Finding a 4 year home

Every year college students drop out or transfer because they are unhappy living in the area that the school they choose is in. May be it is to rural and there is not enough to do or maybe it is too busy and they are overwhelmed by everything.

- The goal of this project is to match a student to a school that has similar venue options as the city they live in.
- For this project I used Waukesha Wisconsin as the city the student lives in and all the 4-year schools in the state of Wisconsin as the choices for Universities.
- I picked the University of Milwaukee Waukesha as the central location for Waukesha and the address of all the 4-year schools as the center for their respective locations.
- I also decided that 3200 meters is a good distance for a student to travel without a car. This is the distance around each school I will use for the project.

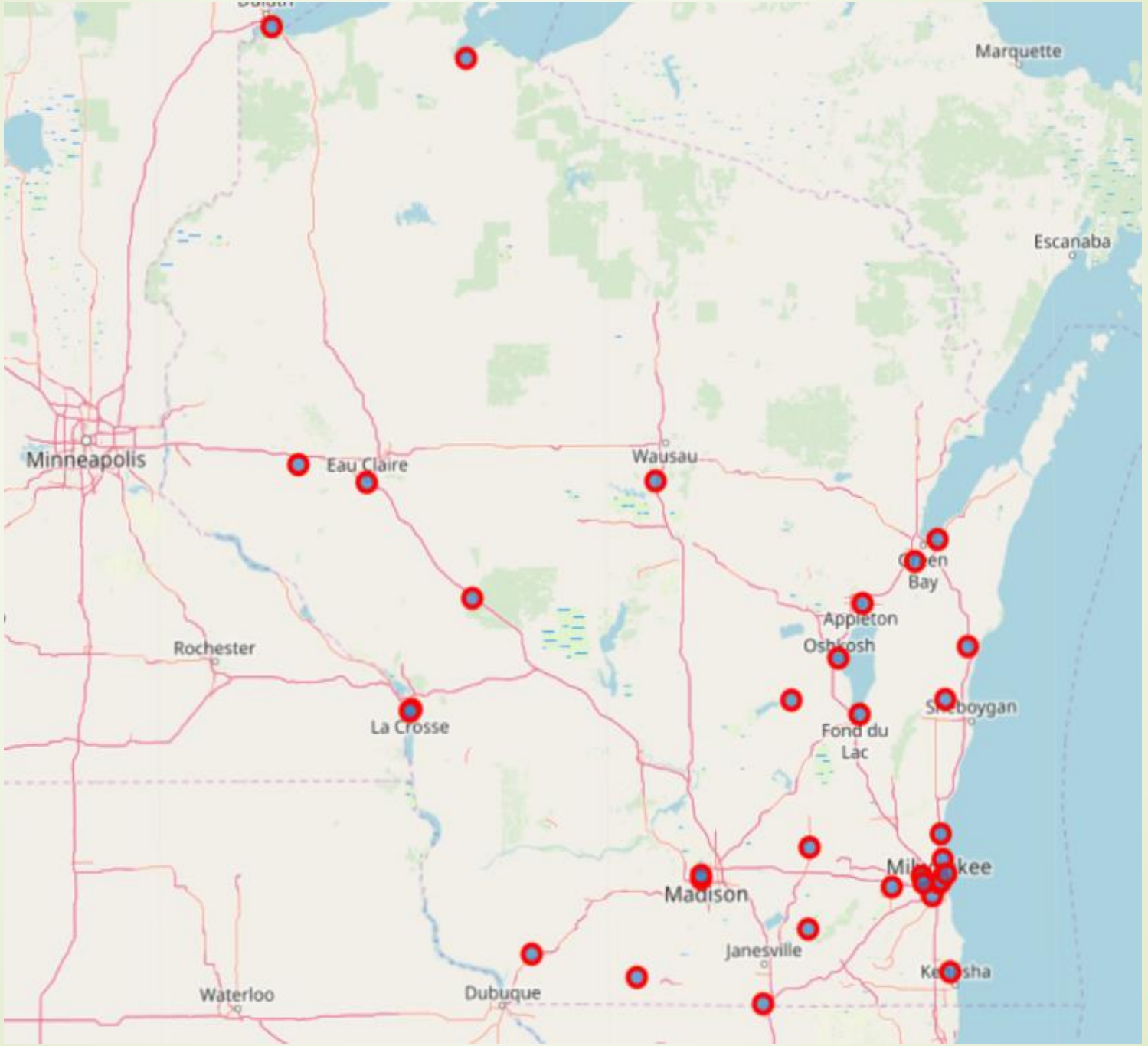
Using the following site I created a list of all universities in the state of Wisconsin.

