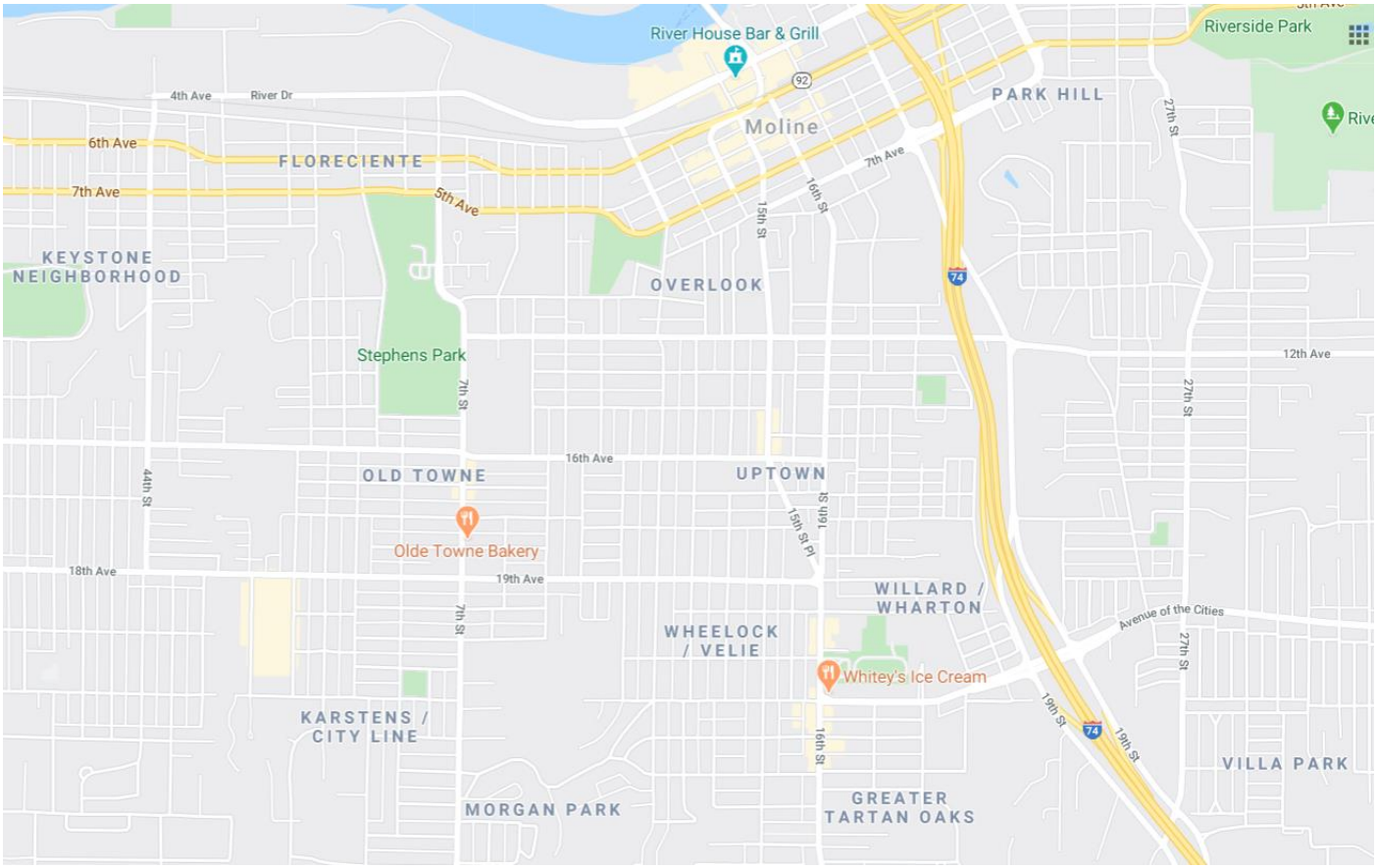
Data

To help determine the perfect location, we will need location data. Moline is made up of many historic neighborhoods (https://en.wikipedia.org/wiki/Moline,_Illinois#Neighborhoods), Unfortunately, there may not be many easily accessible data sources that list or detail these neighborhoods, like there was with the previous Manhattan and Toronto assignments. Once we have a list of neighborhoods and their locations, we can connect to Foursquare and collect data on the most common types of venues in and near these neighborhoods. We will want this data to make sure that we are not building our coffee shop too close to an area that already has lots of coffee options (like other cafés or coffee shops, and to a lesser extent donut shops, diners, and gas stations). To avoid having to go through every single neighborhood in detail, we will use K means cluster analysis to cluster the neighborhoods into similar groups based on location and the most common venues.

