Study of urban green space in relation to happiness

Problem Description

Assisting the urban planning department of Brussels, Belgium to explore the density of urban green space and location for new green space.

A study of the University of Vermont shows a clear relationship between urban green space and happiness. The study was conducted using sentiment analysis (Hedonometer) on Twitter posts. This allowed measurement of the level of happiness or positiveness by comparing the sentiment at the time before visiting, during visiting and after visiting urban parks.

Abstract from the study:

1. With more people living in cities, we are witnessing a decline in exposure to nature. A growing body of research has demonstrated an association between nature contact and improved mood.
2. Here, we used Twitter and the Hedonometer, a world analysis tool, to investigate how sentiment, or the estimated happiness of the words people write, varied before, during, and after visits to San Francisco's urban park system. We found that sentiment was substantially higher during park visits and remained elevated for several hours following the visit.
3. Leveraging differences in vegetative cover across park types, we explored how different types of outdoor public spaces may contribute to subjective well-being. Tweets during visits to Regional Parks, which are greener and have greater vegetative cover, exhibited larger increases in sentiment than tweets during visits to Civic Plazas and Squares.
4. Finally, we analyed word frequencies to explore several mechanisms theorized to link nature exposure with mental and cognitive benefits. Negation words such as ‘no’, ‘not’, and ‘don't’ decreased in frequency during visits to urban parks.
5. These results can be used by urban planners and public health officials to better target nature contact recommendations for growing urban populations.

This is not the only study which has shown that having access to, or exposure to nature makes people kinder, happier and stimulates creativity.

Audience

The study is conducted for the urban planning department of the city of Brussels. This will allow them to locate areas for new green space development in the city. As such they can more effectively use their available budgets.

The results of the study can potentially benefit all residents of the city of Brussels.

Mental health is an important topic worldwide. With an ever increasing healthcare cost, it is important to explore natural and non-medical factors which can have a substantial benefit on the health of the general public. It is far more important to prevent then to cure. Taking the above into account, we will attempt to use maps and location data to visualize the balance between greenspace and urban space in different parts of the city. Just as we are building further on a previous data science study based on sentiment analysis, this analysis can then in turn be used later on to link to public health records to map if there is indeed a link between mental health and happiness and relative access to urban greenspace.

References:

https://www.uvm.edu/uvmnews/news/city-parks-lift-mood-much-christmas-twitter-study-shows

https://besjournals.onlinelibrary.wiley.com/doi/full/10.1002/pan3.10045

https://greatergood.berkeley.edu/article/item/how_nature_makes_you_kinder_happier_more_creative

Data Description

The following data sources can be leveraged to perform this study:

Foursquare:

Parks in Brussels:

https://foursquare.com/explore?mode=url&near=Brussels%2C%20Belgium&nearGeoId=72057594040728802&q=Park

Forest in Brussels:

https://foursquare.com/explore?mode=url&near=Brussels%2C%20Belgium&nearGeoId=72057594040728802&q=forest

The above links are a guideline to initiate the study and will be translated in API queries when performing the actual data analysis using the Foursquare Developer API. In case we find that Brussels has a near perfect distribution of parks, we can leverage the average rating values from Foursquare to decide which parks need changes, maintenance or redevelopment.

Foursquare data items which can potentially be used:

- Location name for identification
- Latitude and longitude for mapping
- Rating of the location for quality indication

- Amount of tips to estimate popularity

Open Data Brussels: Parks: https://opendata.brussels.be/explore/dataset/parks/information/

Parks data from the Open Data Brussels platform can perhaps be used during the analysis to see how good the Foursquare coverage is. Based on this the city council could perhaps leverage the use of social internet tools like Foursquare to assure that residents and tourists can get easy access to info regarding urban green space.

