Business venture to maximize revenue

Abstract:

In an over-crowded city like Bengaluru in India, coming up with a food venture is like a double-edged sword. On one hand we see a high number of restaurants and eateries which would mean a high competition and on the other hand an increasing population in the IT capital of India would mean an opportunity to venture into a profitable business.

Introduction:

As one seeks to venture into a profitable food business, the success and failure of the venture would depend on the present competition in the area, type of the venue, footfall in the neighborhood. This is primarily targeted to individual / groups who would like to venture into a food business and would like to know details which give them an upper hand before start of any business. Data Science with ML will help uncover / segment a city (like Bengaluru) to answer and identify areas which are similar in nature and the opportunity cost to venture into a specific type of food business is low and the returns are high.

Data Acquisition:

The data used will be obtained by scraping through a publicly available Wikipedia page which lists all the regions in a city of Bengaluru and neighborhoods within each region. Using geopy, co-ordinates of the neighborhoods would be identified. Neighborhoods whose co-ordinates are unidentified will be cleaned from the data. Moreover, there are cases when the co-ordinates differ from the actual co-ordinates due to multiple regions with the same name. Such neighborhoods are identified, and the co-ordinates are updated manually. This clean data will include details like **Neighborhood, Region, Latitude, Longitudes**.

This data will be used for further processing with Fairsquare API to identify the nearby venues in each of the neighborhood. The data will now contain details like **Neighborhood, Region, Latitude, Longitudes, <Venue categories>.** The data will be segmented / clustered, and the resulting data can be examined to identify similar areas and number of venues in each of these areas in a cluster. Top 3 venue categories in the areas within a cluster can be identified and the area which lacks such venue categories can be a right opportunity in terms of location / category to venture into.

A brief view of a clean data would be as follows:

	Neighbourhood	Region	latitude	longitude
0	Cantonment area	Central	13.0196	77.5096
1	Domlur	Central	12.9625	77.6382
2	Indiranagar	Central	12.9733	77.6405
3	Malleswaram	Central	13.0163	77.5587
4	Rajajinagar	Central	12.9882	77.5549

Post obtaining the data with the help of Foursquare API and merging with the initial data we should be able to identify the co-ordinates of the venue and the venue categories

	Neighbourhood	Region	latitude	longitude	venue_name	venue_longitude	venue_latitude	venue_category
0	Cantonment area	Central	13.019567	77.509589	Maruthi Theatre	13.018795	77.505283	Indie Movie Theater
1	Cantonment area	Central	13.019567	77.509589	bhat canteen	13.016796	77.504154	Fast Food Restaurant
2	Domlur	Central	12.962467	77.638196	Lavonne	12.963909	77.638579	Café
3	Domlur	Central	12.962467	77.638196	Barbeque Nation	12.962684	77.641599	BBQ Joint
4	Domlur	Central	12.962467	77.638196	Smoke House Deli	12.965584	77.641498	Deli / Bodega

More than 150 unique categories can be identified with the help of Fairquare API.