## Coursera Capstone – “The Battle of the Neighborhoods”

Comparative analysis using location data via Foursquare.

By: Debanjan Ghosal

*Table of Contents:*

[COURSERA CAPSTONE – “THE BATTLE OF THE NEIGHBORHOODS” 1](#_TOC_250005)

[PROBLEM STATEMENT 1](#_TOC_250004)

[DATA OVERVIEW 1](#_TOC_250003)

Source 1: Dubai\_neighborhoods.csv 1

Source 2: Venue data via Foursquare 2

# Problem Statement

Due to its central geographic location and generally welcoming climate, Dubai is known for its international atmosphere. It is home to close to 2.5 Million expatriates who make up 90% of the total population1 and it hosts close to 15 Million visitors per year.2 In addition to its multi-cultural population and capital flows, Dubai is becoming well known as a destination of choice for great food. It is a place where people can rest and try the best of each culture, either while they work here temporarily, or of they are just passing through.

The mission of this project is to use Foursquare location data and regional clustering of venue information to determine what might be the ‘best’ neighborhood in Dubai to open a restaurant. As a westerner who has a passion for good Mexican food, I have found that there is not a lot of selection in the region. This is supported by the fact that a review of Foursquare venues reveals zero Mexican restaurants listed on the site.

My proposal, then, is an analysis of the neighborhoods in Dubai for the consideration of opening a new Mexican restaurant. The objective is to have a location that is within one of the more reasonable rent zones, but also within a close enough range (5km) to a ‘high’ rent zone. The assumption that proximity to high rent neighborhoods would result in takeout opportunities or provide residents an option to travel to our restaurant.

# Data Overview

The data that we will use for this analysis is a combination of a CSV file that has been prepared for the purposes of the analysis from multiple sources (Dubai\_neighborhoods.csv) and the location/venue information in foursquare.

