IBM DATA SCIENCE COURSE 9.

# Applied Data Science Capstone:

# The Battle of Neighbourhoods.

DATA SECTION

# 1. FILES AND SOFTWARES USED.

**FILES AND LANGUAGES:**

We will use an excel file named “Dubai\_Neighbourhoods.xlsx” as our main dataset.

We will convert this file into a pandas dataframe.

We will use the location and venue information from Foursquare.

The code is written in Python.

**SOFTWARE REQUIREMENTS:**

Jupyter Notebook.

Adobe Reader.

Microsoft Excel.

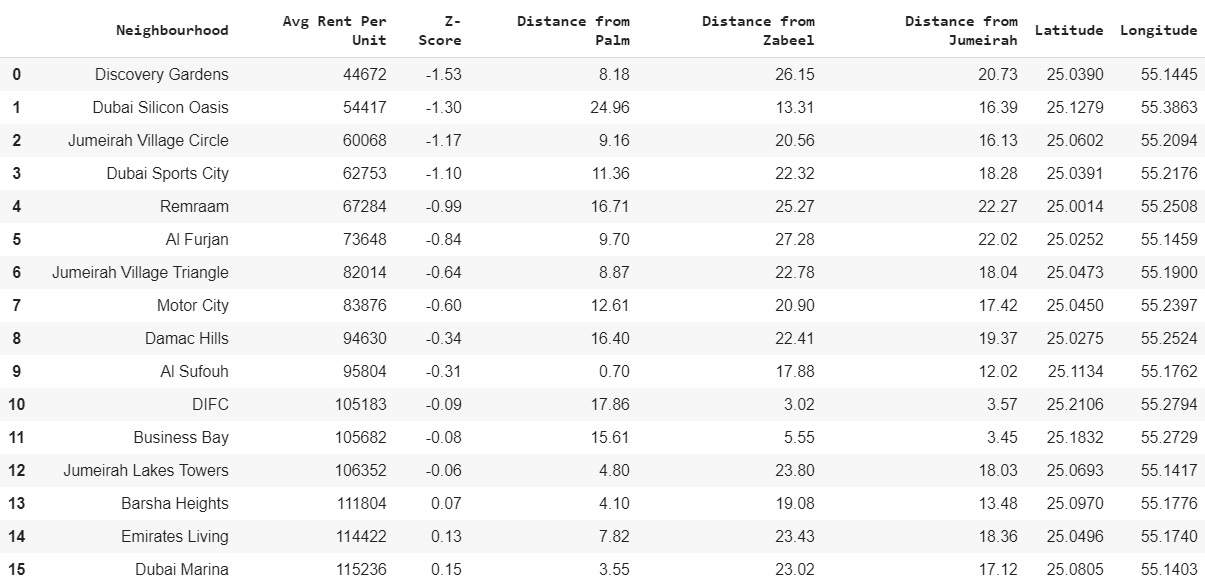
Foursquare, etc.

# 2. DATA DESCRIPTION AND VISUALIZATION.

**DATA DESCRIPTION**

The dataset used for analysis is an Excel file named “Dubai\_Neighbourhoods.” Based on average rent prices and distances from places of high rent, we will determine which neighbourhoods is most suitable for the Mexican restaurant.

The files consists of 8 attributes or columns namely, Neighbourhood, Average rent per unit, Z-Score, Distance from Palm, Distance from Zabeel, Distance from Jumeriah, Latitude and Longitude.



For examination of neighbourhoods, we will make use of Foursquare. With the help of Foursquare, we will cluster most common venues in Dubai.



