Haverhill
Final Deliverable

Juntao Wang U18801732

Yanshan Song U93089867

Yifei Fang U78572119

Tian Ding U90706530

Boston University

May 1st, 2020

Contents

1 Introduction

1.1 Background

The City of Haverhill aims to convert their 311 customer service data (QAlert) to best assist the city government in preparing for first responses to citizen requests.

1.2 Goals

- Merge the city's 311 customer service data (QAlert) with the precinct and ward information in the GIS format to gain further insights on the customers.
- Visualize the refuse routes to optimize the trash collection scheme.

2 Techniques

2.1 Geopandas

GeoPandas is an open source project to make working with geospatial data in python easier. GeoPandas extends the datatypes used by pandas to allow spatial operations on geometric types. Geometric operations are performed by shapely.

2.2 Folium

Folium makes it easy to visualize data that's been manipulated in Python on an interactive leaflet map. It enables both the binding of data to a map for choropleth visualizations as well as passing rich vector/raster/HTML visualizations as markers on the map.

2.3 PyQt5

PyQt is one of the most popular Python bindings for the Qt cross-platform C++ framework. With this tool, we can conveniently develop the GUI software.
The library has a number of built-in tilesets from OpenStreetMap, Mapbox, and Stamen, and supports custom tilesets with Mapbox or Cloudmade API keys. folium supports both Image, Video, GeoJSON and TopoJSON overlays.

3 Data Collection

The client provides us with the geojson format files of the CDBG area, the precincts and wards, the refuse routes of Haverhill as well as a CSV format file storing QAlert requests data.