Open Data Brussels data items which can potentially be used:

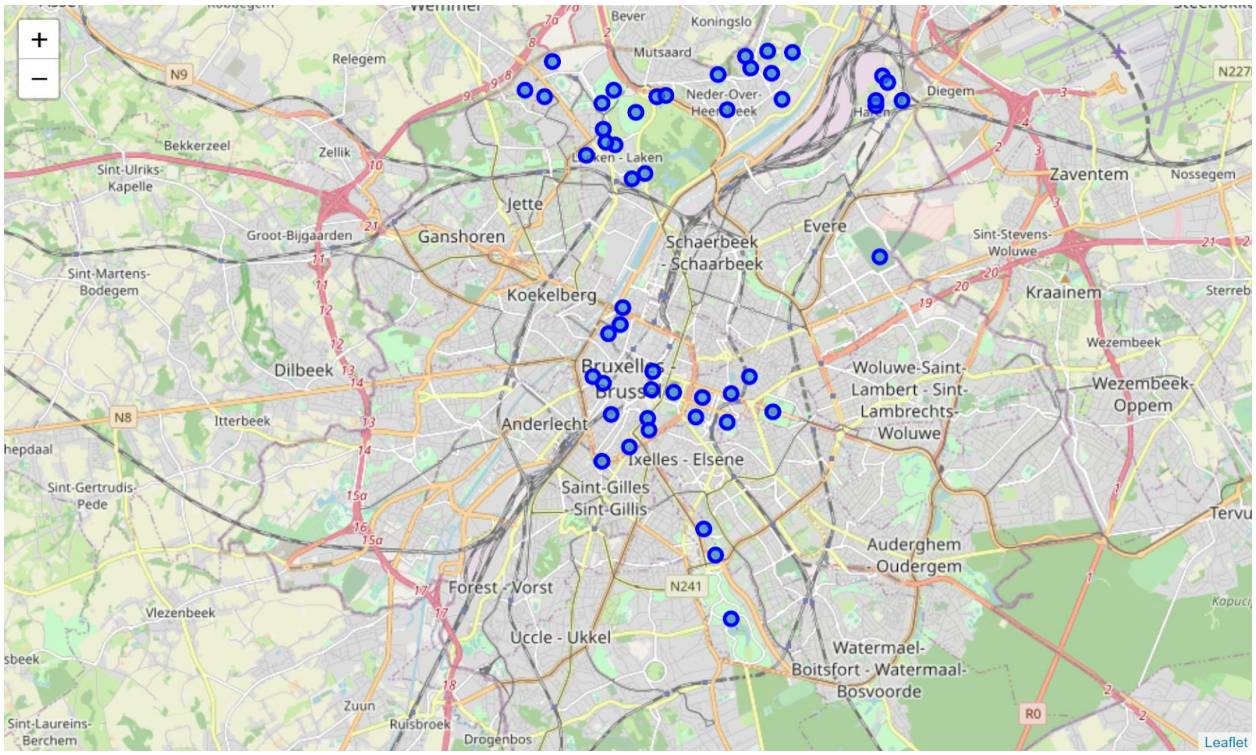
- Location name for identification
- Latitude and longitude for mapping

Methodology

Choosing our most relevant data set

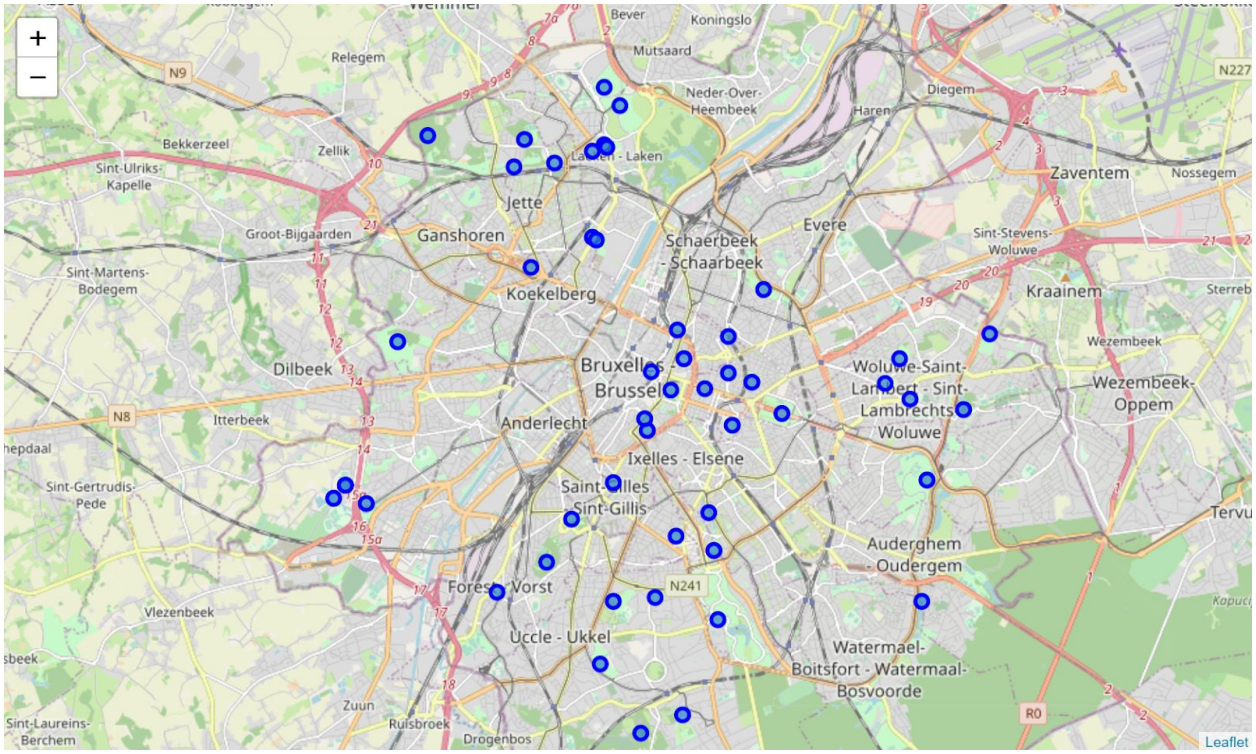
Step 1: Visualize the parks on the map according to Open Brussels Data

