Applied Data Science Capstone

A New Bike Store Location in London, England

Alan Smale

July 2019

Business Problem

In congested cities, bicycles are an efficient, environmentally friendly method of transport. There is a growing demand for the infrastructure to support people who choose to take up cycling as a means of transport. Bike shops are required as part of this infrastructure, providing bicycles, maintenance services, spare parts and safety equipment. Investors have recognised this opportunity and desire to establish a new bike shop in London.

This project aims to seek desired locations within London with which to locate a new bike shop. The search will be concentrated in Central London districts.

The business question to be answered is:

"Where would you recommend in Central London to open a new bike shop?"

Data Sets and APIs

The following data sets and APIs will be used:

Wikipedia

Provides information on Central London Postal Districts. The tables in the following URLs will be scraped to extract Central London Postal District information:

https://en.wikipedia.org/wiki/WC_postcode_area
https://en.wikipedia.org/wiki/EC_postcode_area

ArcGIS

Provides geocoding through the GeoPy Geocoders Python library. Latitude and longitude coordinates for Central London and for each of the Central London Districts will be obtained.

Four Square API

Provides the ability find data about venues of a given type for a specified latitude and longitude. Bike Venues for each of the Central London Districts will be geographically located.

Approach

Obtain Central London Postal District information from Wikipedia. Combine this information with latitude and longitude information from ArcGIS for each district. Finally use bike shop venue locations for each district from Four Square to build a model of existing bike shop locations. We will then cluster the data to locate districts with a low density of existing bike shops.