1.0 Introduction

1.1 Problem

A new investor is looking at the booming Scarborough beach area in Perth and thinking that there is an opportunity to ride the wave and make some real money. In particular they have found a very cheap cafe that is for sale with a completed fit out. The seller insists that there is huge potential and they have simply decided to move to another area.

The investor who is not familiar with the area has asked for advice on the area , in particular specific regions or popular trendy locations in the area which are more popular and ultimately is this in an area that is trendy and well liked. In short: is this going to be a good investment?

1.2 Background

Scarborough is a coastal suburb of Perth, Western Australia, located approximately 14 km northwest of the city centre in the City of Stirling local government area. It was named after the English beach resort Scarborough, North Yorkshire.

It has a population of about 14,300 people (2011 census), of whom about 25% were born overseas, mostly in the United Kingdom. It has a landmark high-rise hotel, the Rendezvous Hotel Perth Scarborough,[2] originally built as Observation City in 1986 for Alan Bond, in anticipation of a demand for accommodation when the 1987 America's Cup challenge was held at nearby Fremantle. Scarborough Beach was the venue for the Australian Surf Life Saving Championships for the years 2007, 2008,[3] 2009 and 2014.

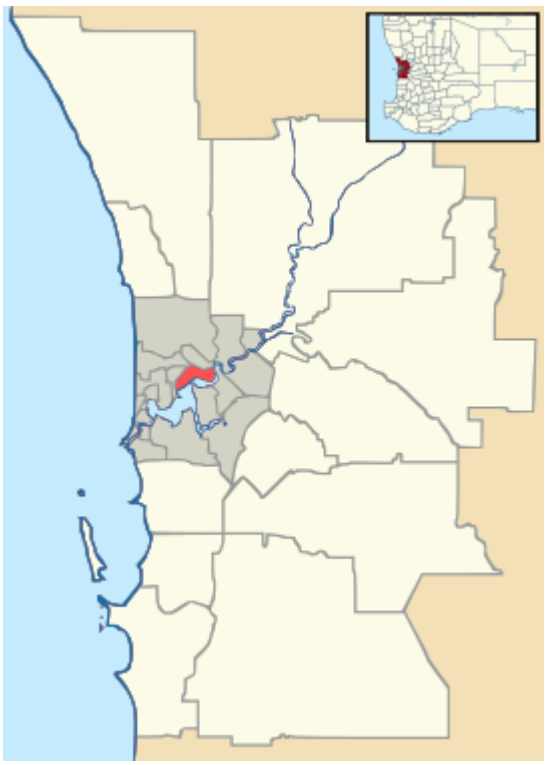

Figure 1 : Scarborough beach
(https://en.wikipedia.org/wiki/Scarborough,_Western_Australia)

2.0 Data

2.1 Data sources

Data will be sourced from Foursquare API and information provided by the investor. The foursquare data via the API is available to the public by creating a developer accounts

2.2 Data Cleaning

Data cleaning will ensure that no NaN entries are in the data list as well as removing outliers that my skew the data. There are also limitations on the radius and amount of data to ensure the focus is on the area surrounding the potential investment.

2.3 Feature selection

The features that are important for this investigation are listed in the table below

Source	Data
Foursquare	Venue ID
Foursquare	Venue Latitude
Foursquare	Venue Longitude
Foursquare	Venue Likes (Count)
Foursquare	Venue Name
Investor	Potential Investment Location (Lat,Long)

3.0 Exploratory Data Analysis / Method

3.1 New Location Pros and Cons

The current proposed location had to be investigated and this was achieved by using the data extracted from Foursquare and mapping it using the Folium Module. This gave a layout of venues around the proposed area