{"type":"FeatureCollection","crs":{"type":"name","properties":{"name":"EPSG:3857"}},"features":[{"type":"Feature","id":0,"geometry":{"type":"Polygon","coordinates":[[[-7913112.7619000003,5280261.8451000005],[-7913020.8235999998,5280320.373999998],[-7912903.1714999992,5280378.6359999999],[-7912923.3326999992,5280430.3316999972],[-7912945.6876999997,5280472.6560999975],[-7912986.0577000007,5280544.7793999985],[-7913067.5719000008,5280685.2479000017],[-7913090.7513999995,5280719.6996999979],[-7913123.5118000004,5280763.0314000025],[-7913151.6294999998,5280797.4878999963],[-7913211.2866999991,5280868.621799998],[-7913301.2138,5280980.3366999999],[-7913361.4719999991,5281055.0978000015],[-7913392.3465,5281085.3456000015],[-7913414.9129000008,5281103.5009000003],[-7913440.6164999995,5281116.6384999976],[-7913568.4078000002,5281167.0056999996],[-7913602.8720999993,5281188.8017999977],[-7913632.0890999995,5281220.2047000006],[-7913722.0702,5281317.4290999994],[-7914012.1572999991,5281631.2663000003],[-7914028.4696999993,5281621.3077000007],[-7914031.0757999998,5281619.7167000026],[-7914045.4023000002,5281611.0002000034],[-7914060.1609000005,5281602.1093999967],[-7914074.2014000006,5281593.5106000006],[-7914088.0864000004,5281585.0820999965],[-7914102.5512000006,5281576.2885000035],[-7914115.5717999991,5281568.3828999996],[-7914117.0234999992,5281567.501500003],[-7914131.0569000002,5281558.8963999972],[-7914145.6190000009,5281550.0561000034],[-7914173.3357999995,5281533.1367999986],[-7914187.8199000005,5281524.2924000025],[-7914202.2455000002,5281515.5857999995],[-7914216.0994000006,5281507.1014000028],[-7914230.4244999997,5281498.3559999987],[-7914244.2285999991,5281490.0439999998],[-7914258.3627000004,5281481.3136000037],[-7914272.6879999992,5281472.5822999999],[-7914287.3835000005,5281463.7150999978],[-7914301.4748999998,5281455.1501000002],[-7914315.4118000008,5281446.6193000004],[-7914330.1948000006,5281437.6420999989],[-7914344.0314000007,5281429.0869999975],[-7914358.5477000009,5281420.3549999967],[-7914372.8495000005,5281411.7383999974],[-7914386.8518000003,5281403.2028999999],[-7914387.0222999994,5281403.0988000035],[-7914415.1502999999,5281385.8979000002],[-7914415.4047999997,5281385.8387999982],[-7914490.1822999995,5281340.2336999997],[-7914515.5477000009,5281324.063500002],[-7914530.3946000002,5281314.5715000033],[-7914546.1414000001,5281304.381099999],[-7914560.1435000002,5281295.2009999976],[-7914574.1048000008,5281286.0649000034],[-7914588.3588999994,5281276.6394999996],[-7914602.1055999994,5281267.440700002],[-7914616.1119999997,5281258.0182999969],[-7914629.8433999997,5281248.6283999979],[-7914643.6805000007,5281239.0495000035],[-7914657.1838000007,5281229.6264000013],[-7914683.8606000002,5281210.6794999987],[-7914697.3805,5281201.1485999972],[-7914710.8116999995,5281191.3442000002],[-7914724.5004999992,5281181.6005999967],[-7914737.8291999996,5281171.7154000029],[-7914751.4581000004,5281161.7250000015],[-7914764.4124999996,5281152.0394999981],[-7914777.8022000007,5281141.9243000001],[-7914791.0547000002,5281131.8997000009],[-7914804.5868999995,5281121.4513999969],[-7914831.1732000001,5281100.9861999974],[-7914834.4500999991,5281098.8616999984],[-7914855.8425999992,5281081.6516999975],[-7914881.6330999993,5281061.1094999984],[-7914907.6948000006,5281040.0302999988],[-7914933.0174000002,5281019.3047999963],[-7914964.7707000002,5280992.2637000009],[-7915027.3494000006,5280939.0191999972],[-7915036.0936999992,5280930.8122000024],[-7915047.9222999997,5280919.7101999968],[-7915072.2675999999,5280897.6059999987],[-7915097.5195000004,5280874.3526000008],[-7915121.5558000002,5280851.8452000022],[-7915145.8425999992,5280828.8078000024],[-7915170.1245000008,5280805.5227999985],[-7915193.3539000005,5280782.8338999972],[-7915217.5585999992,5280758.8150999993],[-7915240.8206999991,5280735.1815000027],[-7915264.2322000004,5280710.9804000035],[-7915287.0094000008,5280686.9587000012],[-7915310.1553000007,5280662.2173999995],[-7915332.4958999995,5280637.9331],[-7915355.3290999997,5280612.8660999984],[-7915377.6453000009,5280588.072300002],[-7915388.9933000002,5280575.4196000025],[-7915400.0248000007,5280563.1196999997],[-7915422.5187999997,5280538.054899998],[-7915444.7017000001,5280513.5864000022],[-7915459.6458000001,5280498.0947000012],[-7915493.7233000007,5280459.250500001],[-7915510.6393999998,5280440.3649000004],[-7915527.0956999995,5280422.6982000023],[-7915495.8092,5280381.8191],[-7915425.2278000005,5280289.5964000002],[-7915365.2597000003,5280211.4416999966],[-7915347.7011999991,5280195.8444999978],[-7915343.4338097479,5280191.9754717937],[-7915338.7710123155,5280188.5934134107],[-7915333.7681044964,5280185.7384331729],[-7915328.4844165044,5280183.4443887072],[-7915322.9826083668,5280181.7384854211],[-7915317.3279268285,5280180.6409538714],[-7915311.5874315826,5280180.1648098445],[-7915305.8291999996,5280180.315700002],[-7915092.4795999993,5280159.1468999982],[-7914907.7259999998,5280113.9763000011],[-7914739.6622000001,5280094.1177000031],[-7914561.9832000006,5279980.0781000033],[-7914514.0734999999,5279935.4129000008],[-7914460.2736000009,5279861.8407000005],[-7914315.1418999992,5279603.699000001],[-7914277.6775000002,5279558.4531000033],[-7914158.7195999995,5279393.0319999978],[-7914107.3155000005,5279268.6532000005],[-7914006.3696999997,5279003.3206999972],[-7914007.3535999991,5279002.9821999967],[-7914123.6817000005,5278962.9632999972],[-7914181.9087000005,5278942.0894000009],[-7914232.8388,5278923.

Figure 1: CDBG Json Format

K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
Request T	Request T	Request T	Request T	Departme	Departme	Address N	Street ID	Street	Address	Complete	City ID	City	Cross Stre	Cross Stre	District ID	District	Escalation	Priority	Longitude	Latitude
330	Traffic/Str	FALSE		11	Highway		629	JUSTIN ST	JUSTIN ST	JUSTIN ST,	1	Haverhill	0		0		0	2	-71.1059	42.76755
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
321	Snow Rem	FALSE		11	Highway		922	SOLITAIRE	SOLITAIRE	SOLITAIRE	1	Haverhill	0		0		0	2	-71.0737	42.80426
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
330	Traffic/Str	FALSE		11	Highway		659	LAKEVIEW	LAKEVIEW	LAKEVIEW	1	Haverhill	777	NORTH AV	0		0	2	-71.0834	42.796
302	Poor Prop	FALSE	Please de	6	Inspectior	44	611	JACKSON	44 JACKSC	44 JACKSC	1	Haverhill	0		0		0	2	-71.0915	42.77307
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
291	Highway -	FALSE	<p>Pleas	11	Highway	114- 116	308	BOARDMA	114- 116 B	114- 116 B	1	Haverhill	0		0		0	2	-71.0617	42.7742
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
343	Trash - Ent	FALSE	Please de	6	Inspectior	2	1153	H Street	2 H Street	2 H Street	1	Haverhill	0		12	Polygon 0	0	2	-71.1085	42.74725
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
274	School De	FALSE		0		685	1036	WASHING	685 WASH	685 WASH	1	Haverhill	0		0		0	2	-71.1032	42.77021
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
901	A - Inform	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0
902	A - Transfe	FALSE		8	311 Call Center		0			, Haverhill	1	Haverhill	0		0		0	2	0	0

