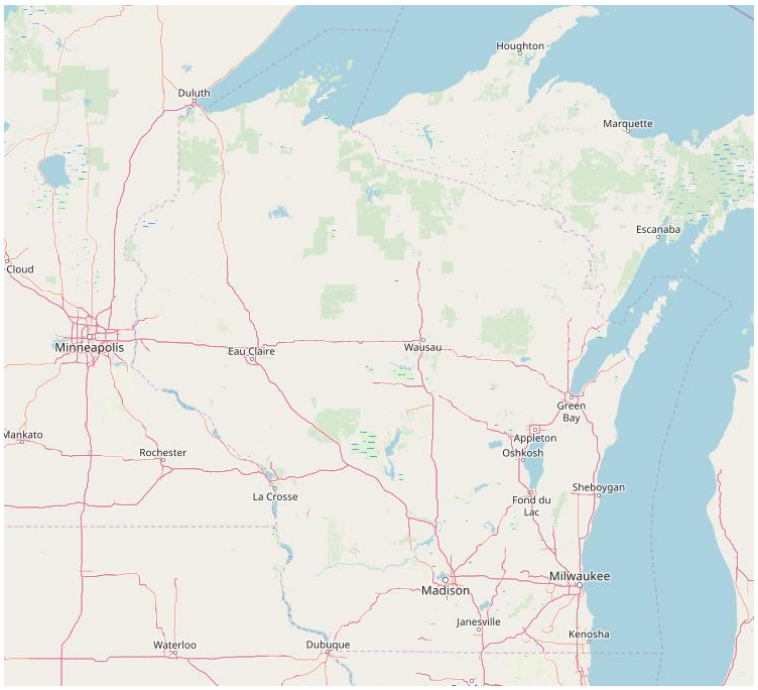


Where Should We Retire?

APPLIED DATA SCIENCE CAPSTONE

Karla K. Arndt
Applied Data Science Specialization
February 22, 2019

1. Introduction

Even though retirement may be far into the future for many taking this class, it is not for some of us. In fact, it's right around the corner! In addition, I belong to an organization that helps seniors plan for these big decisions so know a lot of people that need help in their retirement location planning.

There are many factors that affect when to retire such as financial and health. But, "where to live" is also a major decision. Many choose to live their retirement years from their current residence. Some may become snow birds and live a few weeks or months in a warmer climate, while others may travel to foreign destinations.

"Where to call 'home base'" is the focus of this project. For those who desire to relocate from the city they've called "home" for many years, the decision can be daunting. Even cities that seem similar on the surface can lead to different lifestyles and entertainment opportunities.

This project could be generalized to allow for other seniors to plan for their optimal retirement location, but for this iteration, I want to keep it personal to make it more realistic. That is, I'm going to focus on my husband's and my hobbies and desires for a retirement location. We know we want to retire in Wisconsin, but we don't know where.

I want to find some specifics to help us make our decision such as the following:

- Fishing. There is a lot of available fishing in Wisconsin, but we want to be able to fish for walleye. We also have a boat and need access to a public boat ramp.
- State parks. We hike and camp a lot. I want to explore state parks in my retirement.

Other criteria we are considering:

- Location close to family. We would like to be within 2 hours of Germantown, Wisconsin.
- The city we ultimately chose to call "home" during our retirement needs to have good Italian restaurants. I have Celiac disease and Italian restaurants are generally very accommodating for Celiacs.

2. Data

Finding data to manipulate is key. I want to be able to merge the data so that I have one map with all the information.