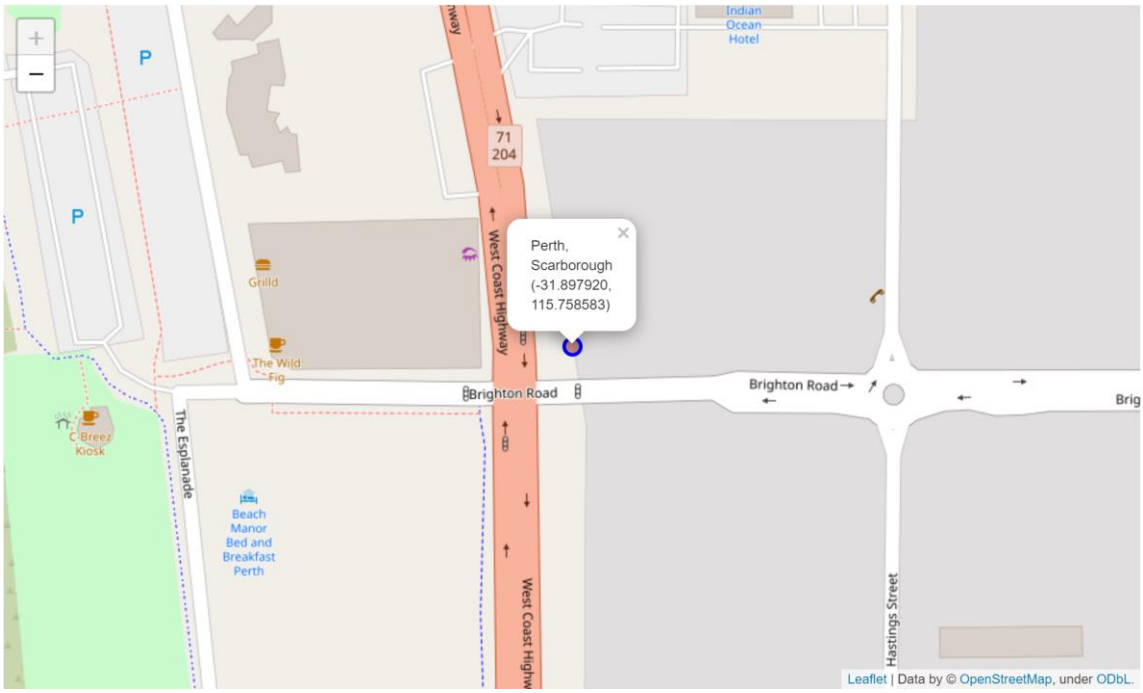

Figure 2 : Proposed Area

3.2 Features

The surround venues were also shown to get an understand of concentration and proximity to the proposed location. This was used by mapping foursquare data and was restricted to 500 m radius and 200 results returned

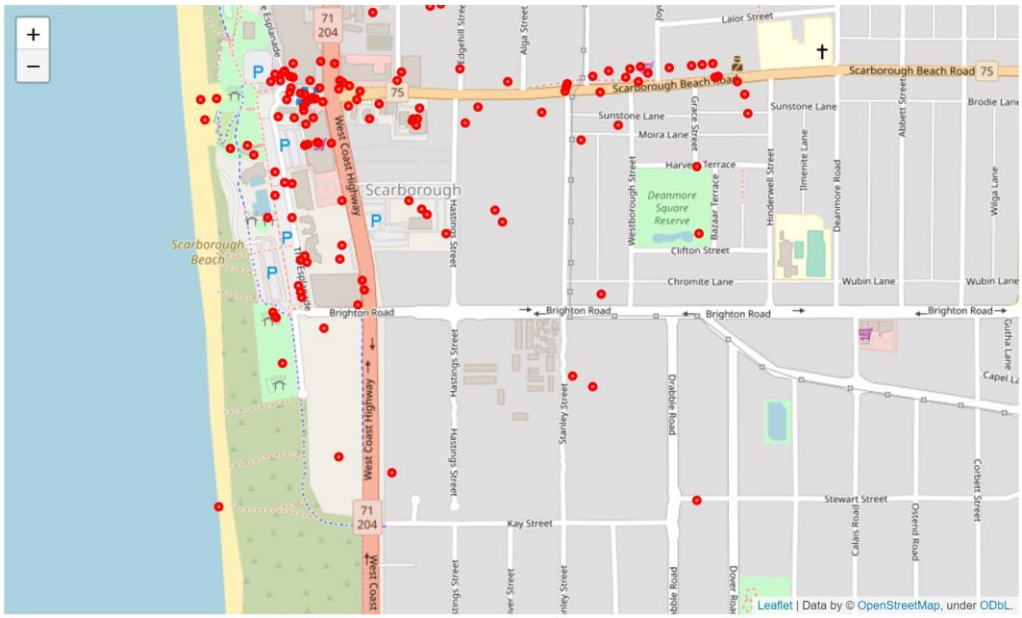

Figure 3: Venues in the surrounding area

3.3 Other features

Understanding popular areas was important and this was achieved by linking data about numbers of likes with venue data. This was extracted from Foursquare and mapped by Folium. Scarborough also had a generic venue which resulted in an outlier was removed to ensure high data quality and accurate results for specific venues

