Geodata sources Lab 1.2 – Quality of geodata sources

Due date: 16th November 2020

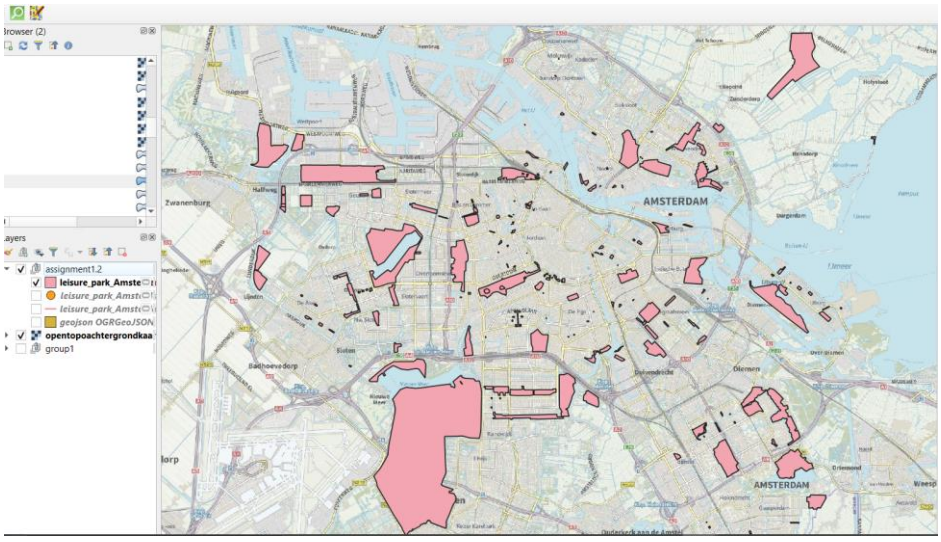
Name: Lena Walcher

Student ID: 2818833

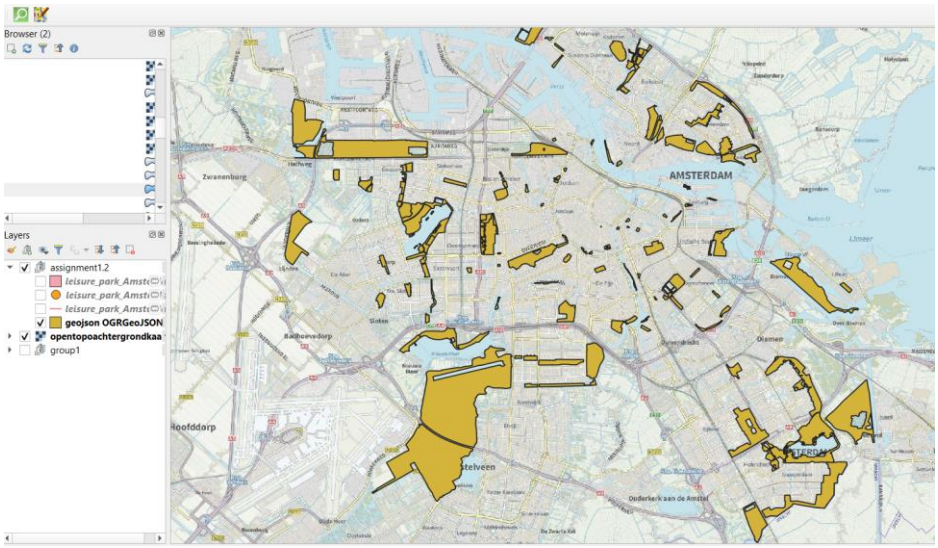
The purpose of this assignment is an analysis and comparison of different geodata sources to compare the different qualities of maps and layers. The goal is to find the most accurate source for further analysis.

With the QuickOSM-tool in QGIS I added the layer “leisure_park_Amsterdam” (screenshot 1) by using the query key:leisure, value:park. To compare this data with another source I used the website (https://maps.amsterdam.nl/) to load a GeoJson with the following source for the file (https://maps.amsterdam.nl/open_geodata/geojson.php?KAARTLAAG=PARKPLANTSOENGROEN&THEMA=stadsparken). Both layers can be seen as visualized below.