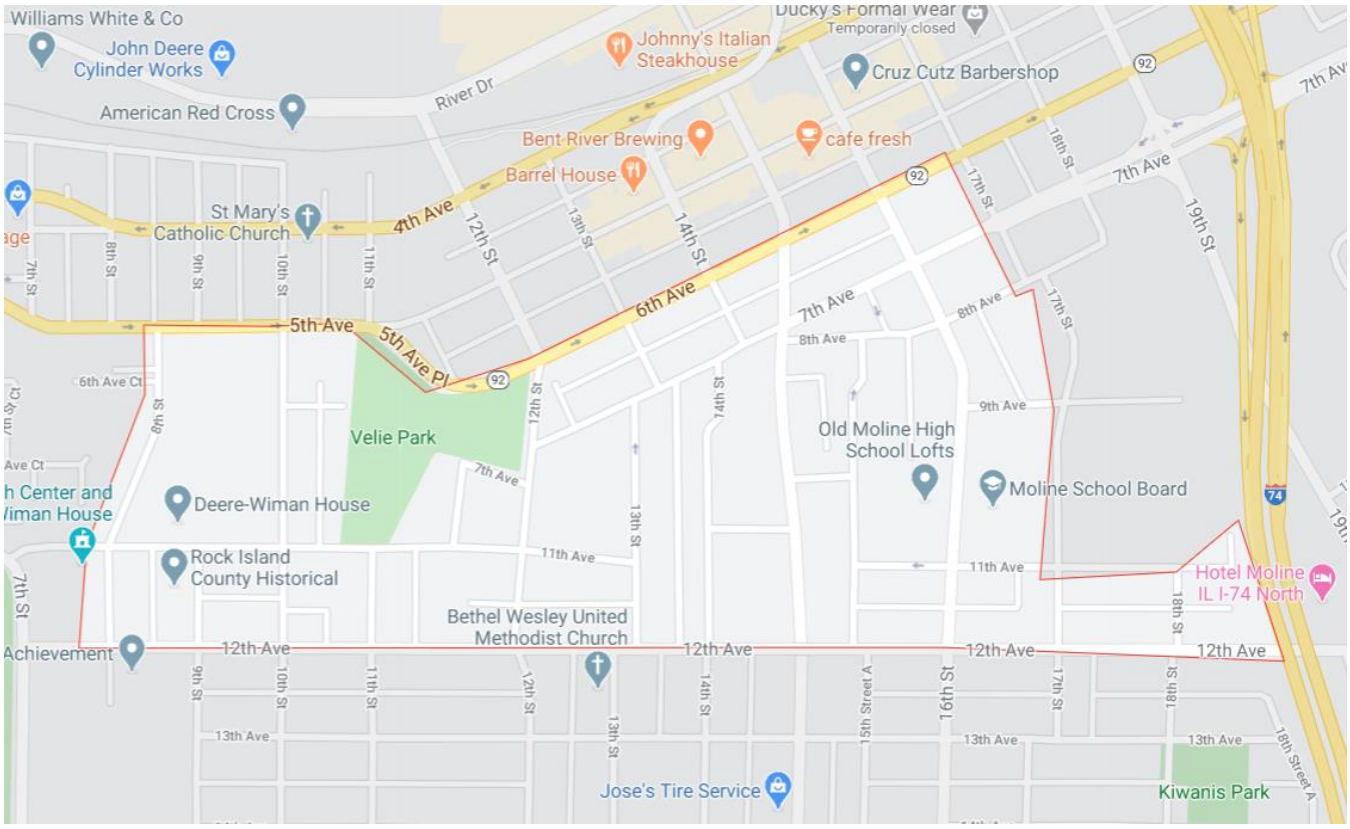
Since robust data on the neighborhoods in Moline did not exist, I ended up creating it myself. First, through Wikipedia and some additional internet searching, I was able to compile a list of 26 historic neighborhoods for Moline:

1. Hamilton Heights
2. Wildwood
3. Prospect Park
4. Park Hill
5. Forest Hill
6. Highland
7. Villa Park
8. Green Acres
9. Molette
10. Rockview Estates
11. Homewood
12. Heritage
13. Stewartville
14. Deerview
15. Walton Hills
16. Olde Towne
17. Uptown
18. Willard/Wharton
19. Karstens/City Line
20. Morgan Park
21. Greater Tartan Oaks
22. Overlook
23. Floreciente
24. Keystone Neighborhood
25. Wheelock/Velie
26. Downtown

After compiling the list, I needed to get estimates for the longitude and latitude of the center of these neighborhoods. This process was rather manual. Luckily, many of the neighborhoods were already labeled on Google maps, as seen below:



Selecting the neighborhoods that were already labeled in Google maps created a nice outline of the neighborhood (as seen in the screenshot below for the Overlook neighborhood), making it easy to select a location roughly in the center and documenting its latitude and longitude.



For locations that were not labeled on google maps, I had to do extra searching on the internet to figure out the rough location of those neighborhoods. Real estate websites were usually the most helpful for finding houses in a given neighborhood, and then I would use Google Maps to find the latitude and longitude of the rough center of that neighborhood. Finally, I was able to create my own csv file that listed all 26 historic Moline neighborhoods, as well as their corresponding latitude and longitude data. To avoid having to go through every single neighborhood in detail, we will use K means cluster analysis to cluster the neighborhoods into similar groups based on location and the most common venues. Using Folium, I was able to create the following map detailing the location of the neighborhoods: