

# Project Proposal

Fai Abdulrahman Albadri

## Question/need:

- Which restaurant cuisine is needed in a specific area?
- What is the average price for the food according to the area and cuisine?
- What is the most liked restaurant cuisine?

## Data Description:

The dataset is provided in .csv format. It contains 19360 Restaurants; each Restaurant has 11 features. The most relevant feature to this project is the text which contains the restaurant name. Some other features are extracted from other features such as the address it is extracted from location, where it contains (Latitude, Longitude). Other important features are the price, Rating and likes that we can extract how these features have relationship between them to effect on the customers.

### **This dataset contains following features:**

- name: The name of the restaurant. Some names are only available in Arabic.
- categories: The list of categories (in English) separated by a comma.
- address: The address according to foursquare. Unfortunately, it is the least useful column in the dataset because lots of restaurants lack a formal address.
- lat: Latitude.
- lng: Longitude.
- price: shows the price category (Cheap, Moderate, Expensive, Very Expensive)
- likes: The number of likes.
- photos: The number of photos. Photos themselves are not included.
- tips: The number of tips in Foursquare (e.g., Don't miss their pasta). Tips themselves are not included.
- rating: The average rating out of 10.
- ratingSignals: The number of rating.

## Tools:

- Numpy and Pandas for data manipulation
- Scikit-learn for modelling
- Matplotlib and Seaborn for plotting
- Streamlit for interactive visualizations