Screenshot 1

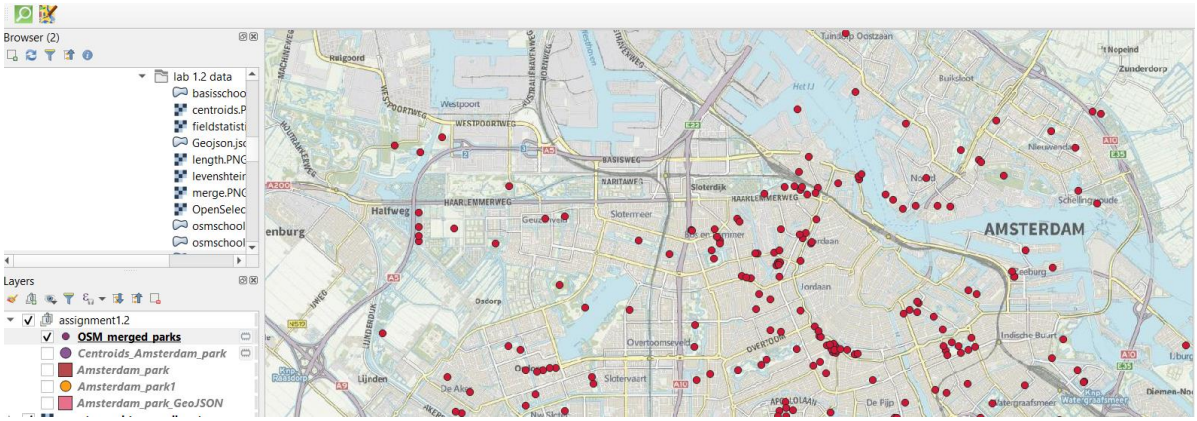


Screenshot 2

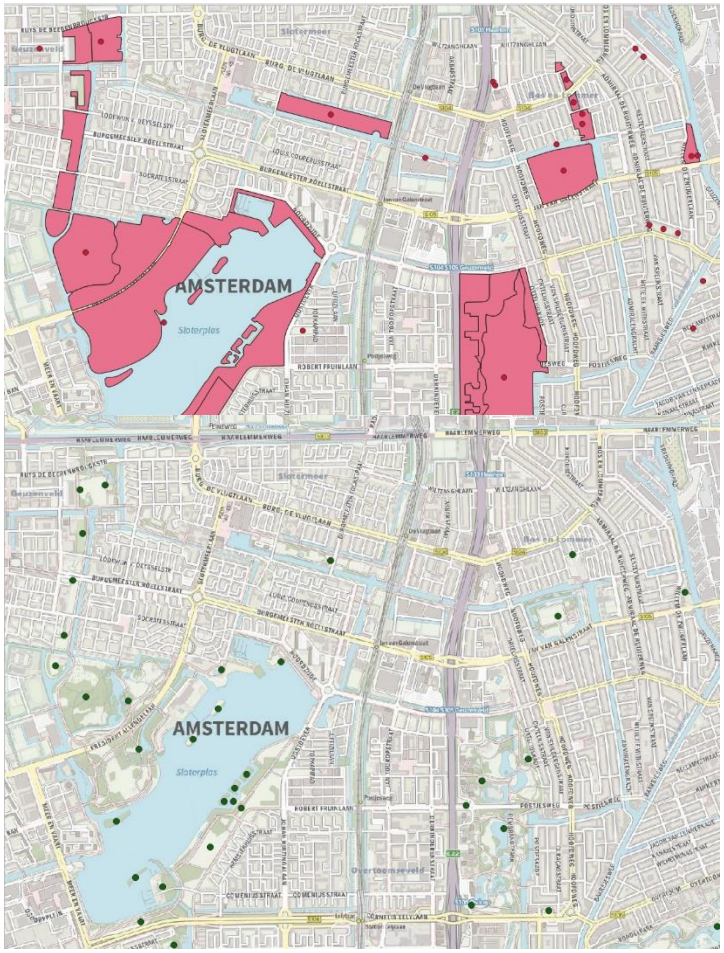


I then saved all layers as ESRI shapefiles and reopened them. To be able to merge the two OSM-layers into one layer I turned the polygon layer into a point layer (Centroids_Amsterdam_park).

