

# Battle of the neighborhoods

---

## Introduction

Let's say you want to start a little pizza place in Stockholm, Sweden. But you don't have the resources to compete with the major fast food companies and you also want as little competition as possible in your local area.

Where should you go, and look for a place to rent?

In this assignment, I will try to find the best place in Stockholm to start a little pizza place where there is little competition but as close as the city center as possible with the assumption that the population density is higher the closer to the city center we get.

To conclude, a pizza place near many possible customers with little competition.

**"Pizza makes anything possible"**

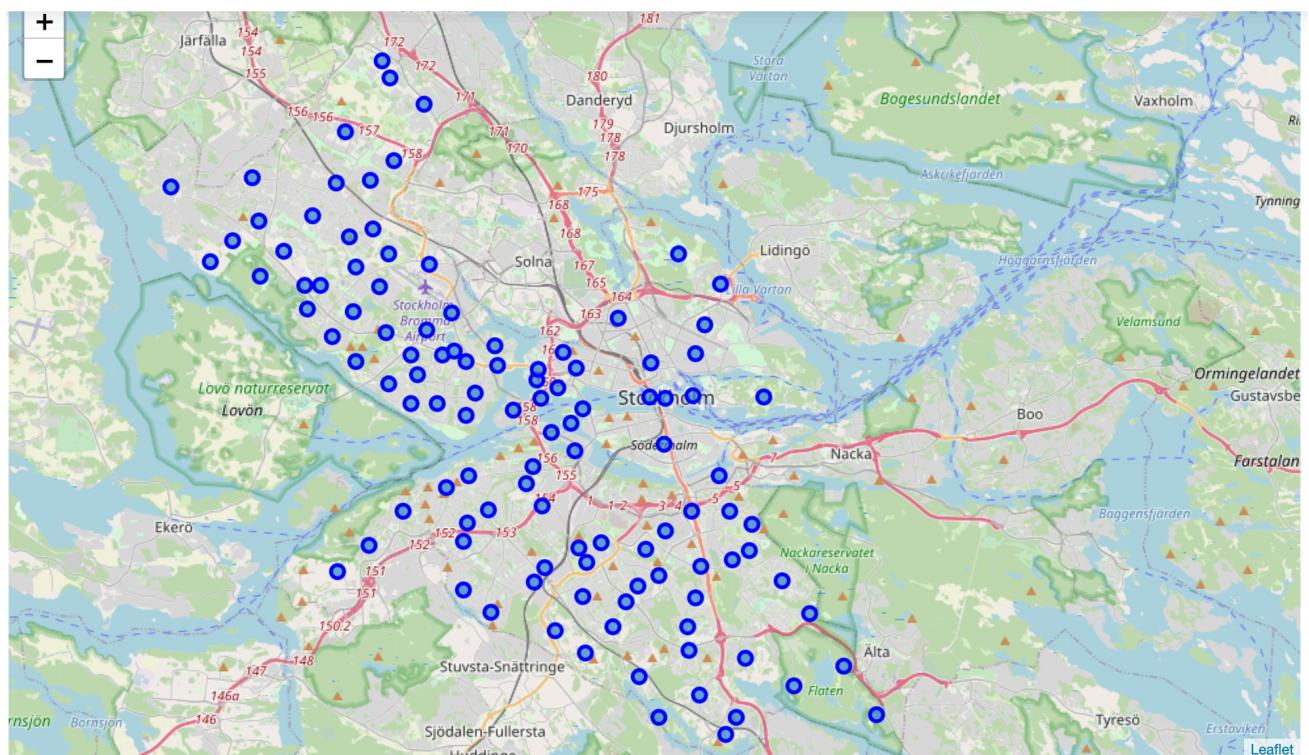
