Exploring Stockholm, Sweden: The battle of neighborhoods!

Ekta Vats

November 7, 2020

Table of contents:

- Introduction: Business Problem
- Data
- Methodology
- Results and Discussion
- Conclusion

Stockholm is the capital and most populous city of Sweden, and the city stretches across 14 islands where Lake Mälaren flows into the Baltic Sea. Stockholm

Introduction: Business Problem

is also the cultural, media, political, and economic centre of Sweden. Stockholm is a vibrant and exciting place, with rich cultural diversity and friendly people all around. Want to explore the Stockholm city? Let's do some fun with data science!

In this project, we will explore the city of Stockholm, and find out the most popular places (such as museums, parks, café, restaurants, etc.), to help guide the tourists, investors and general public interested in visiting Stockholm. We would also prefer locations that are close to the city centre.

We will use the data science skills acquired as part of the IBM Data Science Professional Certification (by Coursera), and generate a few most popular

neighbourhoods with information about interesting venues to visit. The details about each venue will then be suggested for the best possible exploration of the Stockholm city. **Data**

Based on the problem definition, following factors will contribute in deciding which neighborhood is popular:

number of café and restaurants in the neighborhood. whether the neighborhood is close to the city center.

 population density. We will focus around the Stockholm city centre to define our neighborhoods.

number of museums and attractions in the neighborhood.

- Following data sources will be needed to extract the required information: • Wikipedia, https://en.wikipedia.org/wiki/Stockholm_City_Centre, to obtain the information related to the districts/boroughs in the Stockholm city centre. • Foursquare API to obtain the location/venue data (i.e. number of museums, attractions, café, restaurants, etc.)

BeautifulSoup to parse the area information from a HTML table. **Geocoder** to obtain the coordinates of the Stockholm city center. • Folium to visualize the maps.

- **Neighborhood Candidates**
- Let's first install and import the required libraries, and then proceed with finding the latitude and longitude coordinates for the centroids of our candidate neighborhoods around the Stockholm city centre.

In [5]:

So let's get started!

Out[5]:

print(df.shape) df (39, 6)

District Area Population Density Borough **Province** Östermalm 3 0 Djurgården 290 788 Uppland

Following table represents the data extracted from the Wikipedia page of the Stockholm City Centre:

1 130 Kungsholmen Uppland Fredhäll 38 4,958 2 Gustav Vasa 80 12,911 161 Norrmalm Uppland 3 Gärdet 459 18,158 40 Östermalm Uppland Hedvig Eleonora 10,387 179 Östermalm Uppland 58 5 Hjorthagen-Värtahamnen 174 2,225 13 Östermalm Uppland 6 62 201 Norrmalm Uppland Jakob 3 7 Klara 71 1,597 22 Norrmalm Uppland 8 Kristineberg 63 5,572 Kungsholmen Uppland 88 9 Kungsholm 18,465 173 Kungsholmen Uppland 107 Kungsholmen 10 Lilla Essingen 23 4,519 196 Uppland 11 Mariatorget 62 14,099 227 Maria-Gamla stan Södermanland Marieberg 66 Kungsholmen Uppland Mellersta Högalid 13 184 Maria-Gamla stan Södermanland 9,914 14 Norra Adolf Fredrik 19 3,816 201 Uppland Norrmalm 15 137 Norra Johannes 66 9,043 Norrmalm Uppland 16 Norra Högalid 65 13,166 203 Maria-Gamla stan Södermanland 17 58 Norra Sofia 7,721 133 Katarina-Sofia Södermanland Oscars Kyrka 18 116 15,271 132 Östermalm Uppland **19** Reimersholme-Långholmen 52 2,349 45 Maria-Gamla stan Södermanland 20 Roslagstull 9 2,713 301 Norrmalm Uppland 21 Stadshagen 63 2,258 36 Kungsholmen Uppland Stora Essingen 22 72 3,954 55 Kungsholmen Uppland 23 72 Maria-Gamla stan Södermanland Storkyrkan 42 3,017 Stureplan-Lärkstaden 114 24 71 8,104 Östermalm Uppland 25 Södra Adolf Fredrik 35 3,703 106 Norrmalm Uppland 26 Södra Hammarbyhamnen 125 10,615 Katarina-Sofia Södermanland 85 **27** Södra Högalid 72 4,155 58 Katarina-Sofia Södermanland 28 Södra Johannes 15 2,011 134 Norrmalm Uppland Katarina-Sofia Södermanland 29 Södra Sofia 74 11,015 149 Katarina-Sofia Södermanland 30 Södra Station 18 4,844 269 Tekniska Högskolan 31 156 3,442 22 Östermalm Uppland 7 32 Universitetet 476 3,131 Östermalm Uppland 33 Västra Katarina 85 13,220 156 Katarina-Sofia Södermanland 34 Västra Matteus 84 14,272 170 Kungsholmen Uppland 35 Östra Katarina 85 19,855 234 Katarina-Sofia Södermanland 36 Östra Matteus 51 12,325 242 Kungsholmen Uppland 37 Östra Sankt Göran 53 14,079 266 Kungsholmen Uppland 38 Total 3577 296,323 83 Furthermore, using the geocoder, we obtained the latitude and longitude coordinates of the districts in the Stockholm city centre. These are represented as