We are using folium to generate a map of Brussels with all parks plotted according to the Open Brussels Data platform. It appears on the map that this data isn't very dense. Additionally, we note that the last update of the data set dates back to APril 2014, so there is a real chance that this data is not very relevant for today's research.



Step 2: Visualize the parks on the map according to Foursquare

The data retrieved from Foursquare is relatively good in a sense that there are more data points. We can assume that the data is of good quality considering that there is a specific category in Foursquare to mark parks. This increases the confidence in the quality of this data. Additionally since Foursquare is a live platform, we have more up to date information available.



We conclude that the Open Brussels Data map is not the best to use as the coverage of Foursquare is more up to date and better.

The Open Brussels data set is outdated as it was created on April 14, 2014. There is no newer version available. As such we continue the analysis with the Foursquare data.

Step 3 Clustering of the parks

We will apply the K-means algorithm to cluster the parks in groups, based on the quality or rating of the park.

Use 4 clusters

We will organize the parks in 4 clusters. This is a manageable amount to use later on for comparison of quality between different parks. In an initial approach we attempted to use additional variables to guide the clustering, such as number of likes and number of tips for each park according to

Foursquare. A few rounds of testing revealed that there was no additional benefit in using these extra variables. We reached the cleanest and best separation solely based on the rating. The additional variables will be used in the end however for presenting the data points on the map and to provide a data overview. As such the urban planning department can perform specific investigations of parks with a high rating which also have a high like count and a lot of tips. The tips especially can be useful to identify the strong points or the bad aspects of the park.