I then merged the two layers (point and polygon) of from the OSM-query into one layer and named it OSM_merged_parks.



When visually comparing the merged OSM-layers with the layers form the City of Amsterdam (Amsterdam_park_GeoJSON) it is already clear, that even though there is certainly a lot of common datapoints, they do differ.



In order to actually measure the quality of the dimensions I first turned the reference dataset into a point layer using the centroid (Centroids_Amsterdam _park_GeoJSON), then joined the data items from the OSM-layers (OSM_merged_parks) with the reference dataset with a nearest neighbor of 1 and a distance of 200m.

As the green points in the screenshot on the left show in comparison to the picture above, the merge resulted in a lot less points of the merged OSM-data (OSM_merged_parks).

In order to see which of the 122 attributes represent the same parks within the data one can use the Levensthein distance to compare the names of the parks from each source with another.

The new layer Calculated_leven containing the Levenshtein distance between “names” and “naam” of the previously joined layer can then be used to select similar attributes.

Calculated_leven :: Features Total: 122, Filtered: 122, Selected: 0

	Naam	full_id	name	distance	feature_x	feature_y	nearest_x	nearest_y	leven
1	`t Kleine Loopv...	w812982275	't Kleine Loopve...	93,6440817351...	120323,410740...	481740,495981...	120231,370963...	481723,236390...	0,0556
2	Albert Neuhuys	w4842719	NULL	3,51249694707...	118475,232423...	485392,628294...	118475,280525...	485396,140462...	NULL
3	Amstelpark	w26516487	Amstelpark	84,1427691272...	121374,306874...	482519,563069...	121353,492489...	482438,035362...	0
4	Amsterdamse B...	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	Apollolaan	w73998931	NULL	7,05558055383...	119911,061182...	484858,348152...	119907,121895...	484852,494672...	NULL
6	Baanakkerspark	w169173148	Baanakkerspark	150,543993819...	124833,682522...	489920,649115...	124960,361528...	490001,987442...	0
7	Beatrixpark	w26037982	Beatrixpark	7,85401391058...	120555,462035...	483963,273041...	120554,430837...	483955,487017...	0
8	Bella Vistapark	r1043467	Bella Vistapark	0,97624785773...	122947,468373...	483466,264070...	122947,903519...	483465,390166...	0
9	Bergwijkpark	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
10	Bijlmerweide	w817343463	Bijlmerweide	80,4863885082...	127636,197502...	481882,308013...	127582,075202...	481941,880115...	0
11	Bilderdijkpark	w220886085	Bilderdijkpark	9,71973783316...	119796,247581...	487242,496487...	119787,341470...	487238,603610...	0
12	Bogortuim	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
13	Bogortuin	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
14	Bogortuin	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
15	Buiksloterbreek...	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
16	Buikslotermeer...	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
17	Darwinplantsoen	w545029841	Darwinplantsoen	10,5484095253...	123899,168315...	484244,490954...	123903,899924...	484235,063291...	0
18	De Bretten	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
19	De Hoftuin	w702630867	Weesperplantso...	36,4031699503...	122113,953870...	486524,717958...	122149,689418...	486531,657804...	1,300

