*Final Report: Capstone Project – The Battle of Neighborhoods*

# My Kansas City Sports bar Venture

## 1. Introduction/Business Problem

In Superbowl LIV the Kansas City Chiefs trailed the San Francisco 49'ers, 20-10, with 17:40 left in the game. But faithful fans knew that the Chiefs were still in the game. In the divisional round of the playoffs the Chiefs trailed the Texans 24-0 and trailed the Titans 17-7 in the AFC Championship game before coming out ahead in both games. The final score was Kansas City Chiefs 31, San Francisco 49ers 20. Kansas City had won its first Superbowl in *fifty* years again coming from behind to do it.

Following the game, Donald Trump made a gaffe on his Twitter feed congratulating Kansas City and saying they made the 'Great State of Kansas' proud. Kansas City is actually in Missouri, not Kansas. But then I realized that perhaps not unlike Donald Trump I was a Chiefs fan and of their coach Andy Reid but I knew little about Kansas City, Missouri. Being enrolled in the IBM Applied Data Science courses I decided to create a hypothetical 'fan mission' to open up my own Chiefs-themed sports bar in Kansas City. Given that Kansas City borders Missouri and Kansas I have included the requirement that my sports bar, being Kansas City Chiefs themed, should have a proper Kansas City address as a fan might require.

Being completely new to Kansas City, not having even traveled there before, I can't do this by feel or prior knowledge. I must let the data tell me a story on where to locate my new sports bar venture. I will need neighborhood data to show me patterns of where the venues would cater to a sports bar but be careful to avoid competition from other local bars, especially sports bars. This will minimize the risk to the success of my new venture.

## 2. Data

I was able to locate a Wikipedia page on the neighborhoods of Kansas City, Missouri here: https://en.wikipedia.org/wiki/Neighborhoods_of_Kansas_City,_Missouri. Parsing that page to data I created a list of neighborhoods which was easiest just as an Excel csv file attached to the project as a data source in IBM Watson Studio. I used Nominatim geolocator services for latitude and longitude data on these neighborhoods. Nominatim is a search engine for OpenStreetMap data. The neighborhood data needed to be scrubbed a bit as Nominatim could be a bit unreliable. For example, neighborhoods such as 'Waldo' and 'Waldo West' were combined as 'Waldo' to avoid duplicate geocoding. The neighborhood 'Town Fork Creek' was changed to 'South Town Fork Creek' so that Nominatim returned valid coordinates. The Foursquare API was used to obtain a list of venues in those neighborhoods to establish patterns in the neighborhood venue data along with Folium maps for visualization of the data. Visualization of those patterns was used to determine the best location for

the sports bar venture. A discussion of the exact criteria used for selection will follow in Section 5 of this report.

## 3. Methodology

First, the listing of neighborhoods was gathered from the Wikipedia page and worked into an Excel CSV file that became an input file for the project on the IBM cloud. The OpenStreetMap Nominatum Python API was used to geocode the neighborhood data to give the locations of the neighborhoods. The neighborhoods were plotted to get a feel for the layout of the city. Next, the Foursquare API was used to gather lists of venues for each of the neighborhoods and the returned data were grouped by venue frequency. K-means clustering, a form of unsupervised learning was tried on the venue data in order to obtain patterns in the neighborhood venues. Venues that contained sports bars as their top venues were plotted. These neighborhoods might be considered higher competition for the venue.

## 4. Results

A plot of the Kansas City neighborhoods from the geolocation data is presented in Figure 1. Because Kansas City is on the border of Missouri and Kansas, the neighborhoods primarily run north to south. The city is located on the border with Kansas. It has a land border with Kansas to the south and the Missouri river runs northwest as a water border which is the reason why the neighborhoods fan out to the north. Arrowhead Stadium, home of the Kansas City Chiefs, is shown with a location marker.

Figure 2 depicts the visualization of K-mean clusters (k=5) on the venue data for Kansas City. You can see one primary large cluster in blue runs from north to south. The blue cluster is a mix of restaurants, sandwich shops, pizza joints, clothing retail, and grocery stores – exactly as one might expect in a high-density Midwestern town. Another notable cluster is the purple cluster which is on the fringes of the town which is composed of cheaper-rent operations such as pool service companies and yoga studios.

Figure 3 depicts the neighborhoods where “Sports Bar” was a top-10 venue along with a marker for Arrowhead Stadium. The majority of sports bars (as a percentage of venues) run north to south in the main cluster.