District Area Population Density

788

4,958

12,911

18,158

10,387

3

130

161

40

179

290

38

80

459

58

Djurgården

Gustav Vasa

Hedvig Eleonora

Fredhäll

Gärdet

Hjorthagen-Värtahamnen 174 2,225 13 Östermalm Uppland 59.355180 18.100200 Uppland 59.329270 18.066005 6 62 Jakob 201 3 Norrmalm

Province

Latitude Longitude

Askrikefjärd

Hustega

Kappala

Killinge

Brevik

Italian

Restaurant

Harbor /

Seafood

Kungshamn

Duvnäs utskog

Nyckelviken

Skogalund Saltangen

Saltsjö-Duvnäs

Nacka Strand

Lillängen

Hästhagen

Nacka

Lilla Björkna:

Tollar

Restaurant

Marina

Bakery

History

Italian

Restaurant

Museum

Indian

Restaurant

Museum

Clothing

Store

Rudboda

Västra Yttringe

Bo

Högsätra

Larsberg

Björnbo

Skärsätra

Stockby

Näset

Herserud

Hjort Qigen

Baggeby

Uppland 59.324620 18.097800

Uppland 59.331030 18.005450

Uppland 59.342500 18.047750

Uppland 59.333610 18.113360

Uppland 59.335240 18.080460

Borough

Östermalm

Norrmalm

Östermalm

Östermalm

Kungsholmen

7

follows:

In [10]: df.head(15)

0

1

2

3

Out[10]:

In [12]:

Out[12]:

Eneby

Norra Angby

Bromma kyrka

71 1,597 22 Uppland 59.334670 18.068222 Klara Norrmalm 8 63 Kristineberg 5,572 88 Kungsholmen Uppland 59.336620 18.004860 9 Kungsholm 107 18,465 173 Kungsholmen Uppland 59.329010 18.048540 10 Lilla Essingen 23 4,519 196 Kungsholmen Uppland 59.325040 18.006770 11 Mariatorget 62 14,099 227 Maria-Gamla stan Södermanland 59.318307 18.063466 12 Marieberg 66 2,700 41 Kungsholmen Uppland 59.327960 18.018510 13 Mellersta Högalid Maria-Gamla stan Södermanland 59.317260 54 9,914 18.037730 14 Norra Adolf Fredrik 19 3,816 201 Norrmalm Uppland 59.337890 18.060060 Let us visualize the map of Stockholm, along with the districts and Boroughs in the Stockholm city centre. map stockholm + Stora Ursvik Stocksund romsten Lilla Ursvik Nya Ulriksdal Sticklinge Bergshamra Norra Ritorp Rissne. Tranholmen Hallonbergen Islinge Frösunda Flysta Storskogen Torsvik Lidingö Bällsta Hagaparken Norra Djurgården Sundbyberg

