Capstone Project Report
Opening a Bakery in Stockholm

Introduction

Who doesn't love baked goods? In Stockholm, there are many bakeries and an even bigger interest in baking. It could be assumed that many people in Stockholm are considering opening a bakery of their own. However, it is difficult to know in which district there is a lack of bakeries and could use an additional one.

Business Problem

The goal with this project is to analyze and suggest which locations in Stockholm would be the most suitable for a new bakery. Since the location appears to be the most important competitive advantage, knowing where to locate a bakery would be the first step towards success.

Target Audience

The target audience for this project consists of baking entrepreneurs and bakery chains seeking to open a new venue in an optimal location.

Data

In order to make the analysis needed we will need two sources of data:

1. Wikipedia to get a list of all the Districts in Stockholm
    a. https://en.wikipedia.org/wiki/Stockholm_City_Centre
2. Foursquare to get information about Geolocation of Districts
3. Foursquare to get information about venues related to baking

Through analyzing the number of bakeries in each district we will be able to create a number of clusters, which will be the basis for the selection of location a new bakery.

Methodology

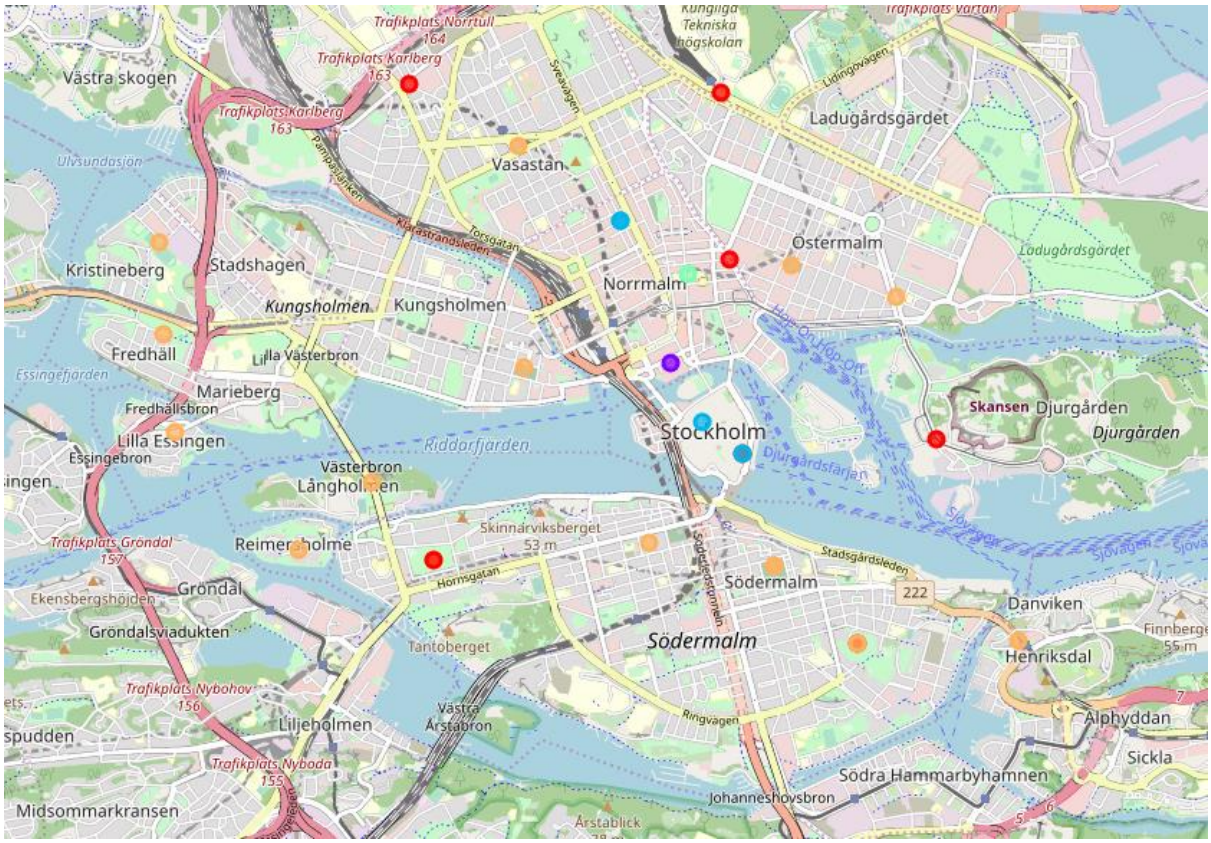
We'll use Wikipedia as a source of information about the district. We'll use Panda's read_html to scrape from the wikipedia-page about Stockholm.