## Data

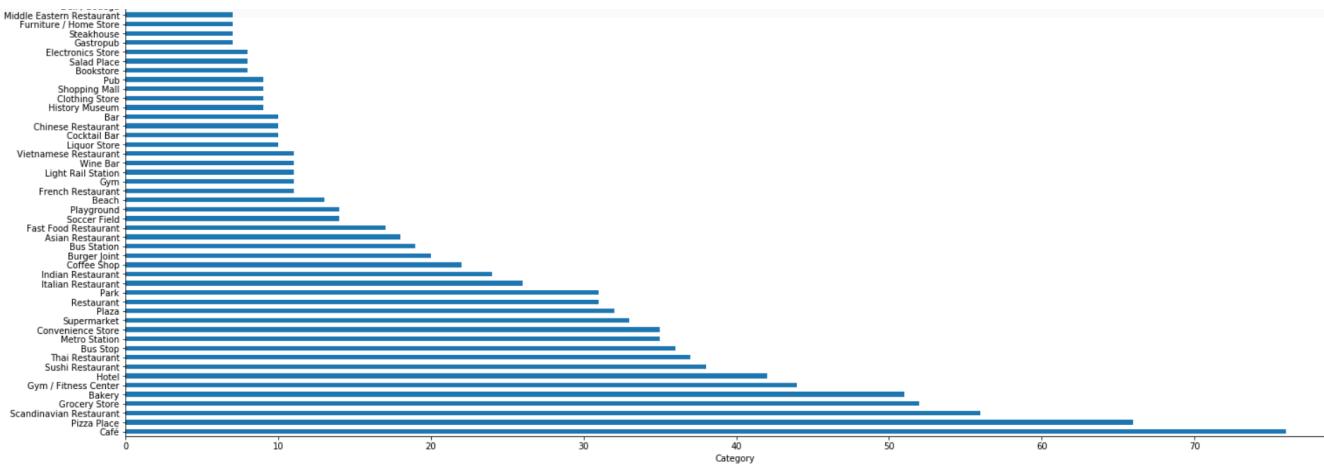
To solve this problem, I will use venues data from Foursquare to get venues. There will also be necessary to get all the neighborhoods in Stockholm, and to do this we scrape all the neighborhoods from Stockholm Wikipedia page. To be able to plot the neighborhoods on a map we also need the coordinates of the neighborhoods which we extract with geopy python library.

- Venues from foursquare API
- Neighborhoods from Wikipedia and coordinates

## Methodology

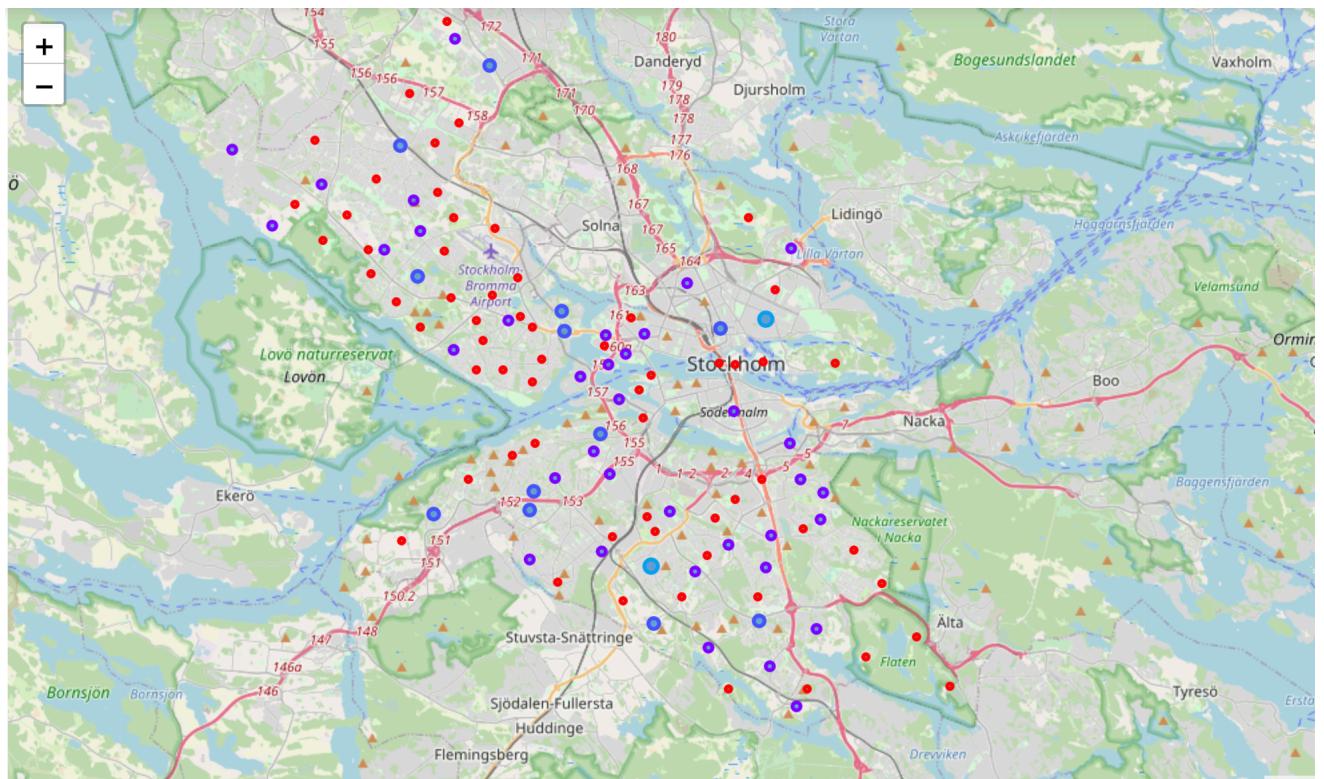
After collecting the neighborhoods, we plot all them on a Stockholm map to see which neighborhoods we are working with. The data tells us that there are 117 neighborhoods in Stockholm:





