

# Coursera Capstone Project – Applied Data Science

Vivek Nichenametla

# Introduction – Business Problem

- Hyderabad, located in the Southern India is fourth largest city in India and is rapidly growing. The city is famous for its varied cuisine and with existing variety of restaurant chains, starting a restaurant is difficult task.
- This project is to find a suitable place in the city to start a restaurant, which of course is dependant on several factors. Some of them are
  - Similar restaurants around the place
  - The connectivity to the place
  - The affordability of the potential customers around the place
  - Availability and real estate price of the location, etc.,
- This project deals with the fundamental analysis based on location and clustering the restaurants obtained from Foursquare API. The other factors mentioned above are not in the scope of this project though their significance cannot be neglected.

# Data

## Neighborhoods:

The data of neighborhoods in Hyderabad are extracted by web scraping using BeautifulSoup library in Python. In this case the source of the data is

[https://en.wikipedia.org/wiki/Category:Neighbourhoods\\_in\\_Hyderabad,\\_India](https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Hyderabad,_India)

## Geo-coding:

After retrieving the data of neighborhoods into a pandas Data Frame, the latitudes and longitudes are obtained using geocoder package in Python

## Venue details:

The venue details are obtained using Foursquare API.

<https://developer.foursquare.com/>