# Introduction

My friend owns a restaurant in Abu Dhabi inspired from the famous Game Of Thrones show. The restaurant has received great coverage from the media as per the article below:

<https://www.thenational.ae/arts-culture/hidden-food-gems-fans-will-love-this-game-of-thrones-inspired-sandwich-shop-in-khalidiya-1.47219>

The interior design resembles that of a typical restaurant/bar in the North of Westeros. It consists of around 6 counter stools for visitors to come and enjoy a fast meal. The restaurant is also registered with delivery companies. The latter accounts for more than 80% of their business. The main reason behind having low physical customers is the location; as the restaurant is in a rundown section of the city. This location was initially selected because he was able to get a great deal on rent. Now my friend has accrued some cash from operations and is interested in expansion. His initial idea was to relocate within Abu Dhabi to a more prime location, but I convinced him not to change the setting as it became part of his restaurant’s brand in Abu Dhabi and hence relocating might not yield the highest ROI. Instead, I convinced him to expand operations in Dubai which has a higher population, much higher influx of tourists and is generally a lot more western than Abu Dhabi and hence would appeal more to the local population.

This time he wants to select an optimal location from the beginning and has requested my help. My recommendation to him was to go for a community that satisfies the following criteria listed below:

* Large population density
* Demographics:
  + Young, single or small family
  + Predominantly European, North American and Western mindset
* Large influx of business visitors and leisure tourists
* Surrounded by other similar communities to facilitate delivery / pick up
* Large number of venues to indicate strong business sentiment

We decided to resort to Data science and Machine Learning to help us segment the different communities in Dubai which’ll greatly narrow down our choices.

# Data

To get census data about the different communities in Dubai, we used the following Wikipedia link:

<https://en.wikipedia.org/wiki/List_of_communities_in_Dubai>

The main table in this link consists of the following structure:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Community # | Community Name in English | Community Name in Arabic | Area | Population | Population Density |
|  |  |  |  |  |  |

There are 130 diverse communities in Dubai. All the data types are of type object or string and hence further processing is needed especially on the population density column.

We plan on using K-means unsupervised Machine Learning method to segment our communities based on the prevalent venue types. Hence, we’ll need to ultimately have a data frame consisting of all communities along with their corresponding top venues. To get the venue information we will be using the Foursquare API. This requires us to get the latitude/longitude data per community; we will use the geopy library to get that information.