Scona

Västra skogen

Huvudsta

Lilla Alby,

Ulvsunda industriområde

167

Solna Kyrkby

165/164

Ladugårdsgärdet Vasastaden Riksby Mölna Halvkakssunde Akeshov Traneberg 600 Ulvsunda Kristineberg Östermalm Åkeslund Nockebyhov Kungsholmer Abrahamsberg Kungshamn Alvik 0 Nockeby-Marieberg Duvnäs utskog Djurgården Stockholm Appelviken Lilla Esengen Höglandet Kärsön Lilla Björk Nyckelviken Smedslätten Reimercholme Södermalm Danviken Klubbfjarden Herriksdal Södermalm • Skogalund Saltangen Nacka. Eknäs Liljeholmen Alphyddan Aspudden-Saltsjö-Duvnäs Lillängen Södra Hammarbyhamnen Hägersten gshättan Midsommarkransen Hästhagen Mälarhöjden Hägerstensåsen Arsta Hammarbyhöjden Västertorp traskogens turreservat Björkhagen Leaflet | Data by @ OpenStreetMap, under ODbL. Methodology After successfully extracting the information about Stockholm city centre from its Wikipedia page, and visualizing it on the city map, we used the Foursquare API to explore venues around the area. Let us explore the 5 most popular/common venues in a given neighborhood. Following table represents the details with respect to a neighborhood, that includes the area, population, density, Boroughs, Province, laatitude, longitude coordinates, and 5 most common venues. stockholm merged.head() Out[27]: 1st Most 2nd Most 3rd Most 4th Most 5th Most **Neighborhood Area Population Density** Borough Province Latitude Longitude Common Common Common Common Common **Venue** Venue Venue **Venue** Venue Theme Park Scandinavian Djurgården 290 788 Östermalm 18.09780 Ride / Museum Tram Station Music Venue 3 Uppland 59.32462 Restaurant Attraction Thai Indian Fredhäll 38 4,958 130 Kungsholmen Uppland 59.33103 18.00545 Café Park Mini Golf 1 Restaurant Restaurant

Uppland 59.34250

Uppland 59.33361

Uppland 59.33524

18.04775

18.11336

18.08046

Pizza Place

Middle

Eastern

Hotel

Restaurant

Café

Science

Museum

Scandinavian

Restaurant

3 459 Gärdet Hedvig 58 Eleonora

Out[35]:

Gustav Vasa

12,911

18,158

10,387

Ulvsunda

Appelviken

Klubbfjarden

Aspudden-

Midsommarkransen

Smedslätten

Hägersten

Mälarhöjden Hägerstensåsen

Västertorp

227

201

133

36

55

106

85

149

269

170

234

266

83

Pub

Café

Café

Plaza

Café

Pier

Café

Café

Hotel

Science Museum Middle Eastern Restaurant

Pizza Place

Pizza Place

Gym / Fitness Center

Gym / Fitness Center

Mariatorget

Norra Sofia

Stadshagen

Södra Sofia

Södra Station

Västra Matteus

Östra Katarina

Total

Östra Sankt Göran

Gärdet

Stora Essingen

Södra Adolf Fredrik

Södra Hammarbyhamnen

Norra Adolf Fredrik

Abrahamsberg

Åkeslund

Höglandet

Nockeby.

ckebyhov

Kärsön

attan

kogens

11

14

17

21

22

25

29

30

34

35

37

38

In [37]:

Out[

In [39]

Out[39]

In [40]

Out[40]

8

10

15

Kristineberg

Lilla Essingen

Mellersta Högalid

Norra Johannes

Södra Högalid

31 Tekniska Högskolan

Thank you for stopping by!

88

196

184

137

22

161

40

179

Clustering the Neighborhoods using K-means

Norrmalm

Östermalm

Östermalm

and popular unsupervised machine learning algorithms, where we specified total number of clusters as 5.

