

IBM Capstone Project

Comparison of 2 St Louis Apartment Complexes

Amy Guenther

● Introduction

- **Business Problem** When companies are looking to choose properties to invest in, they sometimes look for what surrounds the property and the level of desirability of each surrounding venue. According to ApartmentRatings.com and similar sites, when a person is looking for an apartment he/she should take into consideration its proximity to “shopping and other services” as well as “family amenities”.
- **Interest** This project assumes a client is deciding between two apartment complexes and would like to know more about the amenities close by to each. Special consideration is given to food venues as this client does not like to eat at home as well as entertainment such as sports venues or artistic entertainment. Although this example is purely fictional, the process could be repeated for anyone who is looking to assess the location of different apartment complexes.

● Data

- **Data sources** This data is mined from Foursquare, applying different sorts to organize the data and improve its usefulness.
- **Data cleaning** Two apartment complexes were chosen which are located in different areas of the city. The table’s columns were cleaned to better allow for sorting and ability to be understood. For example, the prefix of “location” was removed from the column “location.lat” and “location.lng”.

● Methodology

- Using Foursquare, I searched the entire St Louis area and found the two apartments that I would be analyzing. I noticed that Cityview Apartment Complex was closer to the St Louis Marker. This indicates that it is closer to the heart of the city and downtown. Grandview Tower Apartments however were much farther from downtown. I was hopeful this difference would give good variance in the results.

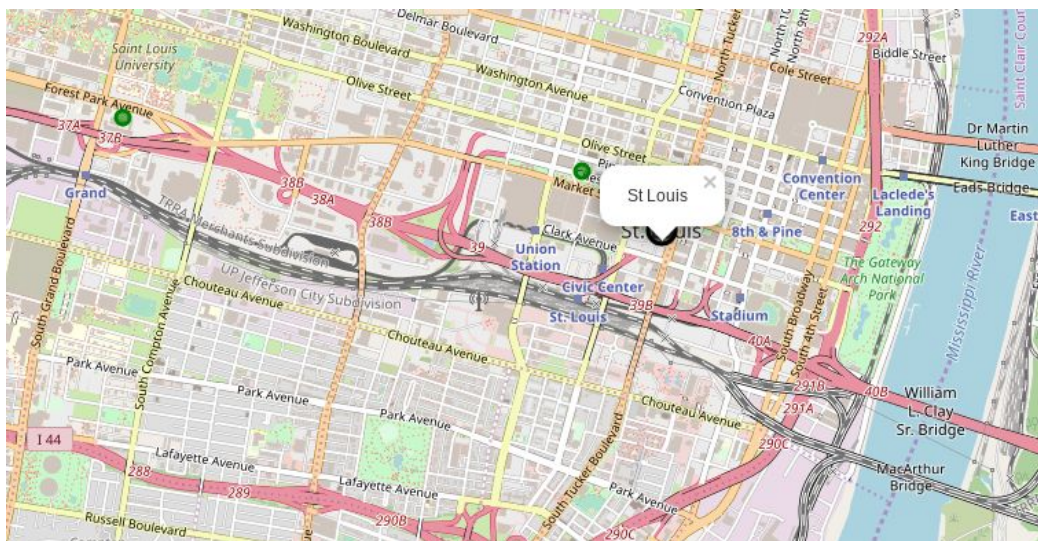


Figure 1- St Louis with pins for two apartment complexes

● Results

● CITYVIEW APARTMENTS

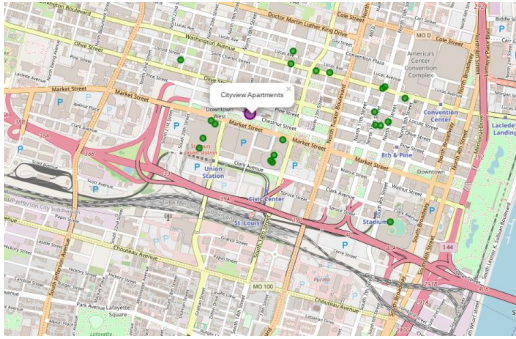
- Calling up the data using Cityview Apartments (cv) as my central point, I called up all the venues close to the complex. There were 30 of them. I began cleaning it using standard methods to make the columns more readable. This included eliminating the “location” indicator before the columns called “location.lat” and “location.lng”