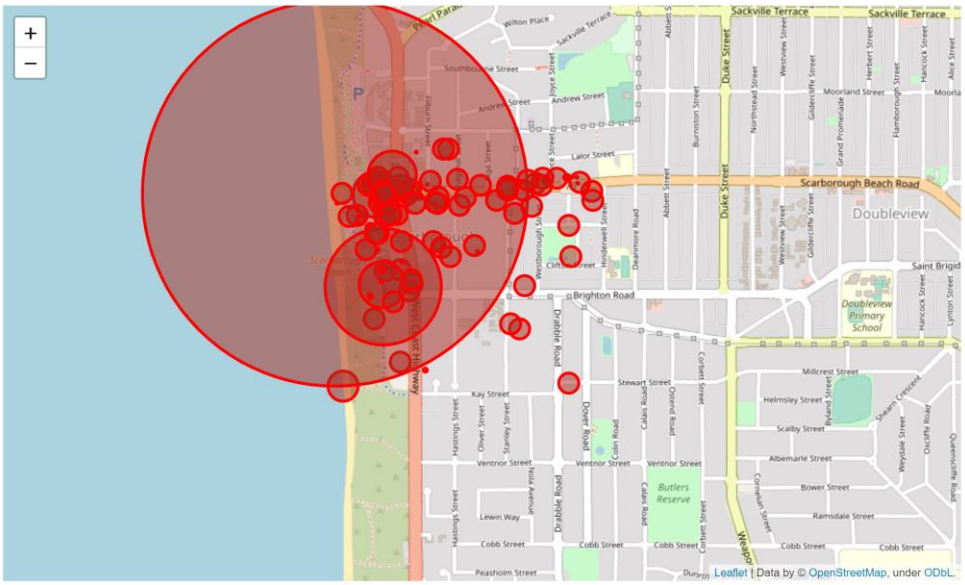

Figure 4: Number of Likes at venues in Scarborough

Outliers were demonstrated using a boxplot and then were removed from the mapping interface

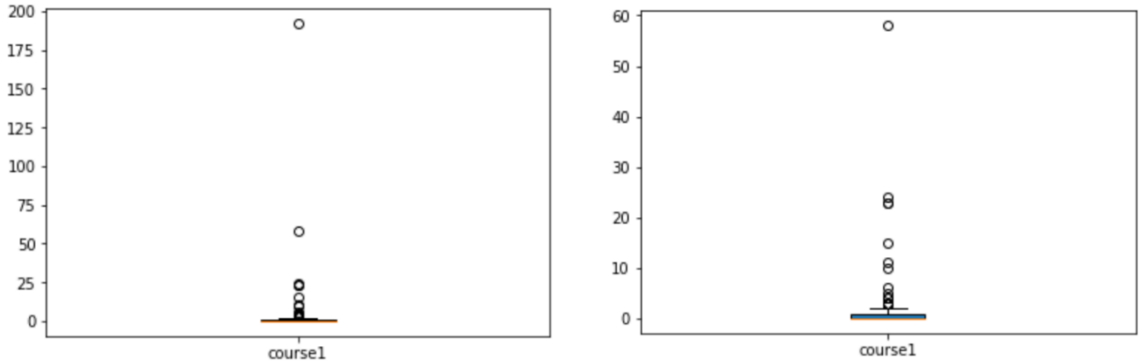

Figure 5: Boxplot of venue like count demonstrating outliers, note the max is now 60 as opposed to 200