Kristineberg

*Lilla Esongen

Marielog

Reimersholme

Liljeholmen

Kungsholmen

Riddarfjärden

Arsta

80

One can visualize these clusters on a map as follows: In [35]: map_clusters 68 169 Tranholmen Hallonbergen Rudboda Bo Näset + Islinge Björnbo Duvbo Frösunda Storskogen Torsvik Västra Yttringe Hustega Lidingö Bällsta Norra Djurgården Sundbyberg Hagaparken) Herserud Eneby Lilla Alby Stockby Killinge 165/164 Baggeby Hjortogen Bromma kyrka Solna Kyrkby Högsätra Huvudsta Kappala Larsberg Ulvsunda industriområde rra Angby Brevik Västra skogen Skärsätra Ladugårdsgärdet Riksby Vasastaden Halvkakssundet Åkeshov

Stockpolm

Södemalm

Södermalm 💮

Östermalm

Djurgården

Danviken

Alphyddan

Björkhagen

Henriksdal

Södra Hammarbyhamnen

Hammarbyhöjden

In order cluster the neighborhoods based on their similar characteristics, K-means clustering approach was used. K-means clustering is one of the simplest

Enskede gård Fruängen Kärrtorp Sätra, iseberg Enskedefältet Nackareservatet | 260 Långbro Orby slott i Nacka Gamla Enskede Solberga Stureby Bagarmossen Segeltorp Alvsjö Svedmyra Leaflet | Data by © OpenStreetMap, under ODbL. **Results and Discussion** As a result, the neighborhoods around the Stockholm city centre were segmented into 5 clusters using the K-means clustering approach (unsupervised machine learning), where each cluster represent a group of similar neighborhods. In the following data frames, one can observe the details about the venues in each of the 5 clusters. First cluster (see Cluster Label 0) represents the neighborhoods sharing similar venues, such as café, pub, pizza place, restaurants, gym etc. being the most popular. This is a nice neighborhood to visit if you want to enjoy a nice variety of food, that includes Scandanavian, Thai, and Italian choices, to mention a few. This neighborhood is densely populated and quite popular amongst the locals and the visitors. It is also well connected with public transport, making it easier to commute around the area. Also, there are hotels nearby, making it a perfect neighborhood for tourists to stay and explore. Other facilities include supermarket, clothing store, bookstore and shopping plaza. Second cluster (see Cluster Label 1) consists of a less crowded neighborhood "Gärdet", which is popular for Science and History museums and, authentic middle eastern restaurant. This is a nice choice of neighborhood if you enjoy visiting museums, and prefer less crowded places. **Third cluster** (see Cluster Label 2) is dominated by neighborhoods with tourist attractions such as theme parks, skate park, beaches, lake, museums, and is also popular for the nightlife. There are a variety of restaurants in this neighborhood, including Sushi restaurant. Also, there are hotels and hostels in the vicinity, and the neighborhood is well connected with the public transport. Fourth cluster (see Cluster Label 3) consists of 2 areas, popular for café, restaurants, coffee shop and bookstore. This area is also famous for an Indian restaurant. Fifth cluster (see Cluster Label 4) consists of neighborhoods popular for restaurants, parks, bar, grocery store, etc. and public transport is available. stockholm merged2.loc[stockholm merged2['Cluster Labels'] == 0, stockholm merged2.columns[[0]+[1]+list(range(5, stockholm) olm_merged2.shape[1]))]] Out[36]: Neighborhood Density 1st Most Common Venue 2nd Most Common Venue 3rd Most Common Venue 2 Gustav Vasa 161 Café Pizza Place Bakery 5 Hjorthagen-Värtahamnen Light Rail Station Pizza Place Thrift / Vintage Store 13 9 173 Scandinavian Restaurant Italian Restaurant Kungsholm Café

olm merged2.shape[1]))]] Neighborhood Density 1st Most Common Venue 2nd Most Common Venue 3rd Most Common Venue

