Start-up recommendation: Province of Rome, Italy

(Capstone project for IBM Data Science Professional Certification)

# Introduction

My first preference would have been to solve a problem relating to my home country (South Africa) but data on the platforms that we were introduced to in the course, is limited for this region. My second preference is Italy given my previous life in the region of Lazio, Rome and surroundings. Looking at the data available, it appears as if there is enough to make this a viable region outside of the US, to target.

The city of Rome has a population of 2856133 but the province of Rome has 1 486079 distributed in about 121 towns around Rome (radius of 50km).

There are many towns around Rome with big populations large that could make a business viable. Office space and taxes in Rome itself is high and capital needed for start-up would be prohibitive for a small business.

## Business question and approach:

The business question that I will try and answer is:

*What would be the best business to start in the province of Rome (outside the capital) and which towns would be most appropriate?*

The approach will be:

1. Find data for all the towns in the province of Rome
2. Analyse all the towns for the following aspects:
   1. Population
   2. Venues already recorded
3. find all the data on Foursquare for those towns
4. Use K-means to cluster the business venues
5. Select the towns with the highest population and propose as business the types that are the least common in those towns.