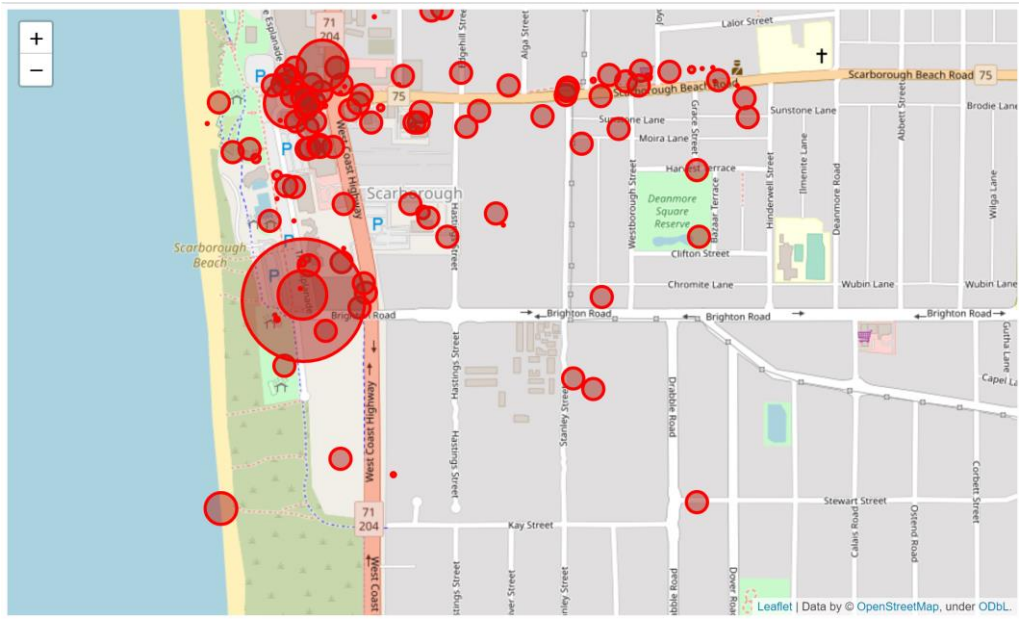

Figure 6: The mapping interface with the outliers removed

4.0 Prediction /Clustering

4.1 Cluster\s that formed

Using an unsupervised prediction technique (kclustering) and mapping the results on folium the epicentres or focal points of activity can be found. Folium mapped the cluster locations which are demonstrated on the map below

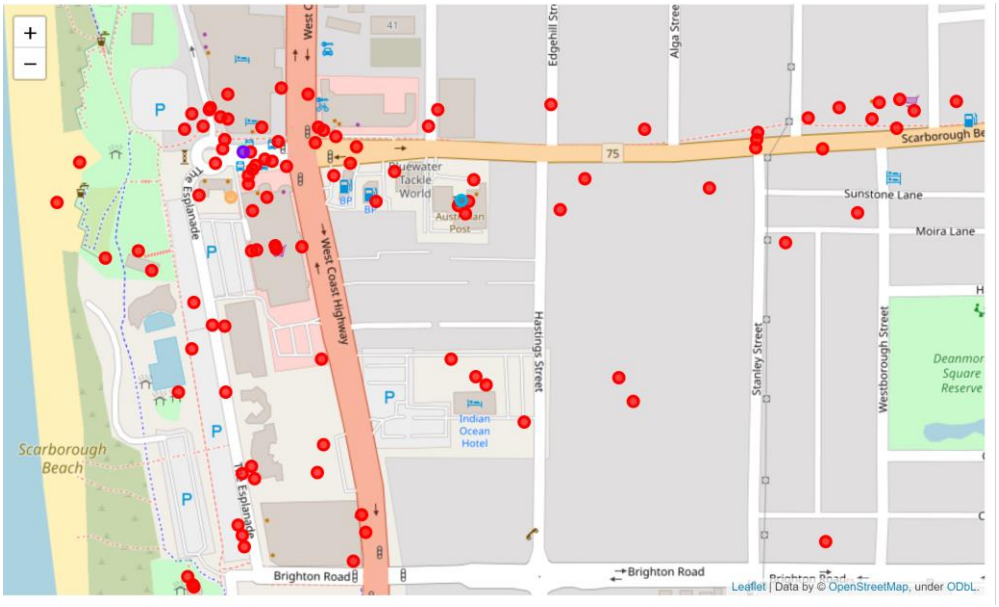

Figure 7: K Clustering results on the mapping interface

4.2 Discussion

From the mapping above it can be seen that none of the clusters are even close to the proposed investor location meaning this would be relying on a unique feature rather than on foot traffic / business from proximity to other businesses. The clusters can clearly be demonstrated closer to norther beach location.

The location of popular or liked venues also supports this argument as although there are some closely located popular venues, the majority of “likes” are located closer to the cluster locations again reduce these as a impact for success of the location

5.0 Conclusions

The recommendation would be to explore other location and opportunity as well as to uncover specifically what the vendor feels drives the supposed success of this location as the data shows that it isn’t due to proximity of other venues or a “precinct” effect.

There are several clusters of business which are popular which are closer to the norther end and this may be a better area to explore opening a new business

While we are using data from a well-documented location, we are also bias as we are only exploring it from one source. For a deeper analysis I would recommending at using another data source to complement these findings

6.0 Future Direction

Investigate other popular locations as well as extract and explore data from other sources which may provide additional information to help explain trending locations and features that predict/confirm this.