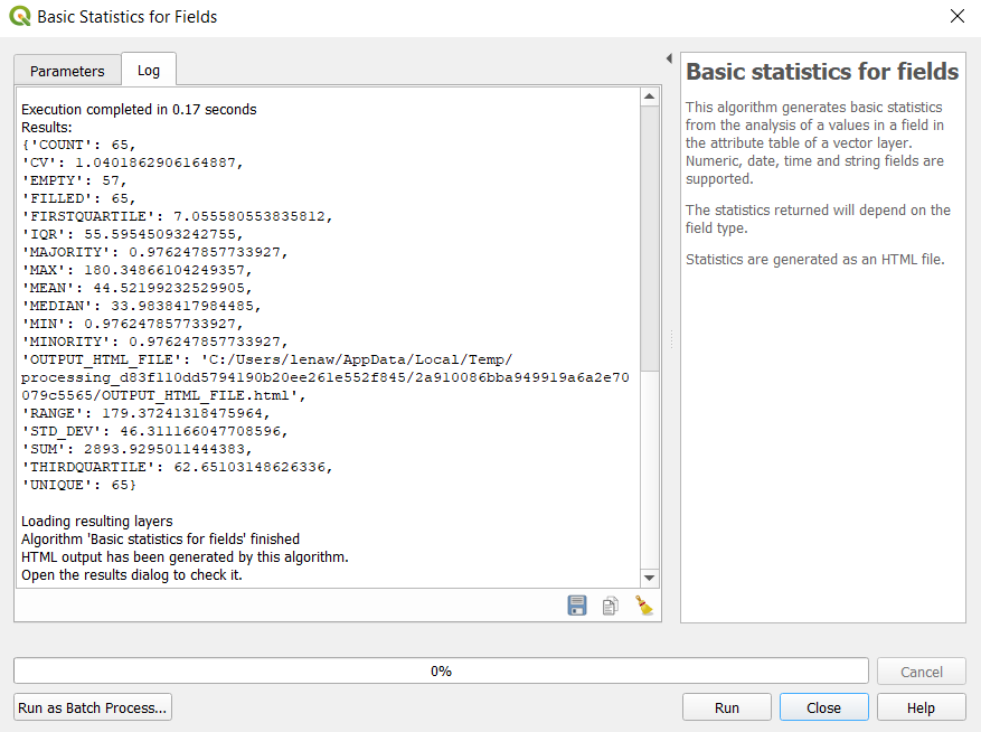
As a good measure for a “valid match” I chose a Levhenstein distance below 1 and a distance between datapoints of a maximum of 200 meters. The resulting table shows 44 attributes that matches these requirements.

Calculated_leven :: Features Total: 122, Filtered: 122, Selected: 44

	Naam	full_id	name	distance	feature_x	feature_y	nearest_x	nearest_y	leven
1	't Kleine Loopv...	w812982275	't Kleine Loopve...	93,6440817351...	120323,410740...	481740,495981...	120231,370963...	481723,236390...	0,0556
2	Albert Neuhuys	w4842719	NULL	3,51249694707...	118475,232423...	485392,628294...	118475,280525...	485396,140462...	NULL
3	Amstelpark	w26516487	Amstelpark	84,1427691272...	121374,306874...	482519,563069...	121353,492489...	482438,035362...	0
4	Amsterdamse B...	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	Apollolaan	w73998931	NULL	7,05558055383...	119911,061182...	484858,348152...	119907,121895...	484852,494672...	NULL
6	Baanakkerspark	w169173148	Baanakkerspark	150,543993819...	124833,682522...	489920,649115...	124960,361528...	490001,987442...	0
7	Beatrixpark	w26037982	Beatrixpark	7,85401391058...	120555,462035...	483963,273041...	120554,430837...	483955,487017...	0
8	Bella Vistapark	r1043467	Bella Vistapark	0,97624785773...	122947,468373...	483466,264070...	122947,903519...	483465,390166...	0
9	Bergwijkpark	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
10	Bijlmerweide	w817343463	Bijlmerweide	80,4863885082...	127636,197502...	481882,308013...	127582,075202...	481941,880115...	0
11	Bilderdijkpark	w220886085	Bilderdijkpark	9,71973783316...	119796,247581...	487242,496487...	119787,341470...	487238,603610...	0
12	Bogortuim	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
13	Bogortuin	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
14	Bogortuin	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
15	Buiksloterbreek...	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
16	Buikslotermeer...	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
17	Darwinplantsoen	w545029841	Darwinplantsoen	10,5484095253...	123899,168315...	484244,490954...	123903,899924...	484235,063291...	0
18	De Bretten	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
19	De Hoftuin	w702630867	Weesperplantso...	36,4031699503...	122113,953870...	486524,717958...	122149,689418...	486531,657804...	1,300

With the given numbers one can measure a percentage of completeness.

44 selected attributes out of 122, result into a completeness of **36%.**



The spatial accuracy as average distance between matches shows a **mean of 44.52**.