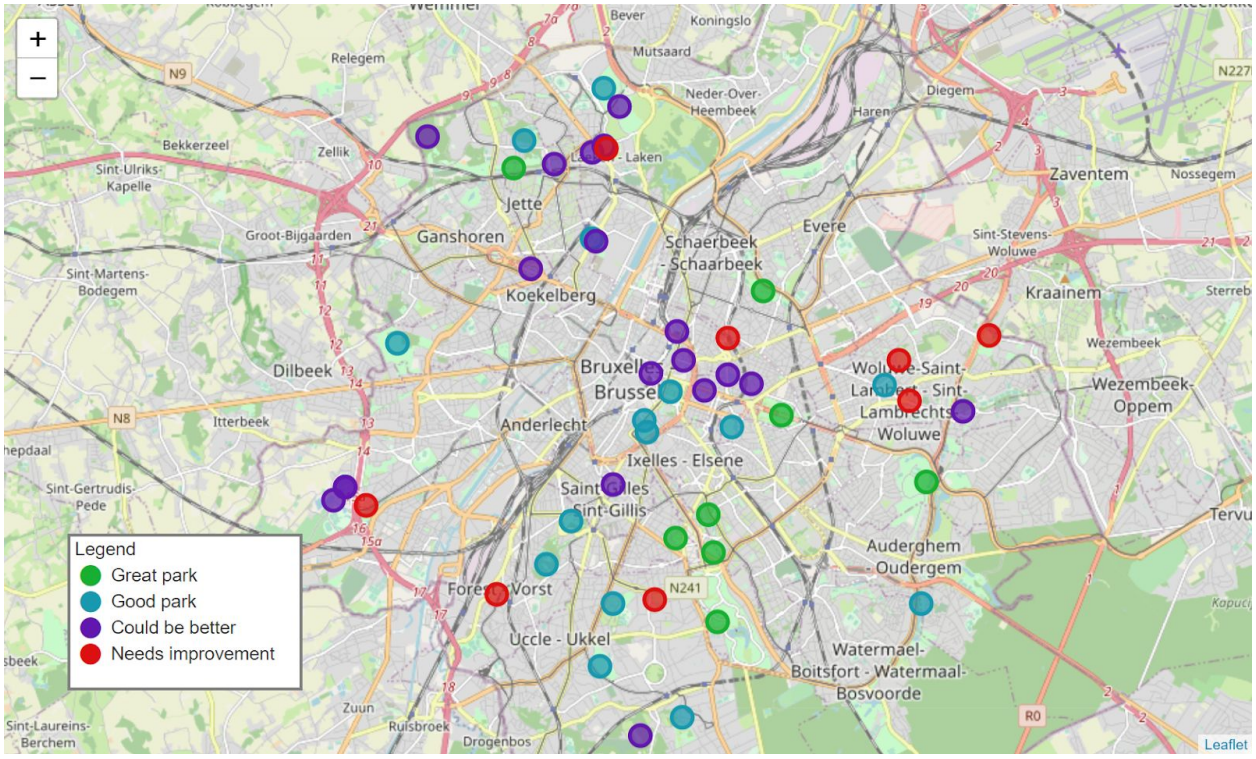
The naming of the clusters will be as follows:

- Great Park
- Good Park
- Could be better
- Needs improvement

Results

After performing data cleansing and merging the list of parks with their relevant details for this study we can go ahead to plot the datapoints on a map.

Clustered map



Below we provide a list of the clusters and the associated details of each park. Note that the numeric cluster labels aren't assigned in order of 0-3 according to best to worst. The algorithm doesn't care about this specific order as it is designed only to separate into clusters.

Cluster 1: Great parks - Rating > 9.0

Cluster Labels	id	name	latitude	longitude	rating	tipCount	likesCount	Cluster Name
2	4ade0d7ff964a520e86b21e3	Jubelpark / Parc du Cinquantenaire (Jubelpark)	50.840654	4.392189	9.3	151	1006	Great park
2	4baf1d46f964a520f4eb3be3	Parc Tenboschpark (Tenboschpark)	50.820295	4.364380	9.1	18	98	Great park
2	4b93c7c6f964a520455134e3	Ter Kamerenbos / Bois de la Cambre (Ter Kamere...	50.806322	4.375214	9.4	74	563	Great park
2	4bc19ce92a89ef3bd292f288	Koning Boudewijnpark / Parc Roi Baudouin (Koni...	50.881646	4.321868	9.1	7	52	Great park
2	4cc6c1e4b2beb1f768042a4c	Parc de Woluwepark (Woluwepark)	50.829686	4.430074	9.5	36	254	Great park
2	4c9c451bd3c2b60c9d74b6bc	Park van Abdij Ter Kameren / Parc de l'Abbaye ...	50.817773	4.374128	9.2	5	54	Great park
2	4b71b434f964a52060572de3	Josafatpark / Parc Josaphat (Josafatpark)	50.861241	4.387321	9.2	30	156	Great park
2	4bba10c8cf2fc9b64a15a102	Étangs d'Ixelles / Vijvers van Elsene (Vijvers...	50.824100	4.372925	9.2	14	170	Great park