- I then created a dataframe with my resulting data and examined it for ways to sort. I realized that the “categories” column had a wide variety of terms. Some, such as “hotel”, “train station” and “post office” did not seem relevant to our intended comparison of food venues and entertainment. I eliminated those and reset the index. Others needed to be grouped into smaller bins in order to be comparable. For example, there were many different venues that could be labeled as “food” such as “Whisky Bar” and “Coffee Shop”.
- I also noticed that several of the category descriptors could fall under an “entertainment” category. These included venues such as sports attractions and parks. This was different from music/arts which I determined after looking at the second apartment complex; it has several places which could be classified as music or art venues. Because I knew I would be comparing the two, I wanted to make sure to keep the grouped categories as similar as possible.

	categories
Food	13
Entertainment	6
Music/Arts	1
Grocery Store	1

Figure 2- Cityview Apartments Dataframe Results

- I wanted to have some visuals to show the clients so in addition to the basic dataframe with numbers, I made a map using Folium to show the locations of all the venues to the apartment complex. I also made a bar graph to show the data from the dataframe in a more visual way.



Figures 3- map of Cityview and venues

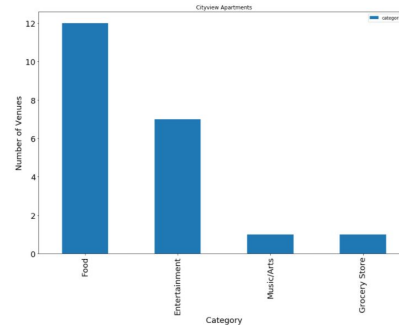


Figure 4- bar graph of results

- Cityview- Results- It is clear that Cityview Apartments have many options for food which is desirable to the client, but not much else. It does have quite a few general entertainment items however.

● Results

● GRANDVIEW TOWER APARTMENTS

- Calling up the data using Grandview Tower Apartments as my central point, I called up all the venues close to the complex. I began cleaning it using standard methods to make the columns more readable. This included eliminating the “location” indicator before the columns called “location.lat” and “location.lng”
- I then created a dataframe with my resulting data and examined it for ways to sort. As with the Cityview Apartments, I realized that the “categories” column had a wide variety of terms. There was one hotel which I eliminated but not as many other outlier venues. As before, I grouped the food venues into a category of just “food”. This complex is close to St Louis University so many of its closest venues were part of the college. While this could be part of food or entertainment, I decided to sort these into their own category of “college”. I did a search for all entries which contained the string “college”, made a new database with those entries, deleted them off the original database and then concatenated the college database with the original.
- After sorting out the “college” venues, I noticed there was a significant number of entertainment options that were specifically music or art related. These were titled as "Theater", "Concert Hall", "Jazz Club", "Rock

Club", and "Performing Arts Venue". I grouped these together as music/art.

categories	
Food	12
Music/Arts	7
College	4
Neighborhood	1
Parking	1
Monument / Landmark	1
Health & Beauty Service	1
Church	1

Figure 5- GVT dataframe

As with before, I wanted to have some visuals to go along with my findings, so I made a Folium map and a bar graph.



Figures 6- map of Grandview and venues

Figure 7- bar graph of results

- Grandview Tower Apartment had a very different range of venues to choose from as compared to Cityview.
- Because of its proximity to the university (a one block walk according to the [website](#)), there are many school related venues and there are also

more artistic venues such as clubs and theaters.

● Discussion

- The two apartment complexes were very different based on their close venues.
- They both had similar amounts of food venues (cv=13 and gvt=12)
- CV had more sports related entertainment and parks listed, but GVT had more arts/music/theater venues and is closer to a major college campus
- It is my recommendation that the client consider investing in the Grandview Tower Apartment complex. I feel that this complex has more to offer based on the desired traits originally stated by the client.

● Conclusion

- In conclusion, there are many factors that go into choosing an apartment location. Limiting the factors to entertainment and food choices doesn't necessarily tell the whole story. It is my recommendation that a further study be done with more input upfront about what is most desirable. Perhaps assigning points values to the different venue categories and trying to calculate in a more detailed way.