Data Sources and Cleaning

I found a fish map by type of fish on the Wisconsin DNR website. I was able to download the information I needed into a csv file and I stored it in github as walleye.csv. The data included the lakes' latitude and longitude and if it had a boat landing. Basically, it had everything I needed right in one file!

```
- https://dnr.wi.gov/lakes/lakepages/Results.aspx?page=walleye
- walleye.csv -- The file on github that I created from the link above.
```

The columns in the table are as follows:

- 'Waterbody ID Code (WBIC)'
- 'Lake Name'
- 'Size (Acres)'
- 'Official Max Depth'
- 'Official Mean Depth'
- 'Latitude'
- 'Longitude'
- 'Public Landing'
- 'Public Beach'
- 'Public Park'
- 'Fish Present'
- 'Lake Type'
- 'Water Clarity'
- 'County'

The columns I was interested in are the following:

- Lake Name -- The name of the lake
- Size (Acres) -- We don't want to fish on anything less than 10 acres
- Official Max Depth -- We don't want to fish on anything that has a maximum depth of less than 10 feet
- Latitude -- The latitude
- Longitude -- The longitude
- Public Landing -- The value must be 'Yes' so that we can use our boat
- Fish Present -- This isn't necessary. It may seem necessary, but the data was already gathered for lakes that have walleye; therefore, it's redundant.

For the state park data, I used the wiki page below that lists the state parks in Wisconsin.

```
- https://en.wikipedia.org/wiki/List_of_Wisconsin_state_parks
```

It lists more than just state parks such as recreation areas. We are not interested in those; therefore, I will need to ensure I am extracting data for only state parks.

3. Methodology

I analyzed each type of data separately and then combined it with other factors that meet our criteria which I discuss in the following section.

Exploratory Analysis

Walleye Data

I built a dataframe with the fields listed above from the walleye.csv file using pd.read_csv. After dropping columns I didn't need, the data looked as follows. Only the first few rows are shown.

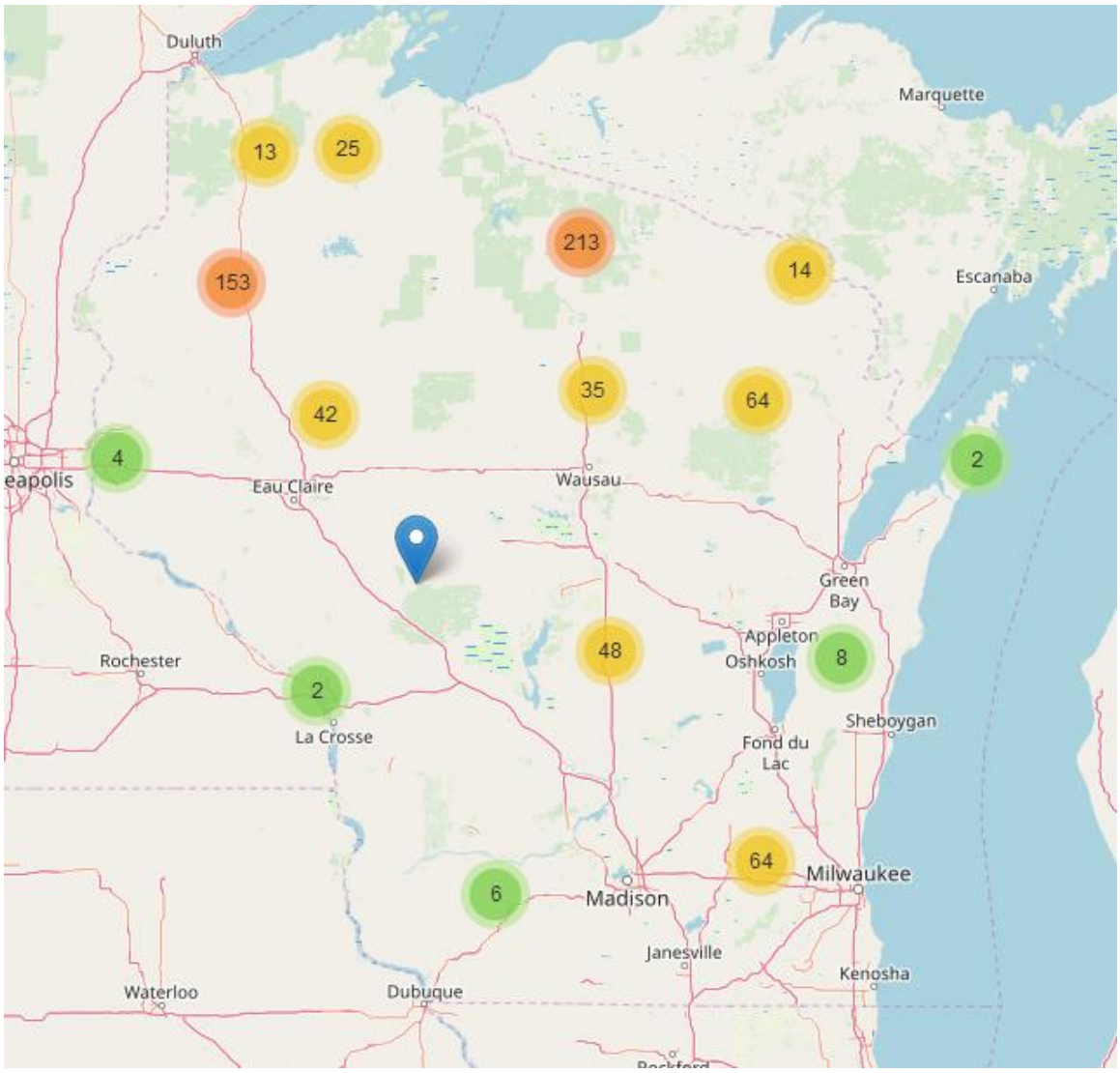
	Lake Name	Size (Acres)	Official Max Depth	Latitude	Longitude	Public Landing
0	Alder Lake	264	30 FEET	46.088099	-89.819586	Yes
1	Aldridge Lake	142	12 FEET	45.865941	-89.301276	No
2	Alexander Lake	618	36 FEET	45.186975	-89.757128	Yes
3	Allequash Lake	406	24 FEET	46.036939	-89.628844	Yes
4	Alma Lake	58	18 FEET	45.909763	-89.428275	Yes

