Peer-graded Assignment:

Capstone Project - The Battle of Neighbourhoods:

Opening an Irish pub in Brooklyn

Report

# Introduction/Business Problem

New York City is one of the most is the most populous city in the United States and the most densely populated city in the United States. It is extremely diverse in culture and it the most linguistically diverse city in the world and has the largest “foreign born” population of any city in the world as of 2016.[[1]](#footnote-1) In a city so big and diverse, many commercial opportunities exist for entrepreneurship, particular in the areas of the city which are already highly commercial, business districts and entertainment districts. It is known that there are many people of Irish descent living in New York[[2]](#footnote-2), and this report therefore focuses on the commonly found “Irish Pub” which appear in urban centres across the United States. In this “battle of the neighborhoods”, we ask where one might open an Irish pub, not in New York overall, but in Brooklyn which has a long-established Irish tradition.[[3]](#footnote-3) Opening a pub/bar is always fraught with uncertainty as that trade can be fickle and extremely difficult to manage successfully. Using data from Foursquare this report assesses which neighborhoods in Brooklyn would be best for opening an Irish pub. This research would be of interest to venue operators such as bar or nightclub owners looking to create a new Irish-themed venue in Brooklyn.

**Research question:**

What is the best neighbourhood in Brooklyn, NY, in which to open an Irish pub? We assess this by looking at the existing commercial and entertainment activity in Brooklyn neighborhoods where there are Irish pubs now, what other entertainment exists and uses k-means clustering to create segments of neighborhoods in Brooklyn based on venues in those areas to establish where might be best to open a pub in terms of related activities that would suggest footfall and an existing trade in nightlife.

For this task I will leverage the Foursquare location data in relation to the location of its wealth of data on commercial businesses as well as other data sources. The audience for this research would be someone wishing to purchase a property in Dublin. If links between property prices and other factors can be established we can understand what facilities are available as related to property prices. The individual therefore can make a decision based on their requirements.

# ****Data****

The data I will be using to solve this problem will come from two sources:

There are two data sources that will be used in this report:

1. Data that will be acquired via a Foursquare developer account.
2. Shapefile data about the neighborhoods of New York that can be plotted on a Folium map for visualisation

I will first use visualisation methods to show the area at interest and then assess the relevance of the various factors to opening a pub. I will use maps, trend analysis and k-means clustering.

# ****Methodology****

Using the shapefile of New York we created a map of neighborhoods in New York, then narrowed that down to Brooklyn. I then used the FourSquare api to investigate Irish “pubs” in Brooklyn. This was quite an unsophisticated analysis, searching for “pubs”, however it is the common term for a bar in Ireland and in Irish-themed venues around the world, therefore a reasonable assumption to make. The locations of those pubs in neighborhoods was considered and neighborhoods where they tend to be clustered were investigated further. Venues in those areas were investigated as well as trending venues and review of the existing pubs to establish the nature of the competition and if those existing Irish pubs have busy and regular clientele which might suggest a popular area for nightlife.

Once that analysis was complete, the neighborhoods were then clustered using k-means clustering based on a ranking of venues grouped in the dataset. This gave five segments based on similarity of venue.

# ****Results****

The FourSquare search for “pubs” There were large areas of Brooklyn in which there were no Irish pubs. Two particularly popular venues were revealed in the neighbourhoods of Clinton Hill and Boerum Hill. These venues were relatively close to Manhattan where there are particularly large concentrations of Irish pubs. A further analysis of the Clinton Hill and Boerum Hill areas reveals that they are particularly busy areas for venues, particularly nightlife. Analysis revealed that the two Irish pubs identified had considerable footfall which suggests popular areas for nightlife.

The cluster analysis was conducted to consider the nature of these neighborhoods in more detail. Brooklyn was clustered into five clusters based on the frequency and popularity of venue types in those neighborhoods. It became clear that the venues in each cluster suggested the nature of those areas in Brooklyn. Segments 0, 3 and 4 are outliers with few or no venues, all three being very small clusters of an island and empty or warehousing districts. Cluster 1 is clearly residential and domestic in character, the most common venues being banks, parks, pharmacies, food stores, etc. This contrasts with cluster 2 which the most common venues were bars and restaurants therefore areas with nightlife and entertainment, suggesting good areas for a bar. As both Clinton Hill and Boerum Hill are in cluster 2, this suggests that these are good areas for such a venue.

# ****Discussion****

Having regard to the findings in section 5, results, above, it is clear that opening a night-time venue in cluster 2 would have the best chance of success given that it appears to be the area of Brooklyn with the most nightlife and entertainment venues. This is more so than cluster 1 which is primarily residential. That is not to suggest that an Irish pub would not succeed in those areas, but that if a busy venue in an area known to be good for nightlife was required then cluster 2 would be the best area.

# ****Conclusion****

This report has been an interesting exploration of data in Brooklyn. It is difficult to generalise about such a large city, but certainly it is clear that many nightime venues exist with considerable footfall and in defined clusters in the city. Cluster analysis is a very useful tool when analysing the types of areas in cities in terms of the nature of the cultures and lifestyles in those areas. Again it is difficult to generalise. And again that is not to suggest that an Irish pub would not succeed in any area, but that if a busy venue in an area known to be good for nightlife was required then cluster 2 would be the best area. Also this report can conclude that Clinton Hill or Boerum Hill would be the best neighborhoods in Brooklyn to open an Irish pub.

1. <https://en.wikipedia.org/wiki/New_York_City> [↑](#footnote-ref-1)
2. <https://en.wikipedia.org/wiki/Irish_Americans_in_New_York_City> [↑](#footnote-ref-2)
3. <https://www.tripsavvy.com/irish-neighbroods-in-brooklyn-441992> [↑](#footnote-ref-3)