# **Capstone Project Week 4 IBM Data Science module**

## **Requirement**

For this week, you will required to submit the following:

1. A description of the problem and a discussion of the background. (15 marks)
2. A description of the data and how it will be used to solve the problem. (15 marks)

## **Problem Description**

We have been asked to write a python program that will explore the City of Birmingham in the UK for a customer who is interested in purchasing a property for his family. The customer would like to know where in Birmingham he should start looking for properties based on the following requirements:

1. The customer has a budget of between 400-500K in GBP
2. He would like the property to be quite central in the city of Birmingham as him and his family enjoy city life.
3. He would like to know what type of venues exist within 10-15 minutes of the property
4. He would like to know what type of facilities exist within 10-15 of the property, such as supermarkets, other shops and hospitals.
5. The customer is also very worried about the safety of his family so he would like to know in those areas that we are going to recommend to him, the crime rates as officially reported by West Midlands Police.

## **Data Description and how they will be used to solve the problem**

To address all the customer requirements we shall develop a python program that will explore some datasets from official sources and will come up with some recommendations for our customer.

In the UK there is an official land registry database that contains information about England and Wales with prices per post code and street name.

The official website is:

<https://www.gov.uk/government/organisations/land-registry>

From there we shall download and process our first csv file that contains all the prices, street names and postcodes for England. We shall process this file as pandas dataframe and we shall extract information for the city of Birmingham, per street name and post code.

From the website <https://www.freemaptools.com/download-uk-postcode-lat-lng.htm>

We shall download the UK csv file that contains all the post codes and maps them to their corresponding latitude and longitudes. We shall then process the file as a pandas framework and extract all the Birmingham post codes that we have identified as per above that we within our customer’s price range. We shall merge the two data sets and map them nicely on a folium map so our customer can have a good visual representation of the information we give him.

We shall then connect to foursquare and extract information for these post codes about venues and facilities in those identified postcodes and present them to our customer so can make an informed decision.

Finally we shall download the csv from the west midlands police website from <http://www.ukcrimestats.com/Police_Force/West_Midlands_Police#>

that contains all the information about the crimes that have occurred in those areas for the past year so our customer can evaluate how safe a particular selection may be. We shall make the crimes onto another folium map and present it to our customer for a full report.