Notice that the maximum depth required some cleaning so that it was a numeric column.

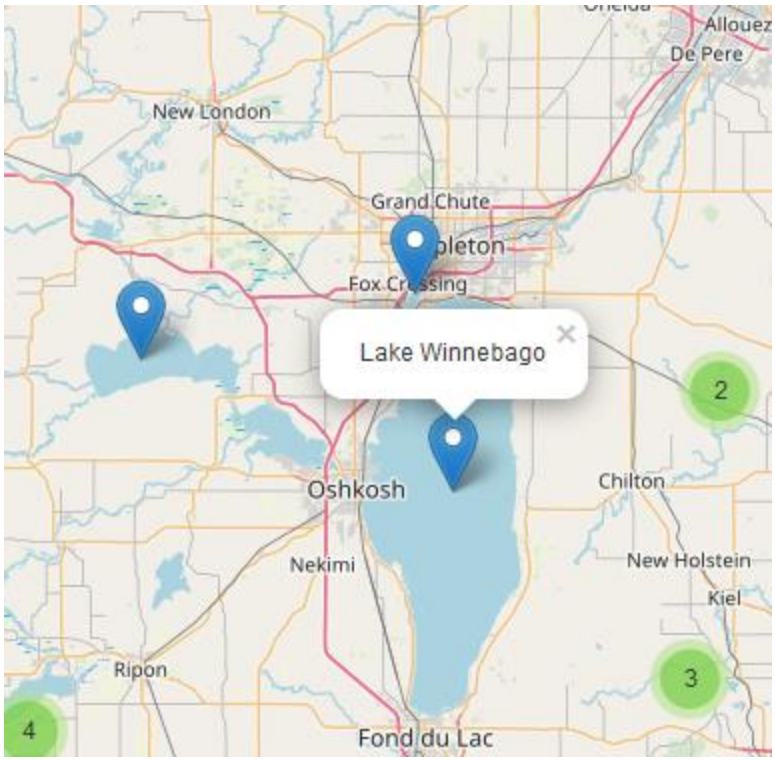
After converting that column into a new float column, the data looked as follows. Only the first few rows are shown.

