

## Contents

Restaurant location recommender.....	1
Introduction .....	1
Business Problem .....	1
Data .....	2
Methodology .....	2
Results .....	3
Conclusion:.....	3

# Coursera Data Science Capstone Project

## Restaurant location recommender

### Introduction

Abu Dhabi is considered one of the fastest growing economies worldwide; it is famous for having a business-welcoming environment, diverse population, many famous tourism attractions and delicious multi-cultural restaurants, but the problem might be sometimes is that restaurants are cluttered in specific places, and it would be nice to have at least local restaurants distributed evenly .

### Business Problem

The question that I will answer throughout this capstone project is; where is the best location to open a restaurant in Abu Dhabi, and to answer that, I will be using Foursquare location data combined with neighborhoods data which I have collected from google maps, along with k-means clustering to determine the best location for a restaurant by analyzing venues from Foursquare and finding out which neighborhood has the least concentration of restaurants.

**This recommender system will be very beneficial for anyone seeking the best location in Abu Dhabi to open a restaurant.**

## Data

The data used in this project will be:

- Neighborhoods in Abu Dhabi with their coordinates, this is collected from Wikipedia and google maps.
- Venues for each neighborhood from Foursquare.

Once venues for each neighborhood is collected, then we will measure the frequency of restaurants in each neighborhood and find the neighborhood with the least restaurant frequency.

## Methodology

In this project, I will be following CRISP-DM methodology (minus the deployment step), which consists of the following six steps:

- 1- Business understanding: Having represented the business case clearly and the question to be answered; deciding the best location in Abu Dhabi to open a restaurant.
- 2- Data understanding: since I will use foursquare maps in this project, and knowing that Foursquare APIs can return categorized venues in a neighborhood given the neighborhood's coordinates, I knew that I will need all neighborhoods in Abu Dhabi along with their coordinates.
- 3- Data Preparation: I collected all Abu Dhabi neighborhoods from Wikipedia, then cleaned the data and made sure that all neighborhoods are correct and actually exist in Abu Dhabi (I found that there are 4 neighborhoods are not actually neighborhoods so I removed them), and then collected the coordinates for each neighborhood through google maps, and combined the said in a csv file.
- 4- Modeling: having collected the needed data, what's remaining is only to feed all neighborhood coordinates to Foursquare API and get the list of venues in each neighborhood, and then find the neighborhood with the least restaurant frequency.

To implement this model, I have used pandas, numpy, request, matplotlib, and Foursquare libraries, following steps shows how:

- Read neighborhoods with their coordinates, then show them on the map
  - Send each coordinate to Foursquare API and get then list of venues
  - Transform venues into frequencies based on the venue category
  - Show the restaurants in each neighborhood on the map
  - Choose the neighborhood with the least frequency.
- 5- Evaluation: testing the model and make sure the results are correct (i.e. the model returns the neighborhood with the least restaurant frequency).

## Results

Having explored the data, I found that venue categorization could be better in Foursquare, my belief is that Foursquare data in UAE might not be consistently updated, and the labeling of venues could be further enhanced.

## Conclusion:

After exploring venues from 38 neighborhoods in Abu Dhabi, I found that there are 5 neighborhoods where restaurants are not in the top 10 venues in terms of frequency, and so those 5 neighborhoods will be a suitable target for a restaurant business.

	Neighborhood	
2	Al Bateen West	
5	Al Manaseer	
6	Al Manhal	
15	Al Reef	
29	Qasr El Shatie	

