Opening a Pub in Bath (UK)

**Introduction**

I will be looking to open a pub in the City of Bath (UK). I appreciate that there is already numerous pubs in this city however I will be looking to open a pub in an area of the city where there is sufficient foot traffic from pedestrians as well as other popular venues that will maximise the number of people who will be in the vicinity of the bar.

Further to that, choosing a location that is close to parking and in close proximity to public transport is key so that both local people and customers from the outside area are able to reach it. Finally, I will consider how much competition there is in the location.

**Data**

I will use the FourSquare data on Bath to find bars, restaurants and other venues to see where the competition is location as well as parking, transport and other busy areas of the city. I will also import postcode data from an excel spreadsheet found from <https://data.bathhacked.org/datasets/bath-north-east-somerset-postcodes> which contains location data on bath postcodes. I will import this spreadsheet via Excel and group the data into their neighbourhoods.

Foursquare API will be used to request information on the top 100 venues and top 10 for each of our assigned Areas. The resulting information will be converted to a pandas dataframe. One hot encoding and k-means analysis will allow me to cluster the venues locations. The venue categories will then be analysed and sub grouped created to encompass multiple venue types. These included:

* Restaurants
* Pubs / bars
* Coffee / cafe's shops
* Shops
* Arts
* Recreation
* Other

**Methodology**

Foursquare developer’s API will be utilised to collect information on the top 100 venues within Bath. This data will then be cross-reference with data on Bath neighbourhoods and the resulting information will be converted to a workable dataframe and locations of each venue visualised on an interactive folium map. The resulting dataframe will be analysed using the one hot encoded and k-means will be applied to the dataset and the resulting clusters will be plotted onto a folium map for visual analysis of where the clusters lie.

**Results**

First examining the top 100 venues in Bath:

A picture containing text, map

Description automatically generated

As you can see from this map, the most densely populated area of popular venues is around Westgate Street up towards John Street. We can also see that just north of the river towards Bath Spa train station and the Bath Bus station there isn't a large amount of competition for our potential sports bar however there is certainly the foot traffic from the transport links and parking available (Avon Street and Bath Spa Station Car Park).

A picture containing text, map

Description automatically generatedNext examining the neighbourhood data and venues within clusters:

Cluster one (red):

A screenshot of a cell phone

Description automatically generated

Cluster two (purple): A screenshot of a cell phone

Description automatically generated

Cluster three (blue):

A screenshot of a cell phone

Description automatically generated

Cluster four (green):

A screenshot of a cell phone

Description automatically generated

Cluster five (orange):

A screenshot of a cell phone

Description automatically generated

After looking at the data from the clusters we can see that clusters one, two, three and five all have either a pub or gastropub within their top 3 most common venues. This gives our first indication that cluster four may be the best area for our pub as it will have the least competition and thus there may be the most need for such an establishment.

**Conclusions**

This gives the indication that cluster four will be the best area for our pub as it will have the least competition and thus there may be the most need for such an establishment.