	Lake Name	Size (Acres)	Latitude	Longitude	Public Landing	Max Depth
0	Alder Lake	264	46.088099	-89.819586	Yes	30.0
1	Aldridge Lake	142	45.865941	-89.301276	No	12.0
2	Alexander Lake	618	45.186975	-89.757128	Yes	36.0
3	Allequash Lake	406	46.036939	-89.628844	Yes	24.0
4	Alma Lake	58	45.909763	-89.428275	Yes	18.0

Using that dataframe, I produced a folium map of all the lakes that met the criteria for boat landings, size, and maximum depth. The folium map clustered the lakes into areas so that the map wasn't so busy.



Zooming into any of the areas produced exact locations of the lakes with a pop-up of the name of the lake as shown in the next figure.



Wisconsin State Parks

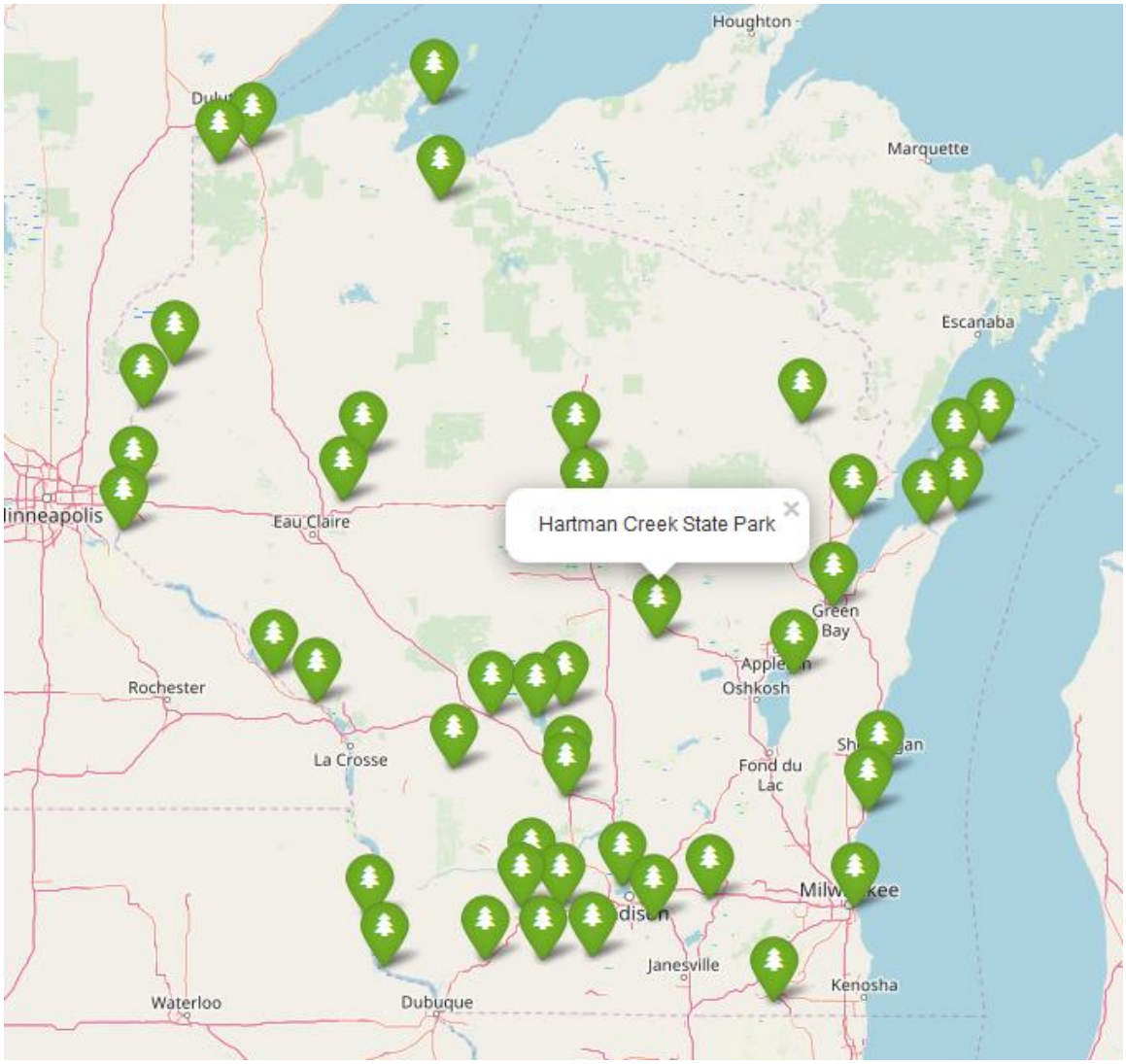
The next step was to do initial analysis of the Wisconsin State Park data. I scraped the names of all the Wisconsin state parks from wiki page using BeautifulSoup. There was no other information on the page that I was able to use.