- https://en.wikipedia.org/wiki/List_of_colleges_and_universities_in_Wisconsin
- Because I have a limited number of calls I can make from FourSquare I choose to only include 4-year institutions.

After creating the list I used google to find all the longitude and latitudes of the schools using there addresses. The following table is the first few rows of the CSV file.

School	Address	City	Type	Latitude	Longitude
Alverno College	3400 S 43rd st	Milwaukee	Master's university	42.983	-87.967
Beloit College	700 College St	Beloit	Baccalaureate college	42.503	-89.031
Cardinal Stritch University	6801 N Yates Rd	Milwaukee	Doctoral/research university	43.142	-87.906
Carroll University	100 N East Ave	Waukesha	Master's university	43.024	-88.221
Carthage College	2001 Alford Pard Dr	Kenosha	Baccalaureate college	42.622	-89.822

Here is a map of all the schools we are using in the project



Making the call to Foursquare and K-means clustering

Using Foursquare we can generate a list of all venues with-in 3200 meters of each school and 3200 meters of Waukesha.

Once we have the list we can use k-means clustering to see which schools are similar to Waukesha.

I did two runs of the test

- Run 1 I used all venues to do clustering
- Run 2 I used the following list to run clustering on a limited number of venues types.

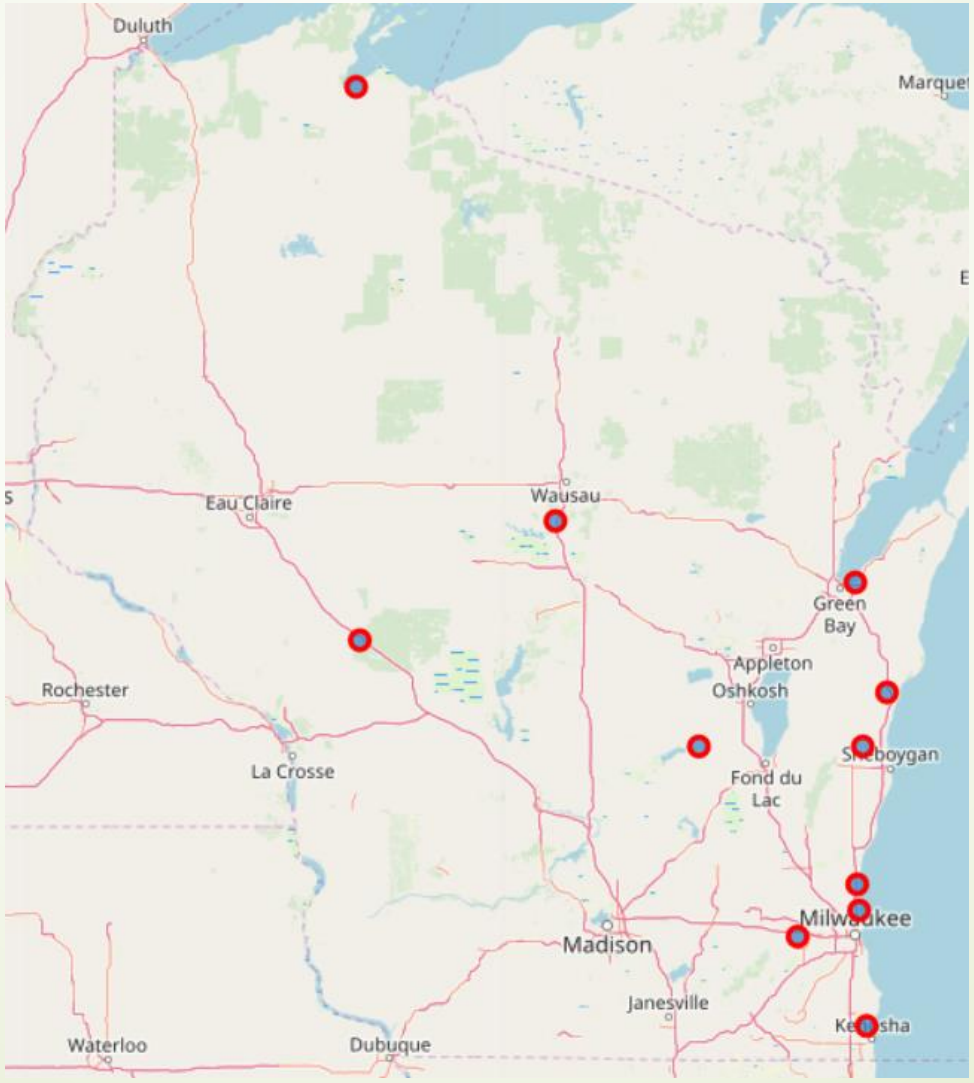
Sports Bar, Coffee Shop, Diner, Gym, Gym / Fitness Center, Mexican Restaurant, Pizza Place, Bank, Golf Course, and Basketball Court.