Cluster 2: Good parks - Rating between 8.0 and 9.0

Cluster Labels	id	name	latitude	longitude	rating	tipCount	likesCount	Cluster Name
0	4ade0d88f964a520176c21e3	Warandepark / Parc de Bruxelles (Warandepark)	50.844442	4.363067	8.6	81	531	Good park
0	4bdfcad70ee3a5936b6735b0	Parc d'Egmontpark (Egmontpark)	50.837804	4.356594	8.8	20	98	Good park
0	4c07d5f0271dc9b6ea6f2b9a	Parc Léopold / Leopoldpark (Leopoldpark)	50.838603	4.379168	8.7	23	185	Good park
0	4bb8c7aab35776b04236c901	Parc Dudenpark (Dudenpark)	50.816022	4.330330	8.9	16	81	Good park
0	4b93c427f964a520925034e3	Parc du Wolvendaelpark (Wolvendaelpark)	50.799073	4.344406	8.8	11	101	Good park
0	4baa84cff964a520ea703ae3	Kleine Zavel / Place du Petit Sablon (Kleine Z...	50.839872	4.356024	8.7	12	115	Good park
0	4e4d17bd62e1e5f4671f145f	Parc Seny / Senypark (Senypark)	50.809460	4.428692	8.9	2	28	Good park
0	4c0a549b3c70b713e0f7275b	Parc de Forest / Park van Vorst (Park van Vorst)	50.823174	4.336691	8.6	15	98	Good park
0	4bc78ac115a7ef3b889a79da	Parc Georges Henri-park (Georges Henri-park)	50.845661	4.419336	8.8	25	104	Good park
0	4c308ee516adc928b149bf9c	Parc de la Sauvagèrepark (Sauvagèrepark)	50.790530	4.365877	8.7	7	34	Good park
0	4d30cb30982f224b9c8bb415	Parc Brugmannpark (Brugmannpark)	50.809496	4.347896	8.6	2	21	Good park
0	4c5e91e385a1e21ed6905c11	Parc du Scheutbospark (Scheutbospark)	50.852493	4.291295	8.4	5	17	Good park
0	4db1831bfa8ca4b3e9f8b8ab	Bois Regional de Dieleghem	50.886164	4.324430	8.3	2	28	Good park
0	4c43194cd7fad13a1f040ada	Ossegempark / Parc d'Osseghem (Ossegempark)	50.894707	4.345377	8.3	6	25	Good park
0	544ab8fd498e954e946304c7	Parckfarm	50.869869	4.342229	8.5	1	12	Good park

Cluster 3: Could be better - Rating between 7.8 and 8.0

Cluster Labels	id	name	latitude	longitude	rating	tipCount	likesCount	Cluster Name
1	4c9b3c2680958cfa407149d4	Maria-Louizasquare / Square Marie-Louise (Mari...	50.847248	4.378141	8.0	9	43	Could be better
1	4bd2b7e6caff9521275cd3f0	Square Ambiorixsquare (Ambiorixsquare)	50.845947	4.384257	8.2	10	96	Could be better
1	4ade0d89f964a5201e6c21e3	Elisabethpark / Parc Élisabeth (Elisabethpark)	50.864884	4.326357	8.1	15	74	Could be better
1	4f22c8d0e4b06176428ecbc7	Les Jardins du Fleuriste	50.885331	4.345573	7.9	3	17	Could be better
1	4bf91d745efe2d7f8cb66b34	Parc Maloupark (Maloupark)	50.841218	4.439753	8.1	12	34	Could be better
1	4dadb1fe1e72b2444e46df02	Parc de Laeken / Park van Laken (Park van Laken)	50.891677	4.349486	8.1	11	91	Could be better
1	536bb49e498e2b47e5c80156	Parc Tour & Taxis	50.869448	4.343234	8.0	1	9	Could be better
1	4cc6e1e91e596dcb662be167	Kauwberg	50.787584	4.355199	8.0	1	10	Could be better
1	4dc51250e4cd169dc642c5fd	Parc Pierre Pauluspark (Pierre Pauluspark)	50.829065	4.347695	8.0	4	18	Could be better
1	4c6556e8d02b20a1ec2c9d90	Jardin Botanique / Kruidtuin (Kruidtuin)	50.854572	4.364791	8.2	44	177	Could be better
1	4d7a3a691bf03704c44bded5	Neerpede	50.828727	4.277332	7.9	3	26	Could be better
1	5186762f498efea876d1f1c7	Etangs de Neerpede	50.826600	4.274421	7.8	2	7	Could be better
1	4e174e1ad4c062b044dcc360	Place de la Liberté / Vrijheidsplein (Vrijheid...	50.849778	4.366510	8.0	9	30	Could be better
1	4bee9e4dd355a593cf0f0b60	Jeugdpark / Parc de la Jeunesse (Jeugdpark)	50.882322	4.332220	7.7	3	12	Could be better
1	4c78defedf08a1cd35cfd85d	Parc de la Pedepark (Pedepark)	50.828718	4.277317	7.7	3	18	Could be better
1	4ba20dc3f964a52087d837e3	Square Frère-Orbansquare (Frère-Orbansquare)	50.844792	4.371751	7.9	8	24	Could be better
1	4fec08b5e4b06e7fdf6f2606	Prins Leopoldsquare / Square Prince Léopold (P...	50.884269	4.342278	7.8	0	8	Could be better
1	4e4d029745dd4c13eda20140	Place Sainte-Gudule / Sint-Goedeleplein (Sint-...	50.847669	4.357880	7.8	11	45	Could be better
1	4bf6e7d44a67c9284ba622cf	Bois du Laerbeek / Laarbeekbos	50.886846	4.299088	7.8	5	30	Could be better