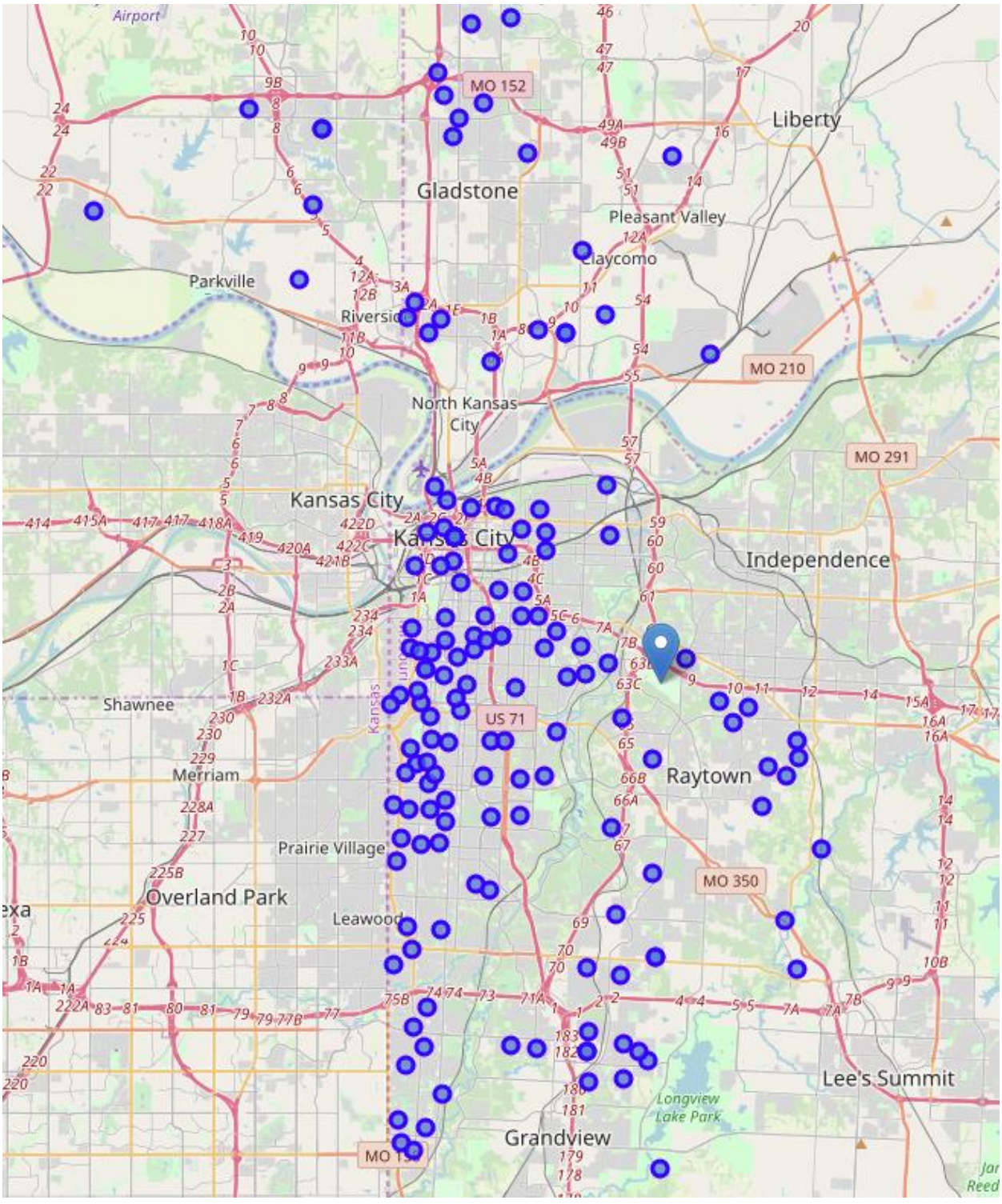


Figure 1: Neighborhoods of Kansas City, Missouri. Arrowhead Stadium, home of the Kansas City Chiefs, is shown by the location marker.