After retrieving the names of the state parks from this page and excluding items that weren't state parks, I used geopy.geocoders Nominatim to get the latitude and longitude of the state parks.

I generated a dataframe with the park names, latitudes, and longitudes as follows. Only the first few rows are shown:

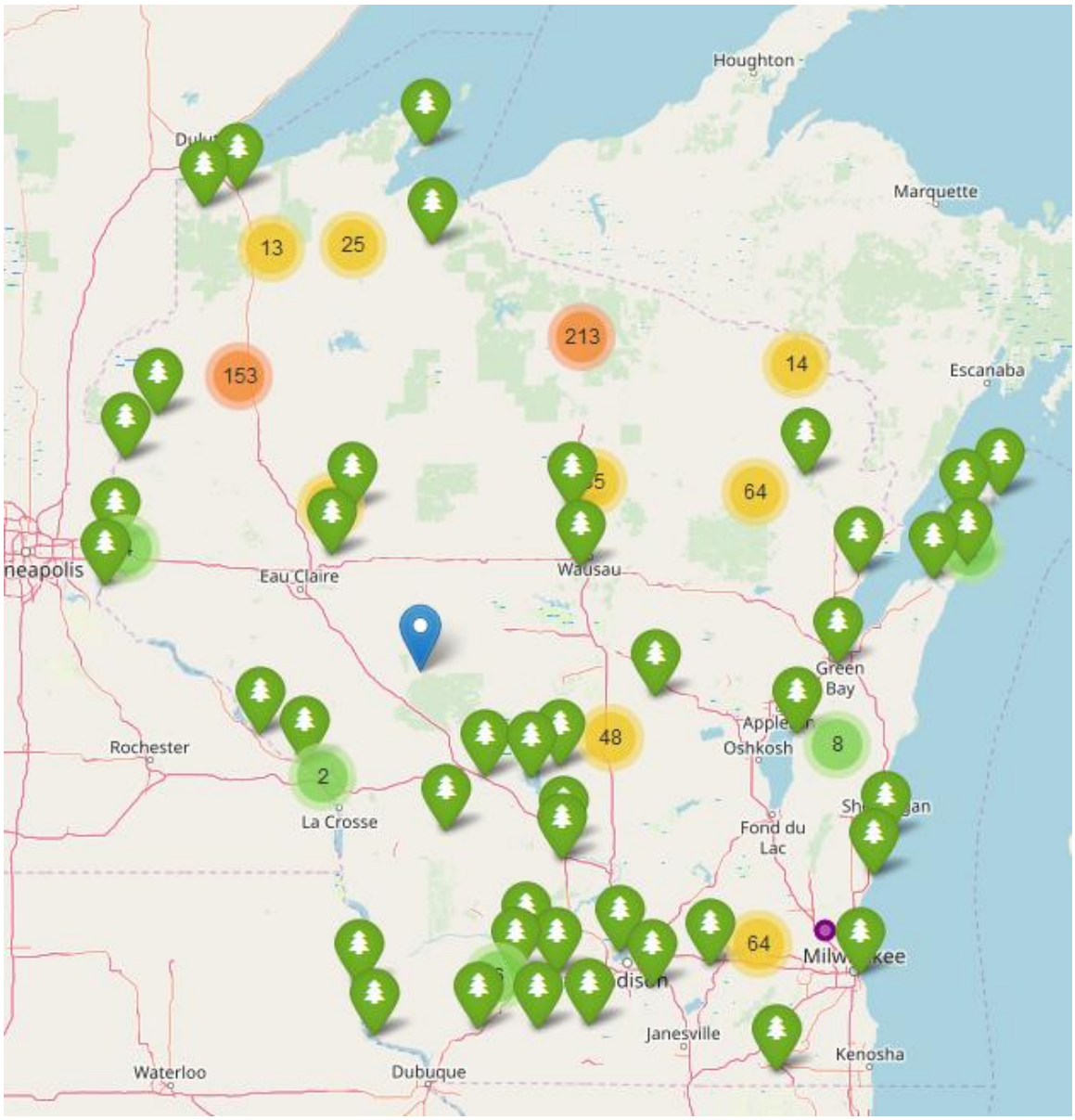
	Park Name	Latitude	Longitude
0	Amnicon Falls State Park	46.614722	-91.898624
1	Aztalan State Park	43.066505	-88.860419
2	Belmont Mound State Park	42.770246	-90.349897
3	Big Bay State Park	46.807026	-90.692332
4	Big Foot Beach State Park	42.569318	-88.426371
5	Blue Mound State Park	43.029746	-89.847331

