Finding the best neighbourhood where to open a restaurant in a foreign city

I. Business Problem

a. Introduction

Immigrating to a new country as a foreigner isn't easy, even more when thinking about opening a new business. The quantity of French restaurants in Chicago is, at first glance, relatively low compared to other city. But even if Chicago might look like a good choice, Chicago is relatively unknown to a future business owner. The aim of this study is to find the best suited neighbourhood for the opening of a French restaurant based on several criteria.

b. Objective

The goal of the study will be to help business owners to export their knowledge to other countries. Here an example with a French restaurant in Chicago. The aim of the study will be to find the neighbourhood fitting the following criteria:

- Affordable lease
- Few competitors
- Dynamic (many venues to go out)
- Safe neighbourhood
- Wealthy neighbourhood (middle range)

II. Data acquisition

a. Community Areas of Chicago

The community areas and names of Chicago will be directly scrapped from the following Wikipedia page https://en.wikipedia.org/wiki/Community_areas_in_Chicago. Other useful information like the population and the area will be used later to create density based features.

b. Affordable lease

The Chicago.gov website has been used to know the affordability of every neighbourhood. The data has been exported as CSV and will be first manually cleaned in excel (quicker); then imported into our Jupyter for further cleaning. The index named NOAH presented in the table is particularly interesting for our needs because it shows the 'Naturally Occurring Affordable Housing' in percentage.

Link:
https://www.chicago.gov/content/dam/city/depts/dcd/general/CITY_OF_CHICAGO_AFFORDABLE_HOUSING_DATA_TABLE.xlsx

c. Few competitors

In order to get an idea about the level of competition (or number of similar restaurants), the Foursquare API will be used. The number of French restaurants of every area will be calculated by

summing up the number of French food (or European food) venues in every area in a radius of 1000m.

d. Dynamism (many venues to go out)

In order to get an idea about the dynamism level, the Foursquare API will be used. The level of dynamism of every area will be calculated by summing up the number of food venues in every area in a radius of 1000m.

e. Safe neighbourhood

The cityofchicago.org website has been used to calculate the level of safety of every neighbourhood. Only the annual reports from 2018 has been used in our case because the data was recent and complete. The data will be imported as excel into our Jupyter Notebook for further cleaning. Then the number of crimes will be counted for every area, without differentiation of the gravity of the crime. A new feature called crime ratio (number of crime divided by population) will also be created for every area.

Link:
https://data.cityofchicago.org/Public-Safety/Crimes-2018/

f. Wealthy neighbourhood (middle range)

The cityofchicago.org website has been used to get the per capita income for every neighbourhood. The data may be a bit old (2008-2012) but still good enough to have an idea about the wealth of every neighbourhood.

Link:
https://data.cityofchicago.org/Health-Human-Services/Per-Capita-Income/r6ad-wvtk