Figure 2: Part of Request CSV File

4 Data Processing

4.1 Draw the CDBG area

(1) Use "geopandas" package of Python to read the "Hav_CDBG_Area_WGS84.json" file and then transform the coordinates stored in it from "espg:3857" format to "epsg:4326" format.

(2) Utilize "folium" package of Python to draw the boundaries of the CDBG area and set up the background color.

(3) Generate a HTML file and store the result in the file.

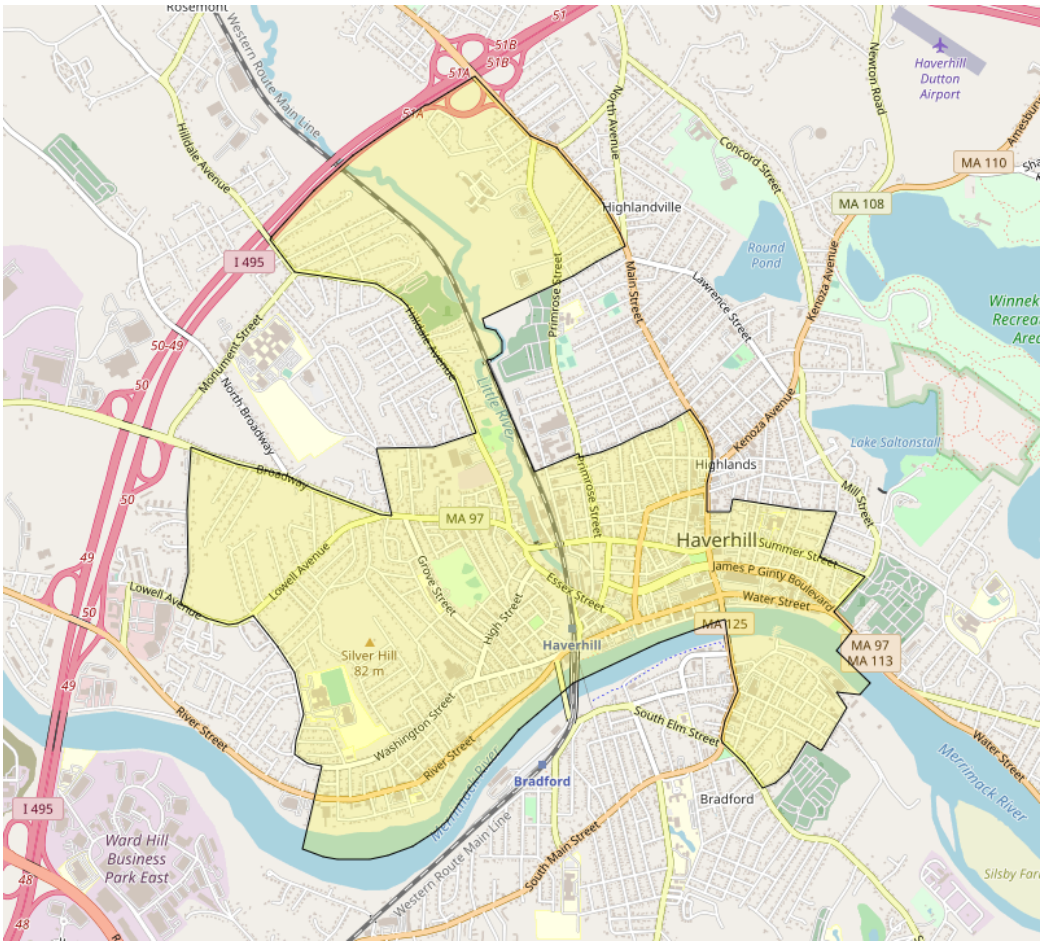

Figure 3: CDBG Area

4.2 Draw the precincts and wards

(1) Remove the data entries representing the old geological data of 2003 in the file.