Using the dataframe of state parks, I generated a folium map. I used a green tree as the icon to mark each state park's location and a pop-up with the name of the park.



Combining the Data

The next step in the analysis was to combine the maps to see where the state parks are in relation to the walleye fishing lakes. We also want to live within 2 hours of Germantown so I marked that city with a purple dot:



Visual Investigation of the Map

There were several surprises when looking at the combined map.

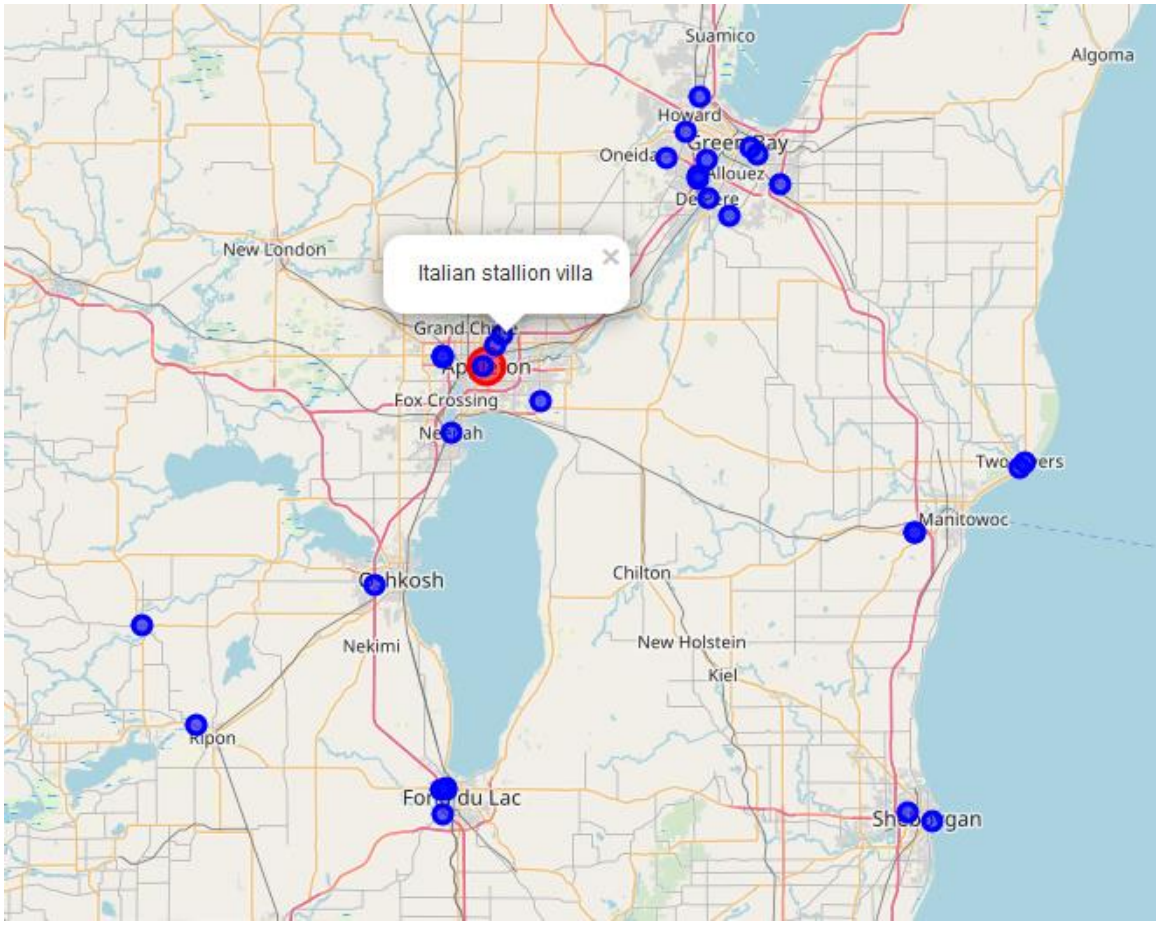
1. In the north central part of Wisconsin which is famous for walleye fishing, there aren't any state parks at all! We were very surprised by this. However, there are a lot of forests. As a future enhancement to this project, I need to look at the state and national forests to see if there is camping and hiking available there.
2. There are a lot of state parks in the central and south-central regions that we hadn't noticed before.