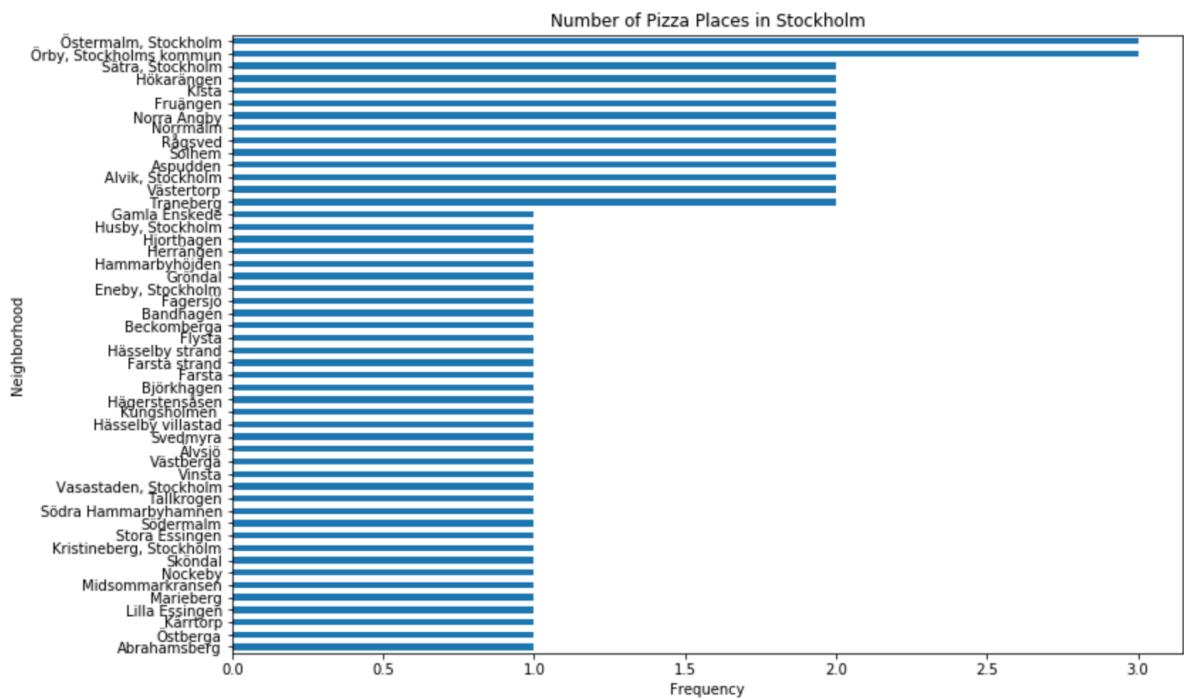
From this plot, we can see that Pizza Places are the second most common venue in Stockholm with around 65 venues. Therefore, it's very important that we find the right location for our new place where the competition is not that big.

