

Battle of Neighborhoods Report

Introduction:

In my project I am going to examine where it is worth opening a new Italian restaurant in Connecticut, US. I will consider the following factors in deciding this. Connecticut is one of the US states with the most Italian-Americans. It has 8 counties and by checking those two with the most people of Italian ancestry, I plan to come to a conclusion in which of them would be a better idea to launch this new venue. In my research, I will take into consideration the Italian restaurants in the selected two counties, check their ratings, explore where the busiest areas are such as, where the cinemas, theatres, hotels, offices are located, as well as markets that could provide the new restaurant with fresh ingredients. I hope to come up with a recommendation where a new Italian restaurant would be best located. This project would interest investors willing to start a new business in the form of an Italian restaurant.

Data:

I will use a .csv file of Connecticut listing its 8 counties along with their geographical details (Longitude, Latitude). For the demographic data, I will create a table to show the distribution of Italian-Americans within the counties based on the source of the United States Census 2000 (<https://www.niaf.org/culture/statistics/italian-american-populations-in-select-u-s-counties/>). After choosing New Haven, CT as the recommended county, I will examine its neighborhoods based on a table listing them by web scraping ("https://github.com/CT-Data-Haven/nhv2016/blob/487ebd54288008f3ecd48a80a2303c136ba68894/public/data/2016_nhv_display.csv"). I will explore the top most common venues by K-means clustering, the already existing Italian restaurants, their ratings, and other venues such as cinemas, theaters, hotels, business service and if any market exists in an optimal location with the help of the Foursquare location data.

Methodology:

I have applied Foursquare location data in my project to decide in which of the two counties Italian restaurants are more popular. After choosing the county for the new business, I have performed segmenting and K-means clustering on its neighborhoods.

By checking the top 5 most frequent venues of occurrence as well as the top 10 venues of the clusters, it could be clearly seen that Italian restaurant is the top venue in the majority of the neighborhoods as I expected given the population's Italian origin.

I have also utilized the groupby method to analyze the data.

Foursquare location data provided the Italian restaurant venues for the two counties and all the venues for the selected county's neighborhoods.

When I decided which county would be a better choice for a new Italian restaurant, I selected two Italian restaurants in two central neighbourhoods based on the map that visualized the venues by Foursquare, in which Italian restaurant is the number one for the ten most common venues. After checking their ratings, I explored the nearby venues for both of them to consider which one could be more frequented.

Finally, I have checked correlation, what the relationship is between the population of Italian ancestry and the number of Italian restaurants.

Results:

New Haven and Fairfield are the counties in Connecticut where most people with Italian ancestry live. Due to the fact that most of the people with Italian ancestry live in New Haven and the number of Italian restaurants is significantly more than in Fairfield (63 in New Haven and 15 in Fairfield), I would consider to open an Italian restaurant in New Haven.

Italian-Americans are represented the most in New Haven and this is reflected by the fact that the majority of its neighbourhoods has Italian restaurant as the number one most common venue. A good quality Italian restaurant

can be successful at a lot of locations in an environment like this. In my project, I have taken into consideration that the optimal location should be one where a market is nearby to provide for the fresh ingredients, museums, concert hall, theatres, movie theatres, hotels, sport fields, schools, and parks etc. as these are all venues that could attract more people to the restaurant.

Within New Haven there are two neighborhoods that seem to be a reasonable choice after exploring the nearby venues and the Italian restaurants. One of them is Wooster Square where there is a Farmer's market nearby that could provide the fresh ingredients for the restaurant, however, there is an already very successful and traditional Italian place in the vicinity with a high rating from the guests. The other one is Downtown New Haven (Central Business District) with business services, parks, hotel, stadiums as well as a centre of cultural life having more museums, theaters, movie theaters while in Wooster Square this does not seem to be the case. The market and several grocery stores are still close to provide for the fresh ingredients.

Finally, my recommendation would be Downtown New Haven as it seems to be a more favourable place to open an Italian restaurant.

Correlation supports that the higher the number of Italian-Americans, the higher the number of Italian restaurants (with a P-value of <0.001 : there is strong evidence that the correlation is significant and the linear relationship is very strong (~ 1.0)).

Discussion :

Italian restaurants and Italian cuisine are very popular not just among people of Italian origin. Opening such a restaurant is less riskier than one with any other kind of kitchen. By ensuring that the selected location is even at a location where the residents are most of the Italian-Americans, it predicts a higher success rate. While any neighborhood in New Haven, CT could be considered for such a new venue, my preference was to choose a central location that is one of the busiest parts of the county. Due to this we can calculate with not just the residents, but with the people who commute to work everyday and tourists too. Offices, hotels, parks, baseball stadiums, golf course may all provide guests as well as the nearby cultural places like the Art Gallery, Art Museum, two theatres, and a movie theatre. Data could be narrowed down by adding more aspects to the research for example, costs like renting costs of a restaurant in the different neighborhoods that was out of my scope this time.

Conclusion:

Gathering data on the population and neighborhoods of Connecticut led me to select one of its counties, New Haven, to explore further the location. I have analyzed New Haven's neighborhoods due to its Italian-American population and nearby venues. After choosing one of the busiest, central neighborhoods, my methods included segmenting, k-means clustering, Foursquare location data analysis and correlation. With taking into account all the data I gathered, I concluded that the Central Business District/ Downtown New Haven would be my recommendation to the investors looking for such opportunity.