

# **Coursera Capstone Project**

## **IBM Applied Data Science Capstone**

### **Opening a New Bakery in Karachi, Pakistan**

By: Roha Farooq

May 2020



# Data

To solve the problem, we will need the following data:

- **List of Towns (neighborhoods) in Karachi:** This defines the scope of this project which is confined to the city of Karachi, the former capital of the country Pakistan.
- **Coordinates (Latitude & Longitude):** The latitude and longitude of these neighborhoods are required to plot them on map and get subsequent venue data.
- **Venue Data:** Venue data particularly related to bakeries. This data will be used to perform clustering on the neighborhoods to identify potential areas for new bakery opening.

## Sources of Data and Extraction Methods

1. Wikipedia page (Wikipedia, n.d.) contains a list of neighborhoods in Karachi, with a total of 18 neighborhoods.
  - a. Data will be extracted from Wikipedia page, with the help of Python requests and BeautifulSoup (Crummy, n.d.) packages.
  - b. Geographical coordinates of the neighborhoods will be obtained using Python OpenCageGeocode (Geocoding API, n.d.) package which will give us the latitude and longitude coordinates of the respective neighborhoods.
2. Foursquare API (Foursquare, n.d.) to get the venue data for all neighborhoods. Foursquare has one of the largest database of 105+ million places and is used by over 125,000 developers. Foursquare API will provide many categories of the venue data, we are particularly interested in the '**Bakery**' category in order to help us solve our business problem.

## References

- Crummy. (n.d.). *BeautifulSoup Documentation*. Retrieved from <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>
- Foursquare. (n.d.). *Foursquare Developers Portal*. Retrieved from <https://foursquare.com/developers/apps>
- Geocoding API. (n.d.). *OpenCage Geocoding API Documentation*. Retrieved from <https://opencagedata.com/api>
- Wikipedia. (n.d.). *Category: Towns in Karachi*. Retrieved from Wikipedia: [https://en.wikipedia.org/wiki/Category:Towns\\_in\\_Karachi](https://en.wikipedia.org/wiki/Category:Towns_in_Karachi)