Now lets plot the pizza places on a map of Stockholm:

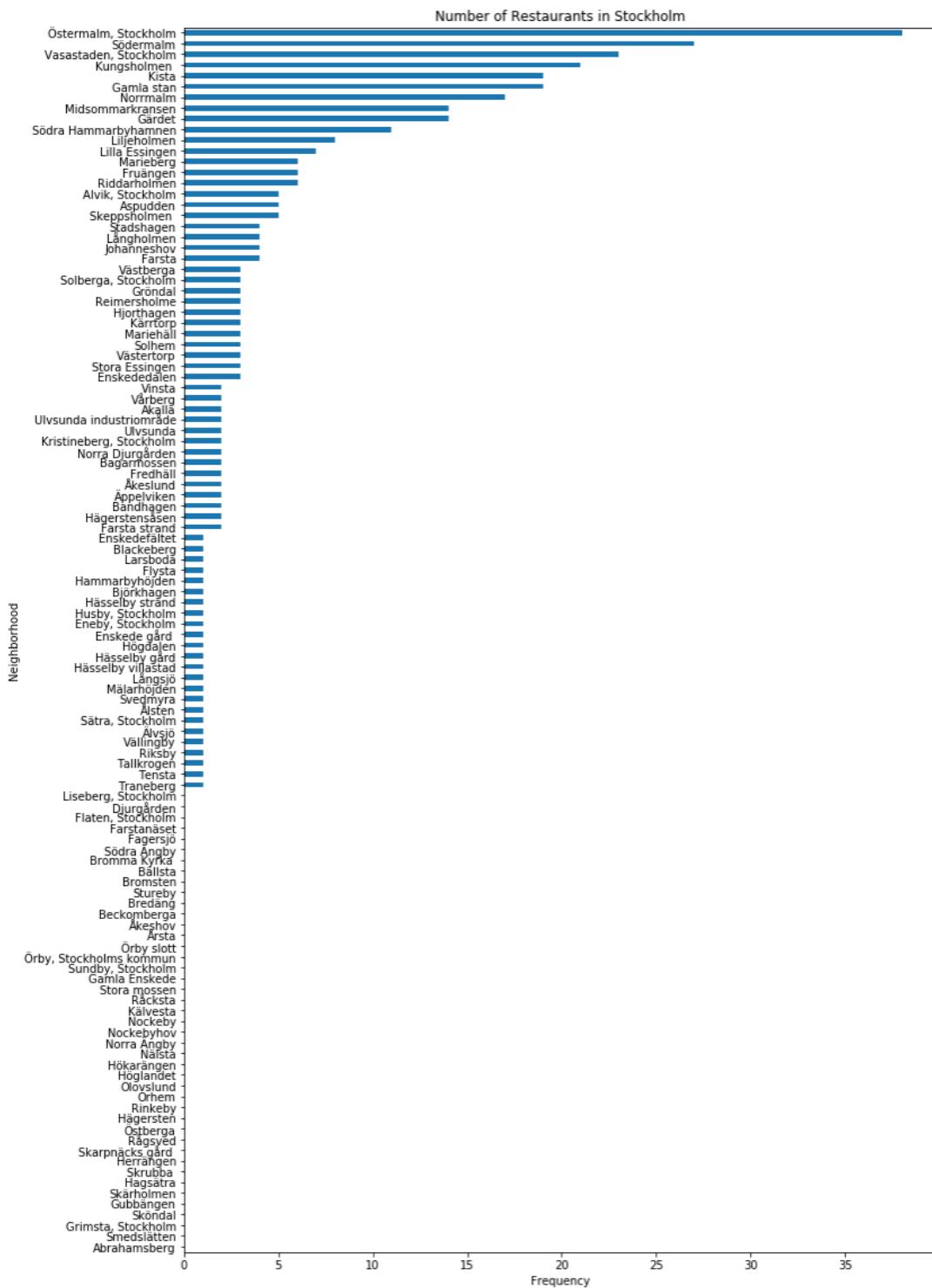


The bigger the dots, the more Pizza Places. This tells us that there are two neighborhoods with 3 Pizza Places, the light blue markers, and the red markers are neighborhoods with 0 Pizza Places.

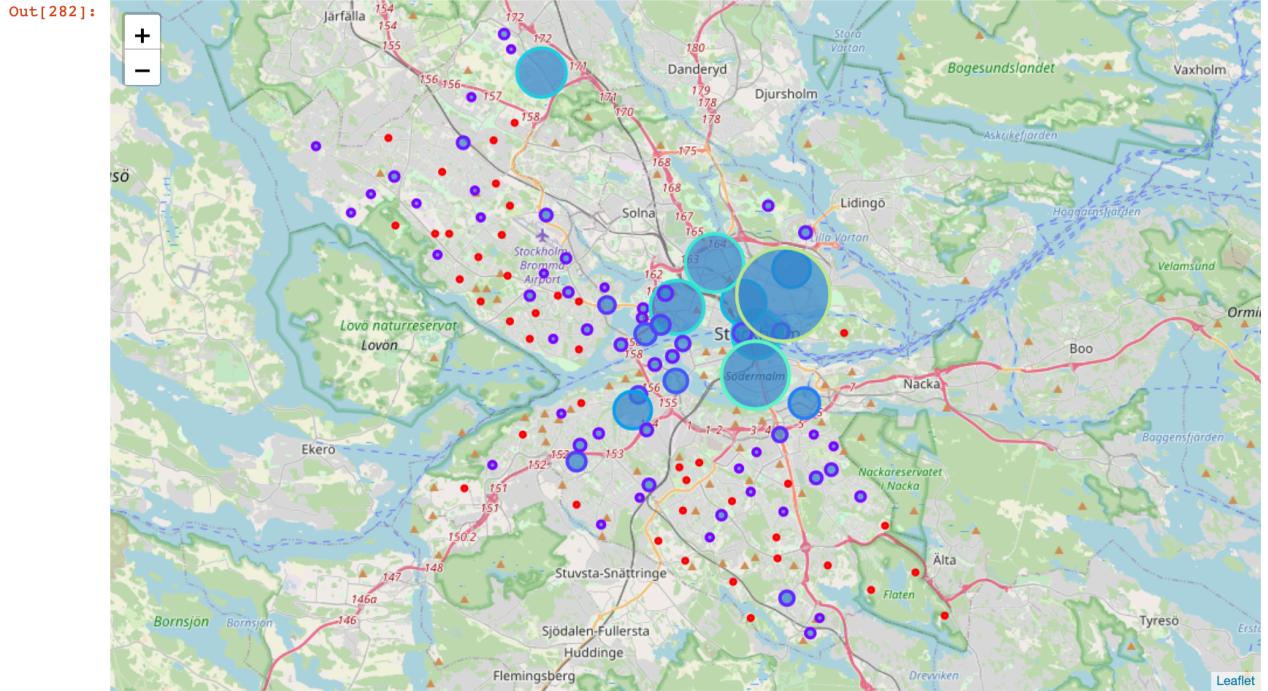
This is also shown in the bar chart below.



There is not only competition from other Pizza Places. Also, nearby restaurants and food joints are competition for customers. First we take a look at how many restaurants there is in each neighborhood. To do so we select all the Venue Categories that includes the word Restaurant.



We also plot these on a map for a better look.



