

The Battle of the Neighborhoods

Pizza Places in Selangor and Kuala Lumpur

1. Introduction/Business Problem

- 1.1. I love pizza. Well, who doesn't? And yet, there are only a few pizza places available in where I live and work, which is the state of Selangor, Malaysia. There are also just a handful of pizza places in Kuala Lumpur (where I work).
- 1.2. Perhaps, this is due to my ignorance of the available choices. Or maybe, there are really few choices available.
- 1.3. Hence, the problem that this project attempts to solve would be:
 - i. What is the distribution of pizza places in Selangor and Kuala Lumpur (referred to as "my hometown" from now on); and
 - ii. Where would be the best location to open a new pizza place in my hometown.
- 1.4. Target audience for this project are as follows:
 - i. Myself. I travel a lot within my hometown for my work., therefore, the result of this project should improve my pizza meal choices moving forward.
 - ii. Existing and future restaurant owners. The project should somewhat benefit this group. Food & beverage industry is very competitive, particularly for Western cuisine restaurants. A Southeast Asian country, rice-based meals are the preferred option for the locals, hence, location is a critical factor for Western cuisine restaurants.

2. Data

Throughout this project, the following data will be exploited through data cleaning, data wrangling, exploratory data analysis and machine learning techniques:

2.1. List of neighborhoods in my hometown

No known single source, hence, we will be using these sources from Wikipedia:

- https://en.wikipedia.org/wiki/Category:Townships_in_Selangor; and
- https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur

Web scraping techniques are used by using Python requests and BeautifulSoup packages.

2.2. Coordinates of the listed neighborhoods

Python Geocoder package should be able to provide us with the latitude and longitude of the listed neighborhoods.

2.3. List of pizza places in the listed neighborhoods

Foursquare API will be utilized as datasource, which should be able to provide to us with all kind of information in relation to venues, which include but not limited to venue name, category, menu, as well as location. Particularly, "Pizza Places" should be available as part of the information.