Run 1 All Venue Categories

Running k-means clustering with all venue categories clustered Waukesha with the following schools.

- Cardinal Stritch University
- Concordia University
- Lakeland University
- Northland College
- Ripon College
- Silver Lake College
- University of Wisconsin-Green Bay
- University of Wisconsin-Parkside
- University of Wisconsin-River Falls
- University of Wisconsin- Stevens Point

Here is a map of the schools that share a cluster with Waukesha

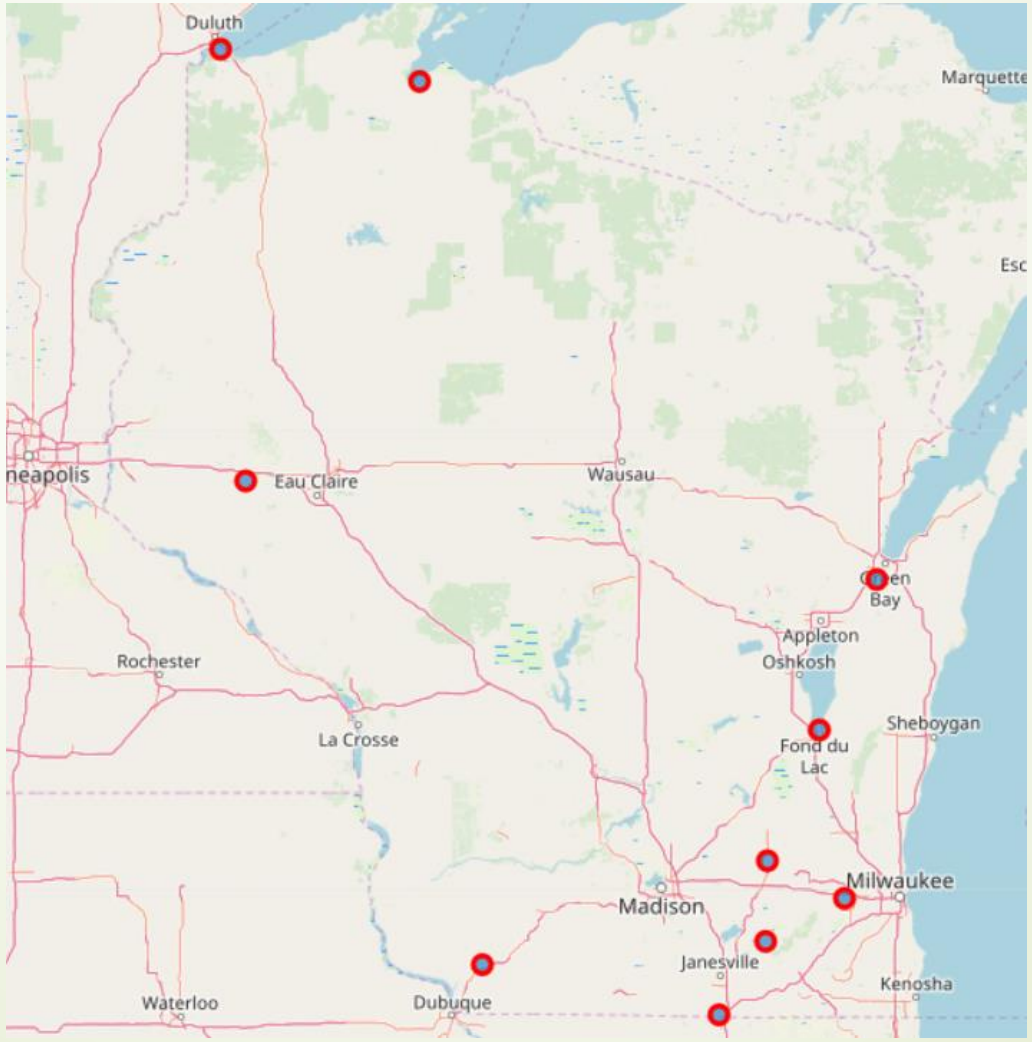


Run 2 Clustering using limited Venue Categories

Running k-means clustering with a limited set of venue categories clustered Waukesha with the following schools.

- Beloit College
- Maranatha Baptist University
- Marian University
- Northland College
- St. Norbert College
- University of Wisconsin-Platteville
- University of Wisconsin-Stout
- University of Wisconsin-Superior
- University of Wisconsin-Whitewater

Here is a map of Schools cluster with Waukesha when the venue categories are limited



Comparing the two list of schools

We can see that the list do not share any schools in common

Schools when cluster with all Venues

- Cardinal Stritch University
- Concordia University
- Lakeland University
- Northland College
- Ripon College
- Silver Lake College
- University of Wisconsin-Green Bay
- University of Wisconsin-Parkside
- University of Wisconsin-River Falls
- University of Wisconsin- Stevens Point

Schools when clustered with limited venues

- Beloit College
- Maranatha Baptist University
- Marian University
- Northland College
- St. Norbert College
- University of Wisconsin-Platteville
- University of Wisconsin-Stout
- University of Wisconsin-Superior
- University of Wisconsin-Whitewater

Which method gives a better result?