3. There aren't very many state parks along Lake Michigan. We would have thought that would be a prime location for them.
4. Waupaca might be a candidate. It's close to a lot of the fishing and it's only 97 minutes from Germantown and it's on a main highway.

Using the Foursquare API to get venues for chosen locations

First, I attempted to retrieve Italian restaurants for Waupaca, Wisconsin, and discovered there weren't any within a 10 mile radius. Then, I chose Appleton, Wisconsin, increased it to a 50 mile radius, and was able to create a dataframe consisting of 30 Italian restaurants.

From that dataframe, I created a map of the Italian restaurants and marked them with a blue dot. I marked Appleton with an apple-red larger dot. The pop-up for the restaurants was the name of the restaurant.



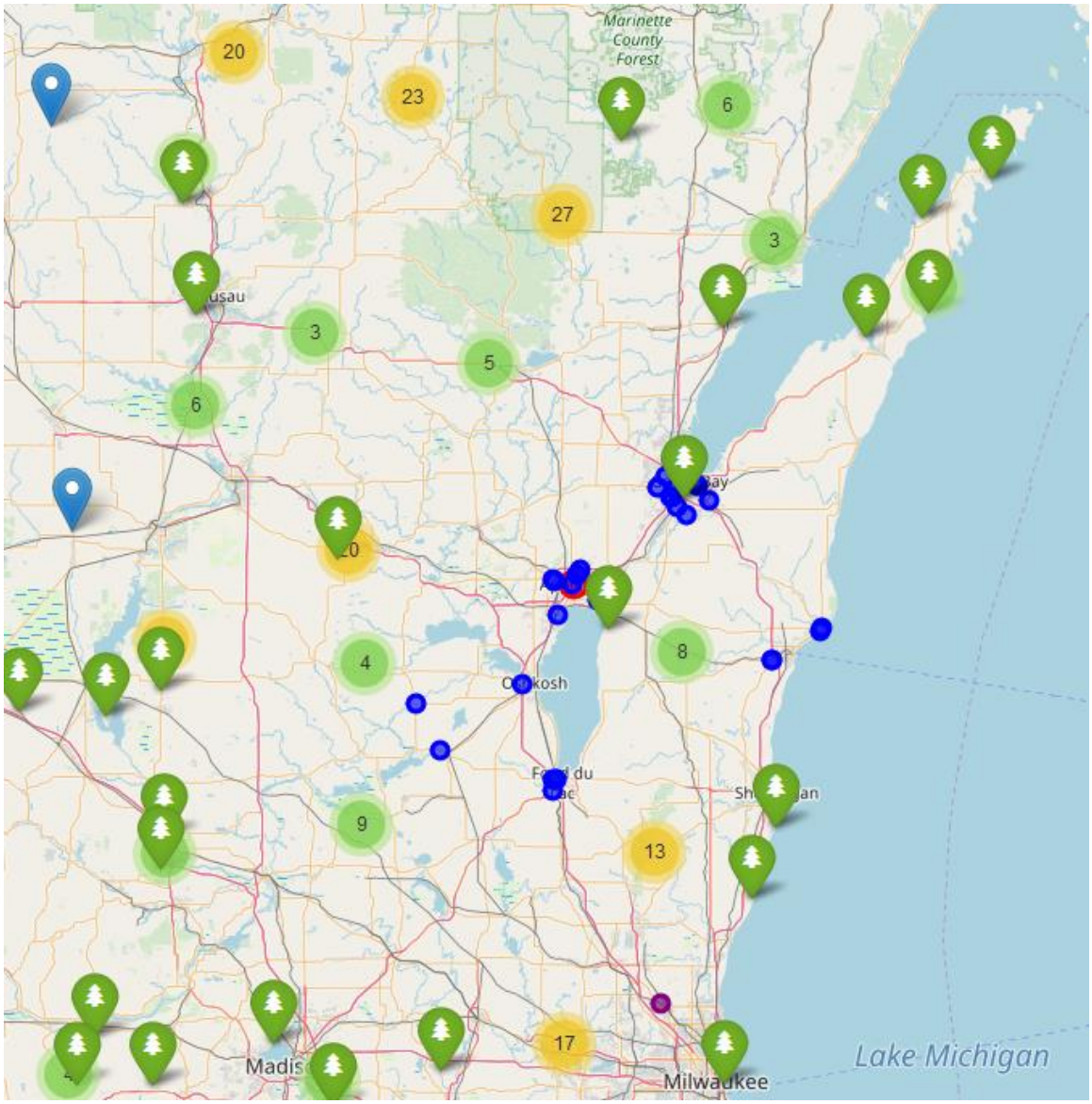
We noticed that there were many Italian restaurants in the Green Bay area. That city is farther from Germantown than we'd want to live and too close to the Green Bay Packers for my husband, but it would be somewhere we could go for a nice dinner even if we didn't live there.

I used Foursquare to further my analysis of a couple Italian restaurants to glean tips to see if there were any that mentioned if they served gluten-free entrees. I didn't find any, so I will need to explore those some more. These are the tips I received for Victoria's restaurant:

	text
0	Great service I came in 15 minutes before closing and the I felt very welcome. A lot of places try to turn you away. By portions are huge .
1	Great food at an affordable price! Weekday 2 for 1 specials in the bar! Be sure to take advantage of that and tip your server/bartender.

4. Results

The last map was a combined map of all the data I had retrieved: walleye lakes, state parks, Italian restaurants within 50 miles of Appleton, and markers for both Appleton and Germantown.



5. Discussion

Some of the observations we had are the following:

1. The walleye lake concentration in the north central area is not adjacent to state parks.
2. Smaller towns could be less likely to have Italian restaurants based on the results for Waupaca.
3. The southern and south-central areas of the state have more state parks than we expected.
4. There are many state trails and state forests that we might also enjoy that I didn't map.
5. We need to do more analysis including some of the following venues:
    - CrossFit gym locations

- Churches in our denomination
- Senior centers (for the future)
- State Forests
- State Trails

My recommendation is to add the items above to the map and then prioritize. We need to factor mileage and number of times per day we would be using these venues and facilities because there will be no perfect place that has everything we want.

6. Conclusions

This data combined into one map is truly helpful in finding locations that meet our needs. It really comes down to personal preference, though, when making a decision like this. The gut feelings and impressions we will have when visiting the area will be major factors in our decision.

We have time before our retirement and can take short trips into Wisconsin in the next few years to spend time in these areas.