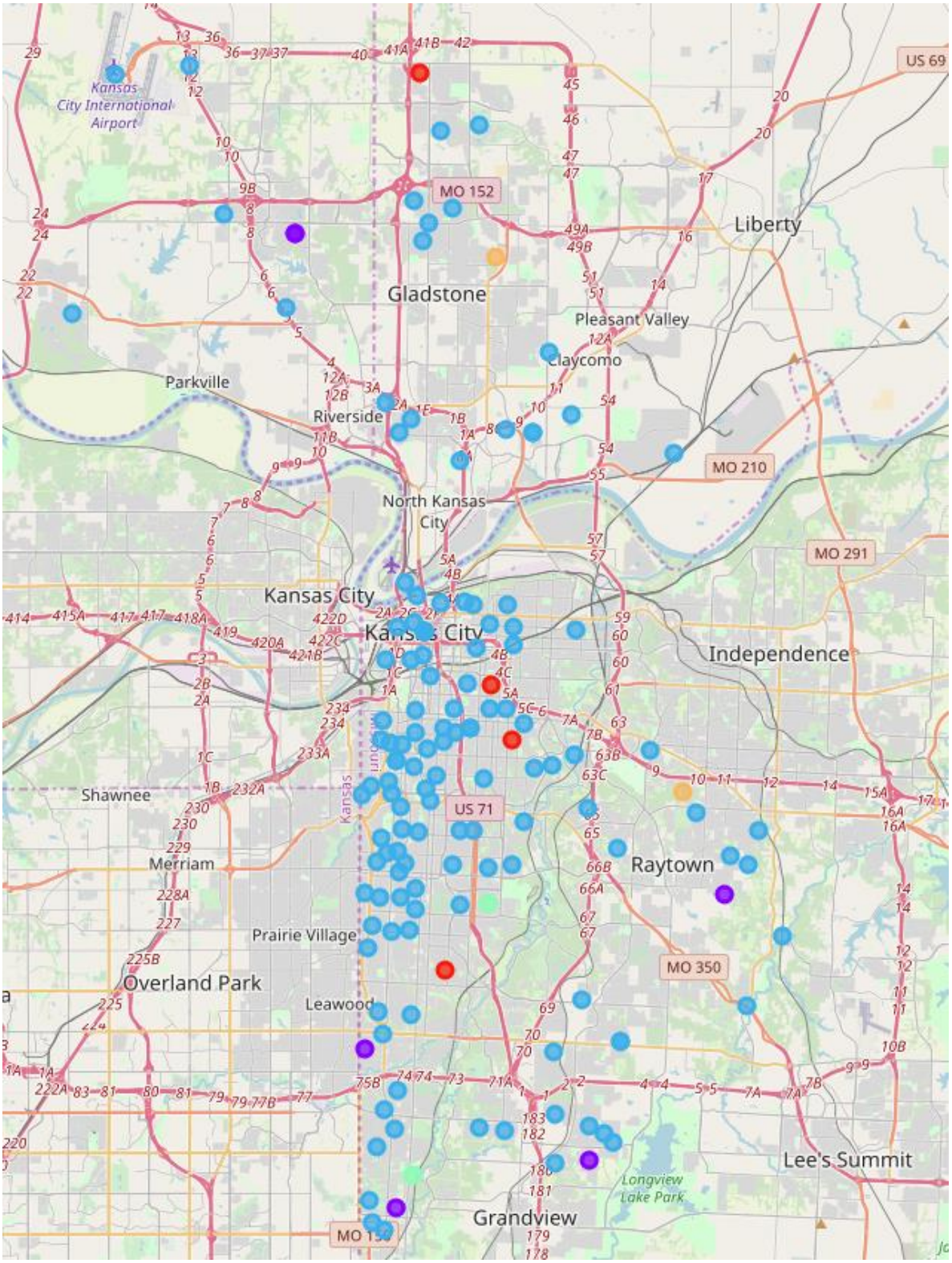


Figure 2: K-Means (k = 5) clusters for neighborhoods of Kansas City, Missouri. The primary cluster in light blue is a mixture of restaurants, bars, groceries, coffee shops, and retail.

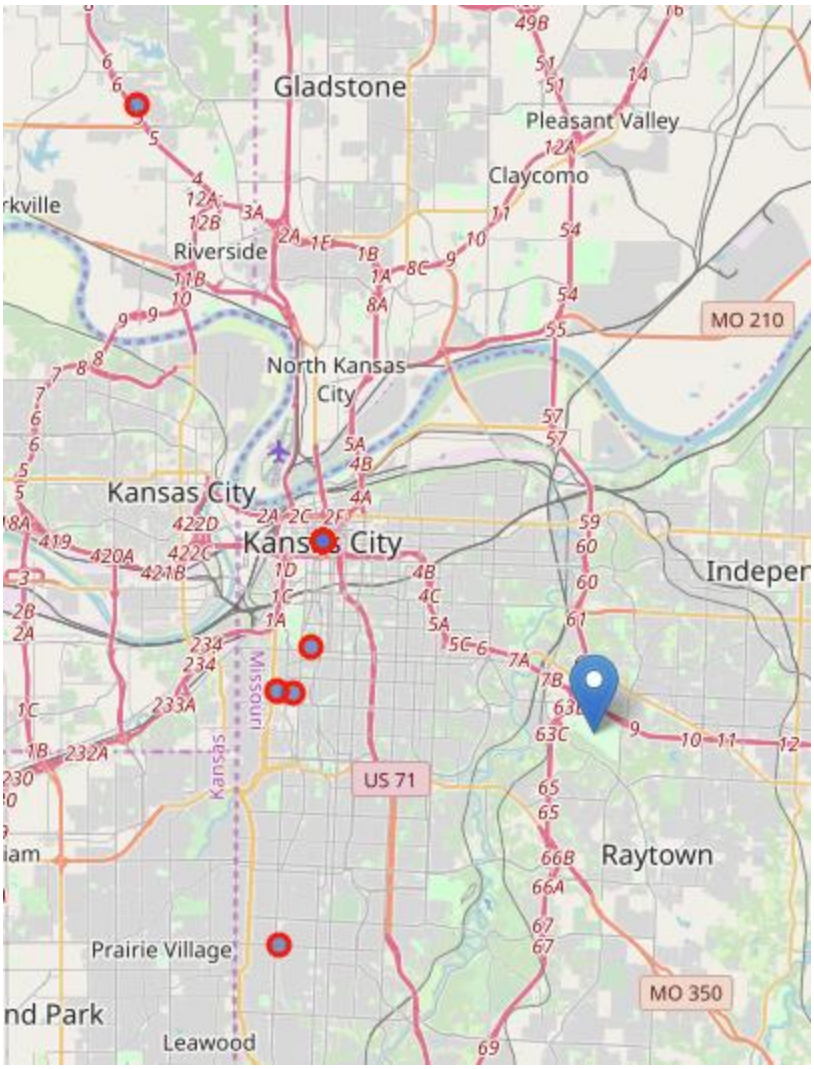


Figure 3: Locations of neighborhoods in Kansas City, Missouri where “sports bar” was a top-10 venue. Arrowhead Stadium, home of the Kansas City Chiefs, is shown by the blue location marker.

## 5. Discussion

The venues run along the neighborhoods mostly north to south along the town lines and likely mix with the venues not included on the Kansas side of the Kansas/Missouri border. The primary cluster is a mix of restaurants, bars, groceries, coffee shops, and retail with no significant defining features although there may be economic boundaries not covered in the location data that could indicate more upscale neighborhoods than others. The large cluster is indicative of a somewhat homogenous downtown at least in the types of venues. This lends to the impression that there may be many suitable locations in this downtown cluster although one should take care to avoid the competition with the bulk of the sports bars running north to south along the cluster. A curious note is that one gets the impression that they are big on their barbeque joints in Kansas City as that venue appeared in several neighborhoods so

any sports bar venue may need to include local barbeque favorites to be competitive which is an unexpected detail revealed by the data about the local culture.

The primary cluster runs north to south as do the prevalence of neighborhoods with sports bars. However, as one travels from city center to the southeast towards Arrowhead Stadium the competition seems to thin out a bit and there may be better opportunities to open a sports bar in the neighborhoods which run along Interstate 70 towards the stadium between Interstate 71 and Interstate 435. These neighborhoods offer the required Kansas City address and also avoid competition from the venues on the Kansas side of the border.

## 6. Conclusion

The neighborhoods along I-70 may be suitable for the new sports bar venture. In this area one could capitalize on stadium traffic during the season yet be close enough to downtown while avoiding some of the competition of the north-south sports bar venues. This area includes the neighborhoods of Knoches Park, Ingleside, Sante Fe, Washington-Wheatley, and Dunbar. One of these neighborhoods could potentially be suitable for a sports bar and would be a good starting point to start looking for locations for the new venture. More research would need to be done looking at rent data in this region to determine if the costs are reasonable. Go Chiefs!