Bar

Hotel

Hotel

Bus Stop

Restaurant

Coffee Shop

Scandinavian Restaurant

Scandinavian Restaurant

Scandinavian Restaurant

Scandinavian Restaurant

Scandinavian Restaurant

Convenience Store

Asian Restaurant

Coffee Shop

Clothing Store

Supermarket

Bus Station

Clothing Store

Supermarket

Thai Restaurant

Clothing Store

History Museum

Bus Station

Hostel

Thai Restaurant

Sushi Restaurant

Sushi Restaurant

Gym / Fitness Center

Bookstore

Café

Hotel

Scandinavian Restaurant

Scandinavian Restaurant

olm	olm_merged2.shape[1]))]]								
<u></u>	Neighborhood	Density	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue				
0	Djurgården	3	Theme Park Ride / Attraction	Scandinavian Restaurant	Museum				
4	Hedvig Eleonora	179	Scandinavian Restaurant	Hotel	Italian Restaurant				
6	Jakob	3	Hotel	Plaza	Café				
7	Klara	22	Scandinavian Restaurant	Hotel	Burger Joint				
12	Marieberg	41	Park	Scandinavian Restaurant	Skate Park				
16	Norra Högalid	203	Bar	Bakery	Sushi Restaurant				
18	Oscars Kyrka	132	Scandinavian Restaurant	Park	Italian Restaurant				
19	Reimersholme-Långholmen	45	Hotel	Beach	Other Nightlife				
20	Roslagstull	301	Sushi Restaurant	Miscellaneous Shop	Lake				
23	Storkyrkan	72	Scandinavian Restaurant	Plaza	Café				

	19	Reimersholme-Långholmen	45	Hotel	Beach	Other Nightlife	
	20	Roslagstull	301	Sushi Restaurant	Miscellaneous Shop	Lake	
	23	Storkyrkan	72	Scandinavian Restaurant	Plaza	Café	
	24	Stureplan-Lärkstaden	114	Scandinavian Restaurant	Hotel	Asian Restaurant	
	28	Södra Johannes	134	Scandinavian Restaurant	Café	Hostel	
	32	Universitetet	7	Café	Convenience Store	Bus Station	
]:		<pre>ckholm_merged2.loc[s _merged2.shape[1]))] Neighborhood Density 1s</pre>	11				mns[[0]+[1]+list(range(5, stockh
	1	Fredhäll 130		Café India	an Restaurant Tha	ai Restaurant	
	33	Västra Katarina 156		Café	Coffee Shop	Bookstore	
]:		ckholm_merged2.loc[s _merged2.shape[1]))]		m_merged2['Cluster	Labels'] == 4, stoc	kholm_merged2.colu	mns[[0]+[1]+list(range(5, stockh
	OTI	5 - 0 - 1 - 1 - 1 - 1 - 1 - 1 - 1					

Conclusion

Park

Bar

Bar

Grocery Store

Sushi Restaurant

the locations to generate the major areas of interest containing the most popular places to explore.

characteristics they like in a neighborhood in every recommended zone (or cluster).

This page will be further updated, as I am still exploring the power of data science!

Scandinavian Restaurant

For example: • If you are a tourist visiting the Stockholm city center, the neighborhoods in the third cluster are a must visit! The venues suggested in the first cluster are also very interesting.

Final decision on the best neighborhood to visit will be made by the tourists, visitors or investors, based on their personal preferences about the specific

• If you are an investor, interested in the restaurant business, do check the neighborhoods and venues in the first, third and fifth cluster. • If you want to visit museums, check out the first and third cluster. • Do check out the map above to explore the neighborhoods, and refer to the tables for information about the venues that interests you.

Middle Eastern Restaurant

Restaurant

Bakery

Bakery

The purpose of this project was to explore the Stockholm city centre and find out the most popular places to visit. Using the Foursquare API, we gathered the location data about the popular venues around the Stockholm city centre. Using unsupervised machine learning approach, K-means clustering, we clustered

Italian Restaurant

Café