

Coursera Capstone Project

IBM Data Science



Alin Pirvan
April 2020

Introduction

- Coffee is the most popular drink worldwide with around two billion cups consumed every day.
- In the UK, we now drink approximately 95 million cups of coffee per day. The coffee industry creates over 210,000 UK jobs. The Gross Value-Added contribution from the UK coffee industry to the economy is estimated to be £9.1 billion.
- In this project, we will attempt to use FourSquare and K-Means clustering to find the optimal location for opening a new cafe.

Areas of Glasgow

- Glasgow has a population density of 3,400 people per square kilometer, which makes it the most densely populated city in Scotland. The larger Greater Glasgow area has an estimated population of 1.2 million.
- This represents about 42% of the population of Scotland.

| | District | Population | Area (km ²) | Density (/km ²) | District_Coord | Latitude | Longitude |
|---|---------------|------------|-------------------------|-----------------------------|--------------------------|-----------|-----------|
| 0 | Govanhill | 9,725 | 0.86 | 11,308 | (55.8363741, -4.2581531) | 55.836374 | -4.258153 |
| 1 | Pollokshields | 9,738 | 1.59 | 6,125 | (55.8422663, -4.2849973) | 55.842266 | -4.284997 |
| 2 | Partick | 8,884 | 0.85 | 10,452 | (55.8699211, -4.3094365) | 55.869921 | -4.309437 |
| 3 | Hillhead | 6,275 | 0.96 | 6,536 | (55.8752091, -4.293281) | 55.875209 | -4.293281 |
| 4 | Govan | 5,860 | 1.63 | 3,595 | (55.860879, -4.3185273) | 55.860879 | -4.318527 |
| 5 | Gorbals | 6,030 | 0.83 | 7,265 | (55.851813, -4.2531625) | 55.851813 | -4.253163 |
| 6 | Shawlands | 7,015 | 0.52 | 13,490 | (55.8292301, -4.2924584) | 55.829230 | -4.292458 |

Data Analysis

- FourSquare API will be used to acquire information on the most popular venues of Glasgow, the most frequented venues and use all this in combination with K-Means clustering of neighborhoods to determine the optimal location for a coffee shop.