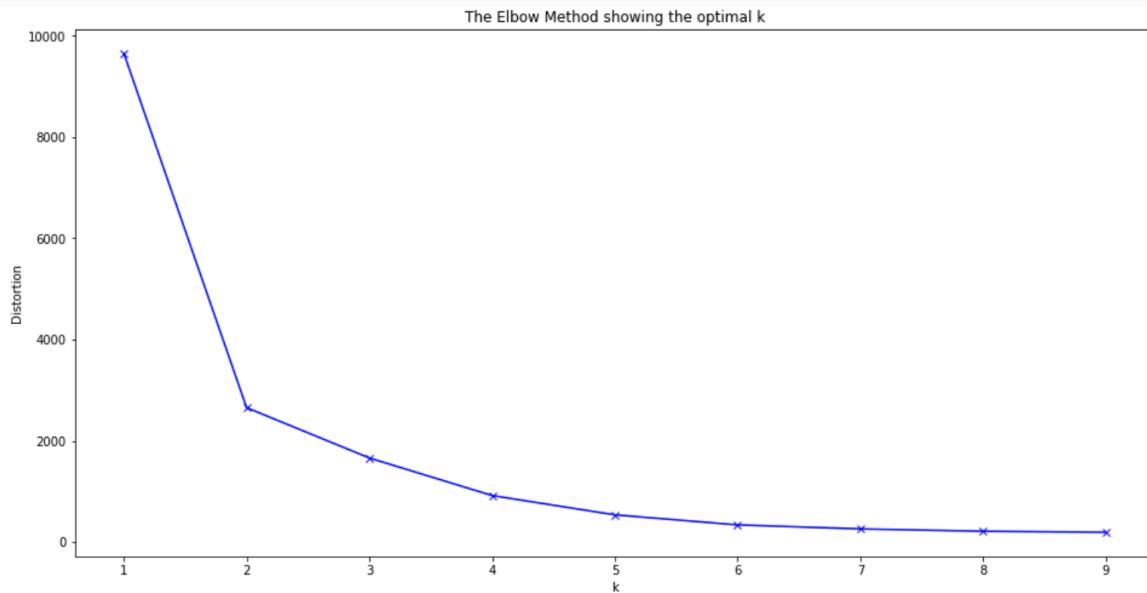
Then we do the same procedure for the venue categories that includes the word food. We can see that two of the categories that includes the word food also have the word Restaurants and has already been counted in the previous section. We therefore remove those two categories from the food plot.

Then we create our last table consisting of the columns neighborhood, food places, restaurants, pizza places and total number of venues.

