1. Introduction

Being the son of migrant workers having to relocate was a constant through my childhood and even now when trying to find employment. For most of the moves we did growing up there seemed to be a relative or family friend living in our next destination. This provided me and my family a sense of camaraderie, a sense of belonging even as we travel and joined new communities. As such to make it easier to other Latinos to find a place of belonging this project aims to cluster and segment Kansas City's neighborhoods to determine potential locations to live or enjoy a night out.

Hispanic migration to the Kansas City area dates as far back as the 1830s, when the Santa Fe Trail opened up trade from Mexico. But the first major wave of immigration came with the establishment of railroads in Kansas City, Kan., and Kansas City, Mo. In the past 20 years, turmoil in Central America and Mexico has brought a new wave of immigration to the Kansas City area, revitalizing the historic Hispanic communities.

Target audiences for this project does not limit to Latino and Hispanic people but everyone. People could simply decide to look for a restaurant all the time because they are enjoy the food.

2. Data

Data Required to resolve the problem

In order to make a good choice in finding a location to reside in the Kansas City area the following data is required:

1. List/Information on neighborhoods from Kansas City with their respective Geodata (latitude and longitude)
2. Listed apartments for rent in Kansas City area
3. Venues and amenities in the Kansas City neighborhoods (e.g. top 10)
   This data will provide a set of parameters to determine the location of Latino centric communities where anyone can relocate. Determining the venues in these neighborhoods will allow others who only seek for a place to eat or enjoy a night out with information to make their decision.

3. Methodology

After acquiring a data set the CSV file was formated to include only the data of importance to determine the neighborhoods of interest. This was done initially by uploading the dataset, this was then read using the pandas library and converted onto a data-frame for further analysis and manipulation. This was followed by establishing a map with an overlay of all the neighborhoods in the list using the Geodata and Nominatim. To determine potential neighborhoods to live location data using the foursquare API was used to determine the most popular venues in the neighborhoods. These were then listed and sorted determining the most popular establishments per neighborhood. Once this was done a further evaluation of the neighborhoods which had in their top ten venues associated with Hispanic/Latin American culture were evaluated to determine potential rental listings. This was done by clustering the neighborhoods and determining which clusters are potentially more relevant to our search. An assumption was made to establish a cut off and limit the rental search using the most common venue as an indicator of community makeup. With this in mind a data-frame with the neighborhoods with Mexican restaurant as their most common venue was established.

4. Results

In this analysis we have explored the locations of Latin or Hispanic centric communities by assuming the most common venue establishes the majority of the make up of the area. Using clustering based on venues within walking distance (defined as 400m radius) we have clustered the neighborhoods into five groups. Of these groups, one has more than one instance of Hispanic communities staples like restaurants and and other amenities that imply the make-up of the community.

The largest clusters share a number of common features. Mexican restaurant, South American venues, as well as bakeries food trucks and grocery stores.

Being specific about the type of each area beyond these shared features is difficult but we note a few differences using the top ten venue types for each cluster: These clusters showed a varied make up, the one which had most in common was cluster 8 where the majority of the determined neighborhoods of interest formed part of.

5. Discussion

Based on our results we have identified five neighborhoods where as per our assumption appeared to be Hispanic or Latin centric communities. It was established that a correlation with rental information could had made this case stronger but the rental data from zillow was unavailable to scrape due to the API restrictions.

In theory, using these results, Mexican Restaurant could characterize other areas where Latin American or Mexican immigrants might consider for a potential area to rent or enjoy a night out eating and enjoying the community.

This data exploration is not significantly strong secondary correlations with other data sets should be employed to establish a strong correlation between the assumptions and the findings of this exercise.

6. Conclusion

The aim of this project was to determine the potential location of Hispanic or Latin centric communities with the assumption that the venues in the area showed the potential make up of the community. With this assumption it was determined that 5 neighborhoods met the description. To establish a stronger correlation a secondary data set should be used to dive deeper in the ethnic make-up of these neighborhoods. As time being a limiting factor for this project it was not possible to establish this stronger correlation with the data set acquired for this exercise.