Furthermore we'll use Foursquare API to collect information about bakeries in stockholm within a 500m radius from each district center.

From there we'll start our analysis utilizing K-means clustering.

Results & Discussion

Below is a map of the various clusters:



Cluster 0

	District	Cluster Labels	Bakery Freq	Population	Borough	Latitude	Longitude	Bakeries Per Capita
0	Djurgården	0	1	788	Östermalm	59.324620	18.097800	0.001269
12	Norra Johannes	0	12	9043	Norrmalm	59.323700	18.074492	0.001327
14	Norra Sofia	0	7	7721	Katarina-Sofia	59.312150	18.088350	0.000907
18	Stureplan-Lärkstaden	0	10	8104	Östermalm	59.335578	18.073046	0.001234
21	Södra Högalid	0	4	4155	Katarina-Sofia	59.317260	18.037730	0.000963
24	Södra Station	0	4	4844	Katarina-Sofia	59.346270	18.034670	0.000826
25	Tekniska Högskolan	0	3	3442	Östermalm	59.345700	18.071940	0.000872

Cluster 1

	District	Cluster Labels	Bakery Freq	Population	Borough	Latitude	Longitude	Bakeries Per Capita
4	Jakob	1	17	201	Norrmalm	59.329274	18.066006	0.084577

Cluster 2

	District	Cluster Labels	Bakery Freq	Population	Borough	Latitude	Longitude	Bakeries Per Capita
11	Norra Adolf Fredrik	2	16	3816	Norrmalm	59.33789	18.060060	0.004193
17	Storkyrkan	2	13	3017	Maria-Gamla stan	59.32563	18.069820	0.004309
19	Södra Adolf Fredrik	2	16	3703	Norrmalm	59.33789	18.060060	0.004321
22	Södra Johannes	2	12	2011	Norrmalm	59.32370	18.074492	0.005967

Cluster 3

	District	Cluster Labels	Bakery Freq	Population	Borough	Latitude	Longitude	Bakeries Per Capita
5	Klara	3	18	1597	Norrmalm	59.334667	18.068028	0.011271

Cluster 4

	District	Cluster Labels	Bakery Freq	Population	Borough	Latitude	Longitude	Bakeries Per Capita
1	Fredhäll	4	1	4958	Kungsholmen	59.331030	18.005450	0.000202
2	Gustav Vasa	4	7	12911	Norrmalm	59.342500	18.047750	0.000542
3	Hedvig Eleonora	4	6	10387	Östermalm	59.335240	18.080460	0.000578
6	Kristineberg	4	1	5572	Kungsholmen	59.336620	18.004860	0.000179
7	Kungsholm	4	6	18465	Kungsholmen	59.329010	18.048540	0.000325
8	Lilla Essingen	4	3	4519	Kungsholmen	59.325040	18.006770	0.000664
9	Mariatorget	4	7	14099	Maria-Gamla stan	59.318302	18.063466	0.000496
10	Mellersta Högalid	4	4	9914	Maria-Gamla stan	59.317260	18.037730	0.000403
13	Norra Högalid	4	4	13166	Maria-Gamla stan	59.317260	18.037730	0.000304
15	Oscars Kyrka	4	2	15271	Östermalm	59.333300	18.092900	0.000131
16	Reimersholme-Långholmen	4	1	2349	Maria-Gamla stan	59.317850	18.021540	0.000426
20	Södra Hammarbyhamnen	4	1	10615	Katarina-Sofia	59.312360	18.107480	0.000094
23	Södra Sofia	4	7	11015	Katarina-Sofia	59.312150	18.088350	0.000635
26	Västra Katarina	4	9	13220	Katarina-Sofia	59.316860	18.078300	0.000681
27	Västra Matteus	4	1	14272	Kungsholmen	59.321946	18.030200	0.000070
28	Östra Katarina	4	9	19855	Katarina-Sofia	59.316860	18.078300	0.000453

Our analysis shows that there are over 200 bakeries in central Stockholm, with five distinct clusters of density of bakeries per capita. It could of course be interesting to include other venues such as cafés and bistros, since these could be competing with eachother. Additional data to analyze would be average income in each district, but I could not find the data.

What we can see from our analysis is that there are several areas in Stockholm that could use some new bakeries - which can be found in cluster 4.

Conclusion

For anyone interested in opening a new bakery in Stockholm, any of the districts in cluster 4 would be recommended!