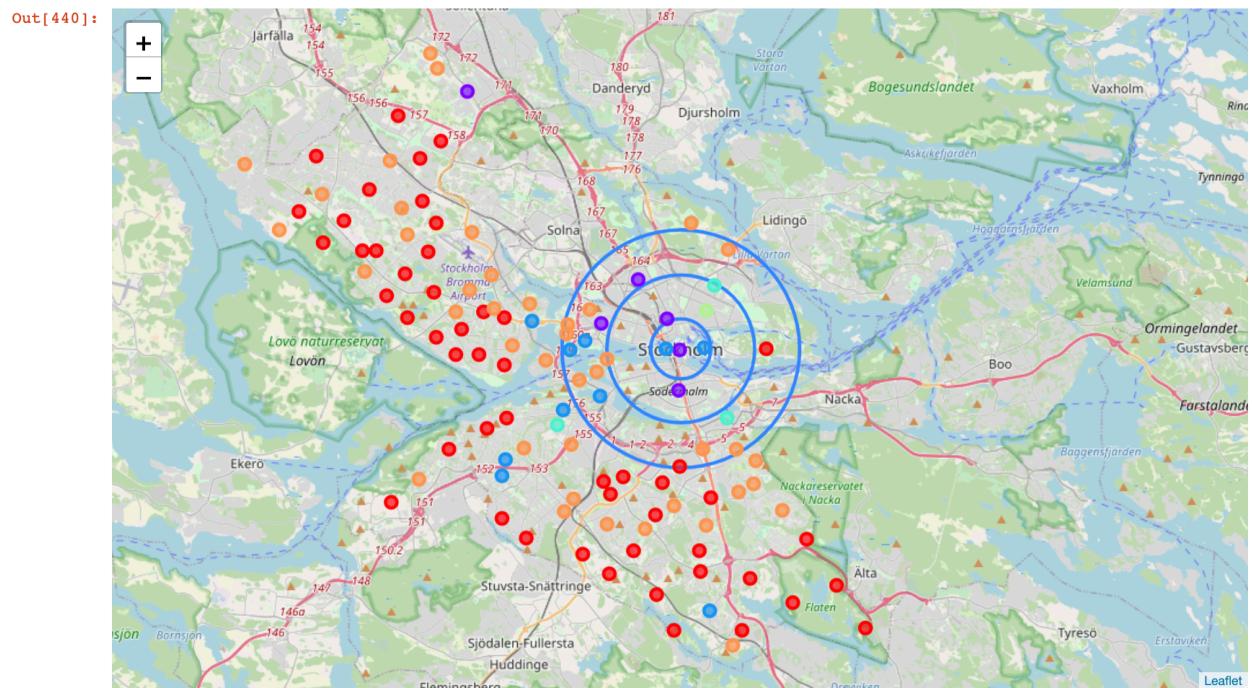
	<b>neighborhood</b>	<b>latitude</b>	<b>longitude</b>	<b>Pizza Place</b>	<b>Restaurants Total</b>	<b>Food Total</b>	<b>Totals</b>
<b>0</b>	Abrahamsberg	59.336468	17.953763	1	0	0	1
<b>1</b>	Akalla	59.414700	17.921818	0	2	1	3
<b>2</b>	Alvik, Stockholm	59.333401	17.982279	2	5	0	7
<b>3</b>	Aspudden	59.306466	18.001373	2	5	0	7
<b>4</b>	Bagarmossen	59.276186	18.131340	0	2	0	2
...	...	...	...	...	...	...	...
<b>109</b>	Äppelviken	59.326225	17.971132	0	2	0	2
<b>110</b>	Örby slott	59.280940	18.029227	0	0	0	0
<b>111</b>	Örby, Stockholms kommun	59.271847	18.027119	3	0	1	4
<b>112</b>	Östberga	59.286210	18.036771	1	0	0	1
<b>113</b>	Östermalm, Stockholm	59.336723	18.085918	3	38	1	42

After this we perform a k-means clustering on our data to cluster our neighborhoods. We use the elbow method to find the optimal number of clusters which in this case seems to be 2. After testing

with 2 clusters we decide to discard this and use 6 clusters.



Finally, we plot our clusters on a map of Stockholm with circles drawn from the city center.



## Results

By looking at this we would choose to look further into the red and orange markers close to the city centre.

Långholmen and Djurgården is our best options based on this analysis. Since I live in Stockholm i know that there isn't many people living on Djurgården or Långholmen and I would rather choose another location based on the map above.

Stadshagen or Johanneshov would be great alternatives with lots of people living nearby but not many restaurants or pizza places. Therefore the next step in this analysis is to add the population density or maybe where people are moving.

---

## Discussion

Solving this problem, we make a bunch of assumptions. Is it true that the population density always is higher closer to the city center? We also use more clusters than necessary to get a good result. It would be interesting to plot on the same map the population density for each neighborhood.

It would also have to be taken into account that not every pizza place or restaurant is available in the foursquare venues API, why more analysis of recommended areas is necessary before making a final decision.

Based on what I know of Stockholm neighborhoods I would say that the results are reasonable.

---

## Conclusion

Based on the results in this report, if opening a Pizza Place in Stockholm it should be opened in Stadshagen or Johanneshov. It can also be said that there are plenty of Pizza Places in Stockholm, it is the second most common venue category in Stockholm from foursquare API and the city center in Stockholm is very crowded with restaurants and food joints.