# Moving from a smaller town to a large city in Scotland

### Exploring similarities and differences between hometown and cities and towns of interest

#### Applied Data Science Capstone by IBM

#### Part of IBM Data Science Professional Certificate

###### By Emil Fuerstenberg Haegg

## Introduction

### Background

Scotland is a country that is part of the United Kingdom and makes up the northern part of the island known as Great Britain. Scotland has a varied nature with the southern and eastern parts mostly consisting of rural lowlands, while the north-western parts are aptly known as the Highlands. The Highlands have a varied geology with mountains, countless islands, forests and bogs, but is sparsely populated. Most of the country's population is instead concentrated to the two larger cities in the south: Glasgow and Edinburgh, and the surrounding areas.

In this project we aim to find a suitable location to relocate from a city to another area, considering some of the factors which would influence the choice. A relative planning to relocate from the town of Stornoway located on the island of Lewis and Harris at the north-western edge of the country. For work reasons they are moving to the area of Glasgow and Edinburgh.

### Business Problem

The problem consists of narrowing down the choice of data for comparing the areas, depending on what the stakeholders, in this case our relative, consider the most important. The number of different comparisons which could be made are vast, and input from our stakeholder will guide which direction to take. Our stakeholder is looking to move to town which is similar to their hometown. As a starting point, we look to compare the cities and town on the following features: venues, population size and GDP per capita. We aim present this comparison visually. This way our stakeholder can get an initial overview on similarities and differences, as a starting point for further exploration.

## Data

Data on the 51 largest cities and towns of Scotland and their population was found on Wikipedia:

https://en.wikipedia.org/wiki/List\_of\_towns\_and\_cities\_in\_Scotland\_by\_population

Data on the hometown Stornoway can also be found at a Wikipedia page, this data can be added manually, which can in this case of a single entry be quicker than other methods:

https://en.wikipedia.org/wiki/Stornoway

GDP per capita for the different areas of Scotland can also be found on a Wikipedia page:

https://en.wikipedia.org/wiki/Economy\_of\_Scotland

Data could also be gathered from official registers but using Wikipedia as a source enables us to make use of this vast collection of knowledge, which is Wikipedia, while demonstrating the usage of commonly used tools for handling unstructured data. Using the package BeautifulSoup we can access the html-code directly and scrape this for the data of interest.

Foursquare Places API can be used to find exploring venues for a given city or town. By connecting to this API we can request up to date information on the locations and categories of venues, within a radius from a certain coordinate. We look to arrange the towns and cities into clusters depending on the most common venues in the neighbourhood.

To create a choropleth map, we need a geo JSON-file representing the borders of the different areas, which can then be used to layer on top of the map and coloured according to GDP per capita. A JSON-file representing Local Authority Districts can be found at the following GitHub repository. The file has not been updated for several years, but represents the current areas, in place since 1994.

https://github.com/martinjc/UK-GeoJSON/blob/master/json/administrative/sco/lad.json

## Methodology

## Results

## Discussion

## Conclusion