Cluster 4: Needs Improvement - Rating < 7.8

Cluster Labels	id	name	latitude	longitude	rating	tipCount	likesCount	Cluster Name
3	4d6f5ecb5b5ca1cd0d6a6b44	Square Armand Steurssquare (Armand Steurssquare)	50.853434	4.377987	7.6	1	15	Needs improvement
3	4f5cdd19e4b075ad7ad8bed9	Parc Tomberg	50.843068	4.425820	7.5	0	7	Needs improvement
3	4b9227ecf964a520c7e933e3	Abdij van Vorst / Abbaye de Forest (Abdij van ...	50.810903	4.317435	7.5	3	18	Needs improvement
3	4b52fc16f964a520fc8b27e3	Parc Sobieskipark (Sobieskipark)	50.884826	4.346192	7.4	4	11	Needs improvement
3	4db151714b22ba776d7f82e1	Parc des Étangs / Vijverspark (Vijverspark)	50.825736	4.283016	7.3	4	22	Needs improvement
3	4c76db7f923ba143324065e6	Parc Montjoiepark (Montjoiepark)	50.809999	4.358759	7.2	1	11	Needs improvement
3	4c20804d8b3aa5931bff9e5f	Hof Ter Musschen Park / Parc Hof Ter Musschen ...	50.853769	4.446855	7.1	4	6	Needs improvement
3	4c835703d8086dcb19657f52	Parc de Roodebeek / Roodebeekpark (Roodebeekpark)	50.849658	4.423155	7.1	8	17	Needs improvement

Discussion

Dead spots

The map visualization allows us to identify dead spots where little to no parks are available. The urban planning department can use this to identify the locations of new parks to be developed.

Quality of parks

The clustering of the parks according to the rating allows us to identify parks ranging from "Great Park" to "Needs improvement".

This data easily allows the urban planning department to identify existing parks that need to be inspected for improvements or potential redevelopment.

Best parks as an example

The parks in cluster "Great Parks" can be used for a further study on what makes these parks so great. The final data set also includes how many tips and likes a park has on Foursquare. We can see that the best parks have more tips and likes, so this could be relevant. This allows us to inspect what people find good or bad about the park. Eventually this allows us to identify how to improve the other parks in the city by applying these same elements.

Conclusions

IMPORTANT Unfortunately the Foursquare API only allows 50 premium calls, these are required to fetch the details like the park rating. The data is heavily influenced by this rating.

We are well aware that this study is incomplete, but it could easily be repeated with all data instead of the 50 limit if budget is available to pay for the API.

As such we will consider these preliminary results, but we know this can be repeated on a full data set with equally relevant outcomes

The clustering approach based on parks ratings gives us good indicators on where to start with planning of new parks and improvement of existing parks.

On a long term this could have a likely benefit to the general well being of the population as several studies have indicated the impact of parks and urban green space on happiness. Overall this would lead to better quality of life, reduced health care costs and potentially a more safe and higher quality environment overall. Something which is important to the capital of Belgium and of the European Union.

Further steps

As a next step, it would be great to combine additional data to this report to make it more robust.

- With some budget we can get all relevant data out of the Foursquare API
- It would be interesting to combine police reports on violence and drug use in or around these parks to see the effect on the rating. Maybe the parks are great in facilities, but problematic in terms of safety/crime.
- It would be interesting to take local surveys from the residents about these parks. It is possible that Foursquare is biased towards mostly tourists and likely younger people, but not (elderly) residents.