Sri

Abstract

Using Clustering technique to identify best place to open a Hotel

Capstone Project

Explore Bristol city, segment and cluster its post codes

**Business Problem**

The objective of this project is to explore Bristol city, segment and cluster its various post code areas and identify a suitable place to open a Hotel Business.

**Background**

Bristol is one of the UK's most popular tourist destinations and was selected in 2009 as one of the world's top ten cities by international travel publishers Dorling Kindersley in their Eyewitness series of travel guides. The city is steeped in history and the historical and heritage sites, like Bristol Cathedral and the Lord Mayor's Chapel, are very popular among tourists. World famous Festivals like Bristol balloon festival, Bristol Harbour Festival draw lot of tourists all year. Bristol is represented by professional teams in all the major national sports so there are regular sporting events held here as well. It is well connected by Air, Rail and Road making it ideal city to open a hotel business.

According to the British Hospitality Association (BHA), the hospitality industry now supports more than 2.7 million people. The hotel industry represents a hugely important part of it, with more than 45,000 establishments across the UK responsible for jobs in hotels and related services. Total combined turnover for the hotel industry is estimated to exceed £40bn – a significant portion of the £127bn tourist economy.

By 2025 this figure is forecast by accountancy firm Deloitte to rise to £257bn, which is around 10% of the UK's GDP. It will support 3.8 million jobs at that point, which is around 11% of the UK workforce.

Actual tourism spending in 2013 reached £113bn, with £24bn via international visitors and £89bn from domestic residents. And for every £1,000 generated directly from tourists into the industry, a further £1,800 goes into the economy via the supply chain and consumer spending, Deloitte's report Tourism: jobs and growth states.

This suggests that despite threats of terrorism, flooding, and bad summer weather in recent years, which have undoubtedly hampered the UK tourism industry, there has been strong growth, particularly following some key international events.

**Target audience**

Based on the background, it looks like opening a hotel in this city is a good business proposition. If a hotel chain wants to open its new hotel or a startup is trying to open their first hotel, this project will provide recommendation on best place to open one.

**Data**

To identify various post codes and cluster them we will be using the following data:

* List of Post codes and their coverage - This will define the area/city in scope for this project which is Bristol, UK. [Wikipedia](https://en.wikipedia.org/wiki/BS_postcode_area) will be used to source this data. This page has a table with Post code districts, Post town, their coverage and the Local authority area. Python’s Beautifulsoup package will be used to scrape data from this page.
* Latitude and Longitude co-ordinates of these post codes which will be used to plot the map and also to get the popular venues list. This can either be found using Python’s Geocoder package or pulled from a compiled list readily available on various sites. 2-Outcode file from the [link](https://www.theinformationlab.co.uk/2011/03/14/hopefully-the-last-word-on-uk-postcode-mapping-in-tableau/) was used to get the co-ordinate details for Bristol.
* Venue data will be used to perform clustering of the post code areas - Foursquare is a location technology platform which has one of the largest data of over 105 million+ locations and is used by over 150k developers. This API will be used to get details of various venues in and around the post codes in Bristol. Once we have the venues, we will understand the categories and then use clustering using K-means to cluster the various post codes based on similarity to gain insights.