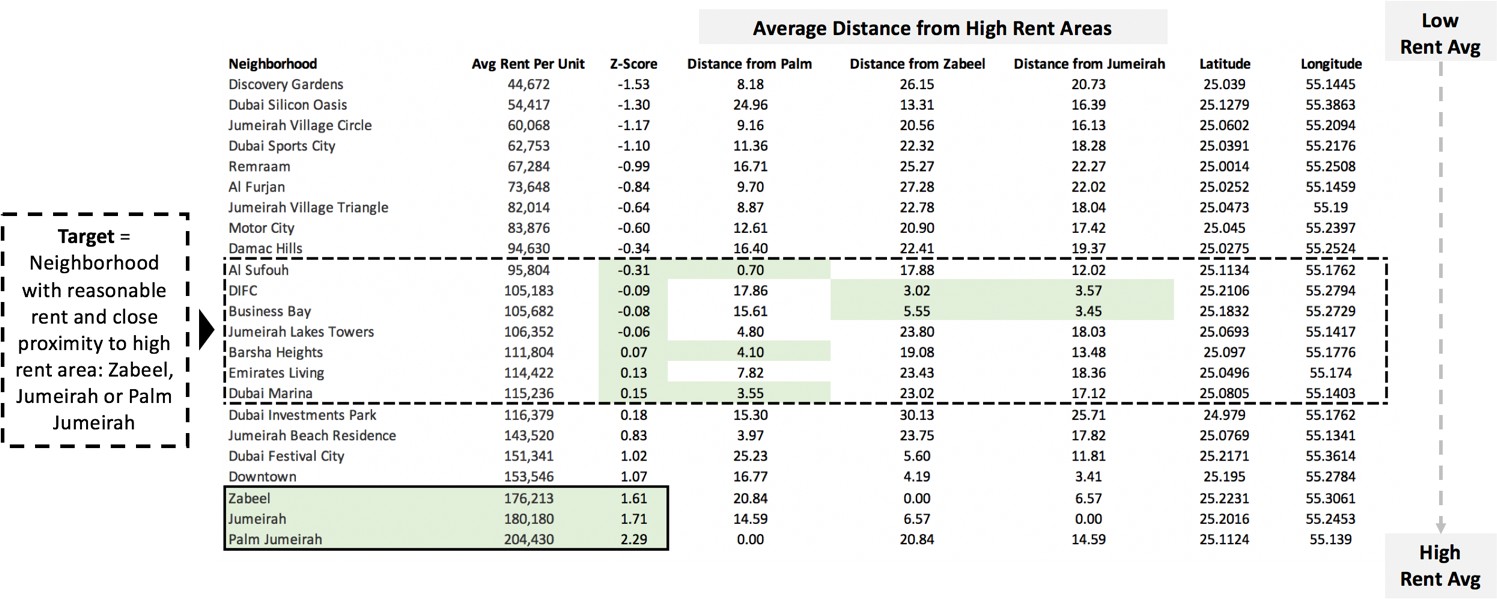
Source 1: Dubai\_neighborhoods.csv

Link: <https://github.com/deb201819/Coursera_Capstone/blob/master/Dubai_neighborhoods.csv>

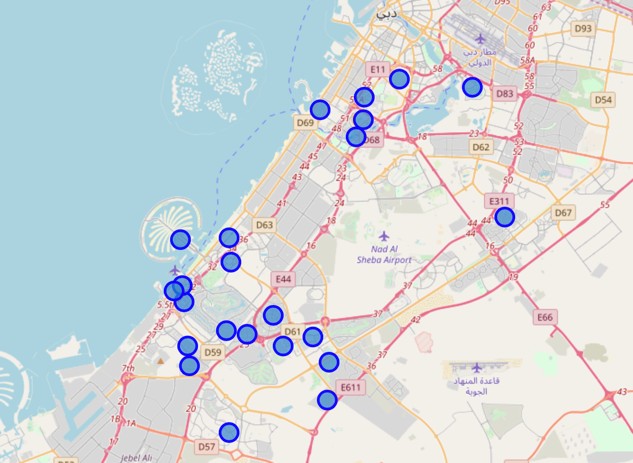
We will first determine the most likely neighborhoods for a restaurant based on average rental prices and relative distance to a high rent center. In the case of Dubai, the three highest rent averages can be found in Palm Jumeirah, Jumeirah, and Zabeel. We will then consider the total number of venues and additional criteria such as proximity to a shopping center or offices for mid- day traffic to make a final determination.

1 https://[www.dsc.gov.ae/en-us/Themes/Pages/Population-and-Vital-Statistics.aspx](http://www.dsc.gov.ae/en-us/Themes/Pages/Population-and-Vital-Statistics.aspx)

2 https://[www.statista.com/statistics/284636/visitor-arrivals-in-dubai-from-international-destinations/](http://www.statista.com/statistics/284636/visitor-arrivals-in-dubai-from-international-destinations/)

The first step is to establish the neighborhoods in Dubai and a summary of their average rental prices. The average rental index is published annually and can be referenced by a number of different websites.3 We then make a calculation of the ‘Z-score’ to standardize the data and sort from lowest to highest average rent. Using google, I looked up the latitude and longitude of each neighborhood and entered it. This information is used to calculate the distance of each neighborhood from the highest rent regions and will also be entered to the Foursquare database for venue query later.

*The csv is stored can be found at: https://github.com/talentrics/coursera\_capstone/*

Source 2: Venue data via Foursquare:

Using clustering techniques developed in prior exercises in the capstone project, we will examine the most common venues by neighborhood listed in Foursquare.

When we cluster the data together and rank by ‘most common’, it appears that Mexican venues not registered in any of the neighborhoods, so it is safe to assume there would be minimum competition.

*Map of Dubai neighborhoods using Folium*



The remaining steps in the capstone project will be to evaluate which neighborhood is most suited using the criteria shown above: the rent index of the neighborhood is close to the mean

3 https://[www.propertymonitor.ae/research/uae-communities-index/dubai-rental-index.html](http://www.propertymonitor.ae/research/uae-communities-index/dubai-rental-index.html)

average rent in Dubai, the neighborhood’s is located close to a ‘high rent’ neighborhood, the general restaurant frequency in the neighborhood is reasonable, and proximity to ‘other’ venues such as business centers or malls is maximized.