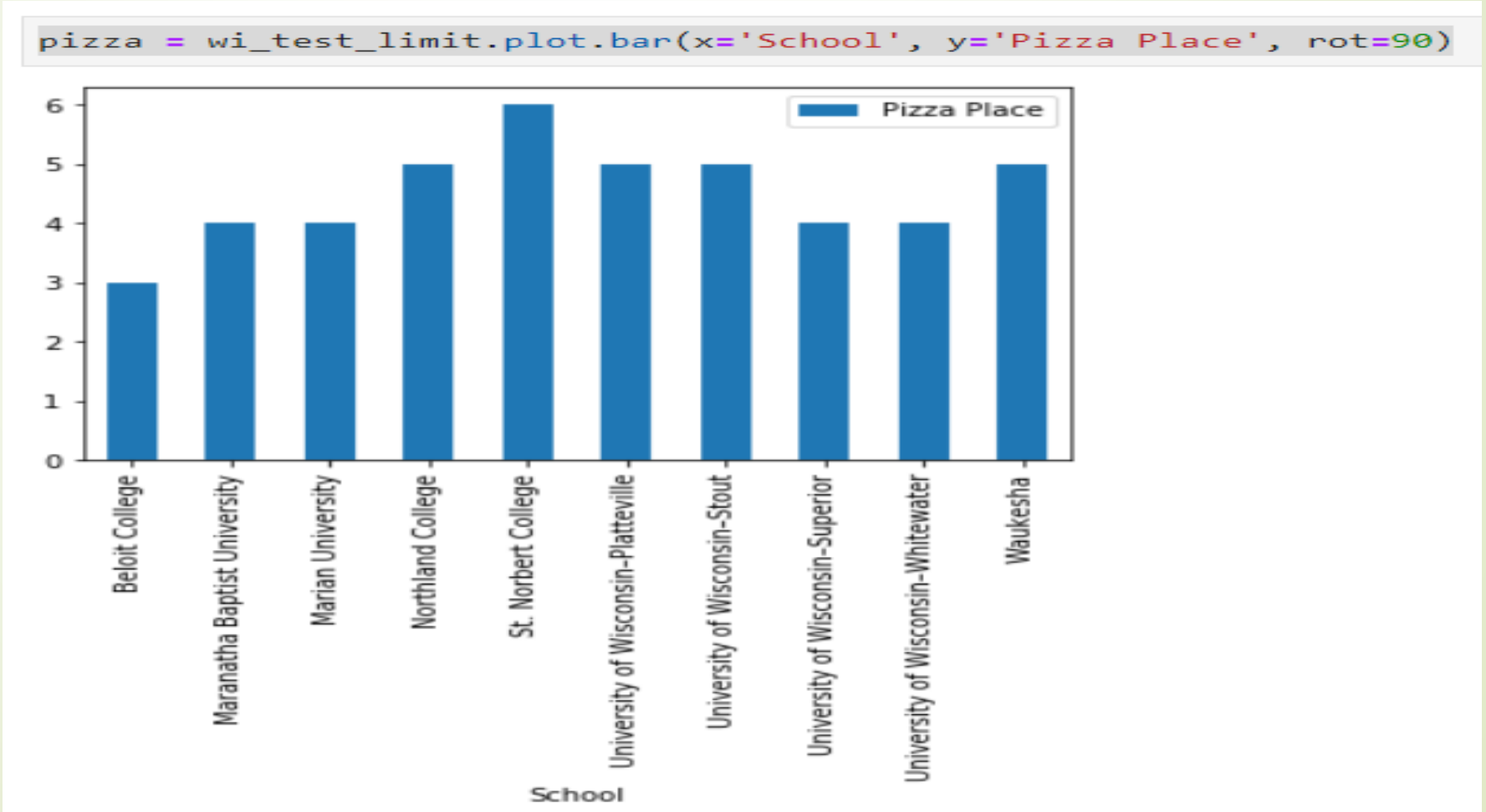
Lets look at some of the venues and compare how similar they are to Waukesha.

We will look pizza places

Ideally we should probably compare all of the graphs.

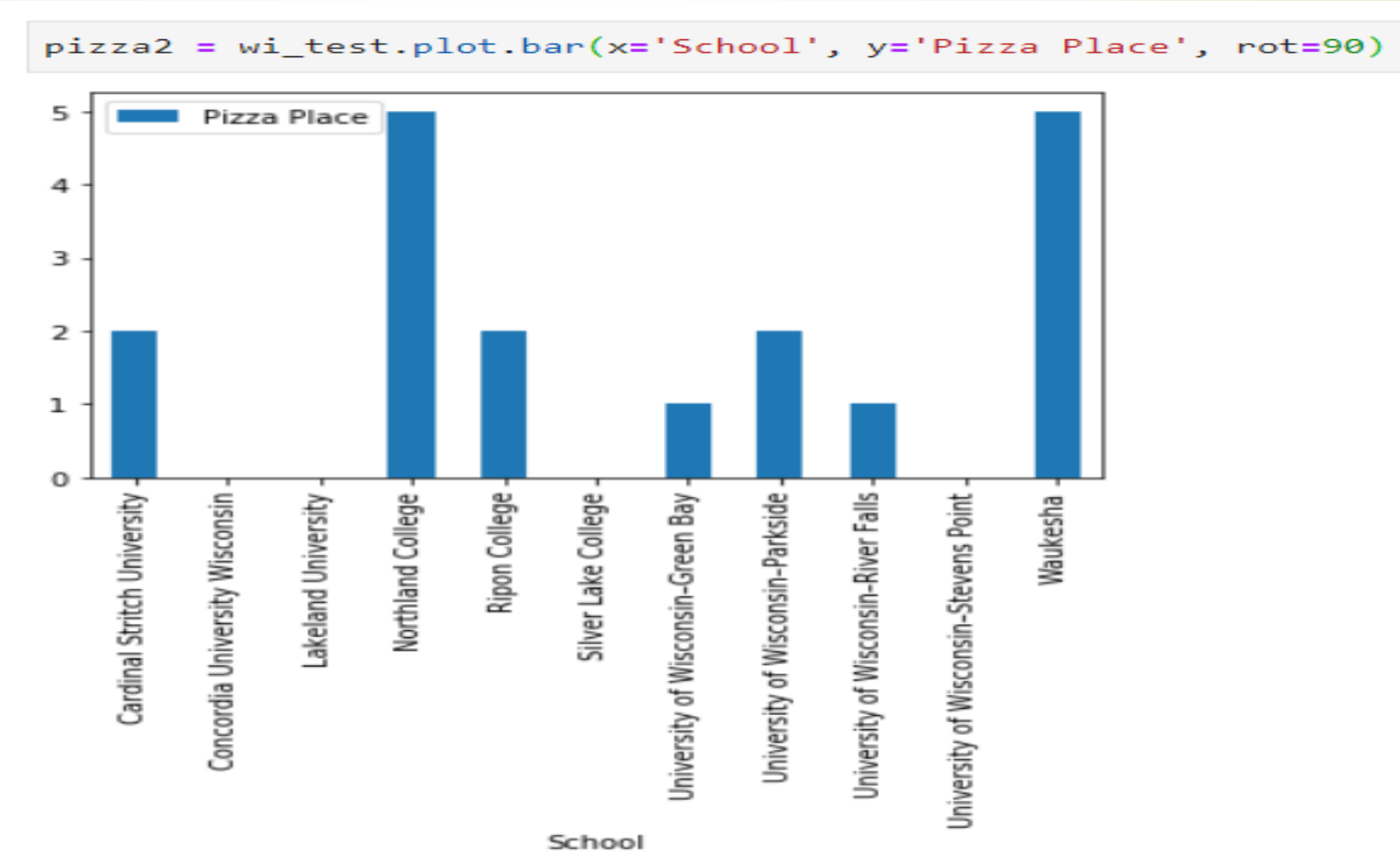
Bar Graph Pizza Places when the categories are limited

```
pizza = wi_test_limit.plot.bar(x='School', y='Pizza Place', rot=90)
```



Bar Graph when categories are not limited.

```
pizza2 = wi_test.plot.bar(x='School', y='Pizza Place', rot=90)
```



- We can see that the clustering method that did not limit the categories clustered Waukesha with schools that do not have any pizza places with-in 3200 meters. This would be a problem if you really like pizza.
- The clustering method that used limited categories seems to be a better match then the other method. There are at least 3 pizza places near each school which is close to the number at Waukesha.

Information that can be returned to Client

- Now that we have located some schools with similar area to Waukesha we can give the client information on the venues located around each school.
- In the following slides I will show examples of the following.
  - Top of the list for all venues around Beloit College
  - List of all Pizza Places around Beloit College
  - List of all venues from the limited list of venues used for both Beloit College and Maranatha Baptist Bible College
- Ideally these list would be populated for all the schools and given to the client to help them make a decision on where to live.

Top of list for all venues around Beloit College

ool	School Latitude	School Longitude	Venue	Venue Latitude	Venue
ege	42.503	-89.031	Merrill & Houston's Steak Joint	42.501016	
ege	42.503	-89.031	615 Club	42.498880	
ege	42.503	-89.031	Bushel & Peck's	42.499538	
ege	42.503	-89.031	Ironworks Hotel	42.501129	
ege	42.503	-89.031	Tilley's Pizza House & Ballyhoo Tavern	42.507632	
ege	42.503	-89.031	Beloit Farmer's Market	42.499642	
ege	42.503	-89.031	Zen Sushi & Grill	42.499553	
ege	42.503	-89.031	Everett's Wines, Spirits And Beer	42.492459	
ege	42.503	-89.031	Barnes & Noble	42.499620	
ege	42.503	-89.031	Jerry's Cafe	42.508382	
ege	42.503	-89.031	Club Impulse	42.501302	
ege	42.503	-89.031	Neli's Restaurant	42.492549	
ege	42.503	-89.031	Suds O'Hanahans	42.499688	
ege	42.503	-89.031	Riverside Park	42.511158	
ege	42.503	-89.031	Turtle Tap	42.497630	
ege	42.503	-89.031	M & M Dari Ripple	42.490882	
ege	42.503	-89.031	Grand Avenue Pub	42.501250	
ege	42.503	-89.031	La Casa Grande	42.502361	
ege	42.503	-89.031	Walgreens	42.498332	
ege	42.503	-89.031	Bagels & More	42.499480	

Venues from limited list for Beloit College

School	Venue	Venue Categor
Beloit College	Tilley's Pizza House & Ballyhoo Tavern	Pizza Plac
Beloit College	La Casa Grande	Mexican Restaurar
Beloit College	Bagels & More	Coffee Sho
Beloit College	Planet Fitness	Gym / Fitness Cente
Beloit College	Pizza Hut	Pizza Plac
Beloit College	Stateline Family YMCA	Gyr
Beloit College	Taco John's	Mexican Restaurar
Beloit College	Papa John's Pizza	Pizza Plac
Beloit College	Taqueria Azteca	Mexican Restaurar
Beloit College	Denny's	Dine

	School	School Latitude	School Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
104	Beloit College	42.503	-89.031	Tilley's Pizza House & Ballyhoo Tavern	42.507632	-89.038463	Pizza Place
137	Beloit College	42.503	-89.031	Pizza Hut	42.505454	-89.038283	Pizza Place
146	Beloit College	42.503	-89.031	Papa John's Pizza	42.493209	-89.037700	Pizza Place

Pizza places Beloit College

Limited Venues Maranatha Baptist College

	School	School Latitude	School Longitude	
565	Maranatha Baptist University	43.194	-88.739	
567	Maranatha Baptist University	43.194	-88.739	
568	Maranatha Baptist University	43.194	-88.739	Tribeca
569	Maranatha Baptist University	43.194	-88.739	
573	Maranatha Baptist University	43.194	-88.739	
574	Maranatha Baptist University	43.194	-88.739	
585	Maranatha Baptist University	43.194	-88.739	
590	Maranatha Baptist University	43.194	-88.739	
591	Maranatha Baptist University	43.194	-88.739	W
592	Maranatha Baptist University	43.194	-88.739	
594	Maranatha Baptist University	43.194	-88.739	
597	Maranatha Baptist University	43.194	-88.739	
603	Maranatha Baptist University	43.194	-88.739	
615	Maranatha Baptist University	43.194	-88.739	

Conclusion

The method which limits the venue categories seams to work better then the method that uses all categories. Most people would want to be matched on venues they visit so if they could choose the category it would give them better results.

Going forward

I should try other clustering methods but k-means is the only one I have has time to practice.

I also limited the number of venues and the distance from each venue so I did not use up all my calls to foursquare. If I paid for an account I would increase both these numbers.

It would also be nice if people could rank things they enjoy instead of matching them to a location but that programming skills is beyond my current skill level.