area. The city of Paris is large and is composed of 20 arrondissements, also known as administrative districts. This project aims to explore these 20 districts and identify the suitable one for opening an Indian restaurant.

Target audience of this project

This project aims to assist the growing Indian origin entrepreneurs in the city of Paris who plan to set up an Indian restaurant and is looking for a suitable neighborhood. This will also aid any other enthusiastic entrepreneurship who wishes to start an Indian restaurant in Paris.

Data

To address the business problem stated above, the following data are needed:
- List of administrative districts or arrondissements of Paris, France in Western Europe.
- Geographical coordinates – latitude and longitude – of these districts of Paris, in order to plot the map and get venue data.
- Venue data – that is, data related to Indian restaurants in Paris. This will be used to perform clustering on the districts.

Data sources and extraction methods:

- List of arrondissements of Paris will be scraped from the Wikipedia page:

  https://en.wikipedia.org/wiki/Arrondissements_of_Paris using Python requests and

  BeautifulSoup package.

- Geographical coordinates of the districts of Paris will be obtained from the open data provided

  by Government of France in csv format from this link:

  https://www.data.gouv.fr/en/datasets/arrondissements-1/#_

- Foursquare API will be used to get the venue data for the districts

Skills applied

Following skills learnt in this course will be applied in the Capstone project:
- Web scraping for data extraction

- Pandas DataFrames for tabular data

- Foursquare API for fetching location data of required venues

- Data cleaning and wrangling

- Machine Learning | K-Means Clustering

- Data Visualization for maps using Folium