(2) Use "geopandas" package of Python to read the "Hav_Precincts_Wards_WGS84.json" file and then transform the coordinates stored in it from "espg:3857" format to "epsg:4326" format.

(3) Utilized "folium" package of Python to draw the boundaries of each precinct respectively.

(4) Generate a HTML file and store the result in the file.

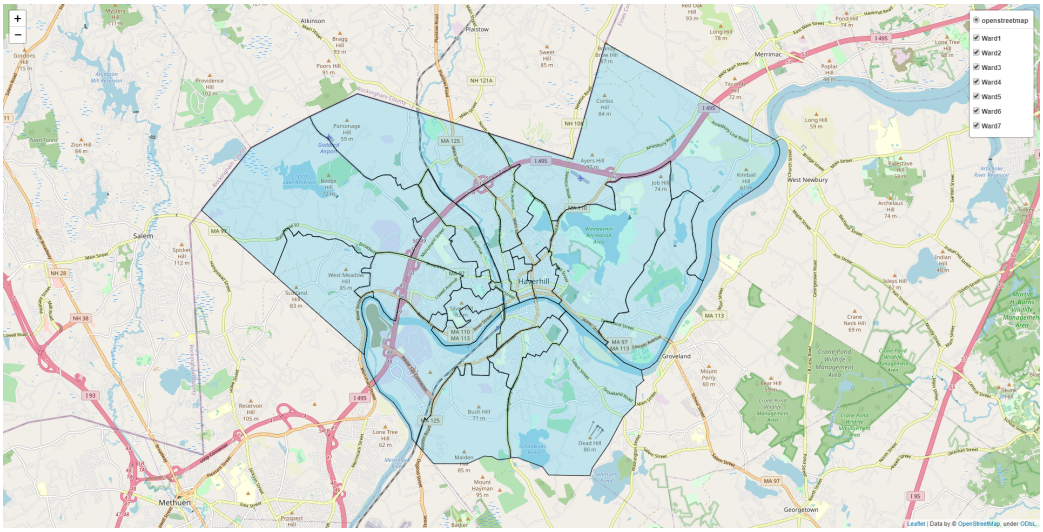

Figure 4: Wards and Precincts

You can see the layer control on the side, so that you can choose to show or hide some of the wards as below.

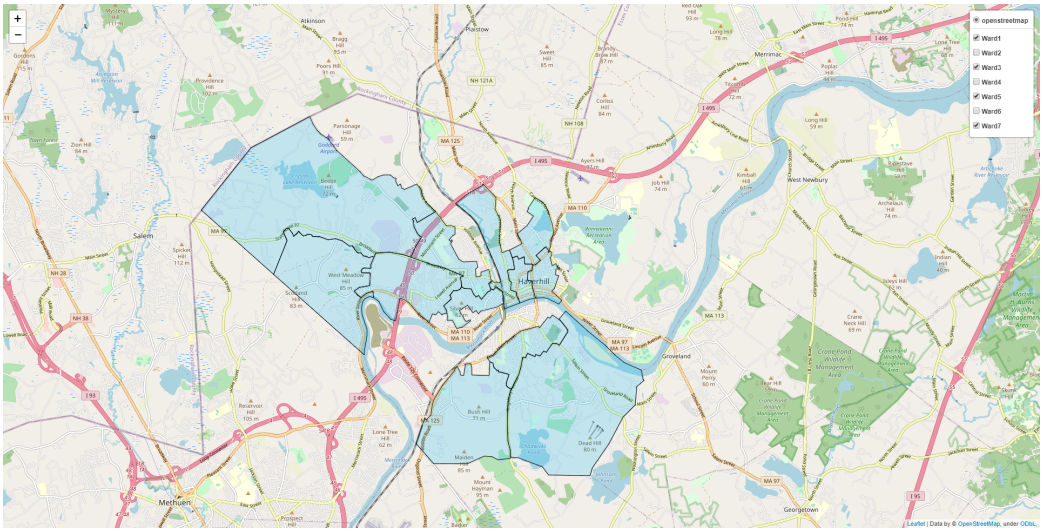

Figure 5: Selected Wards and Precincts

4.3 Map the QAlert requests data(CSV format)

(1) Use the "pandas" package of Python to read the "haverhill-request_updated.csv" file into the python to process the data.

(2) Remove all the entries that do not have address or their latitude or longitude equal to 0. Then we can get the valid data.

(3) Remove all the entries that contain back quote in its address, because the back quote will lead to error when we map the requests.

(4) Traverse the file and get all types of requests.

(5) Group the coordinates of all requests by their request types.

(6) Use "folium" package in Python to map the requests and generate the HTML file.

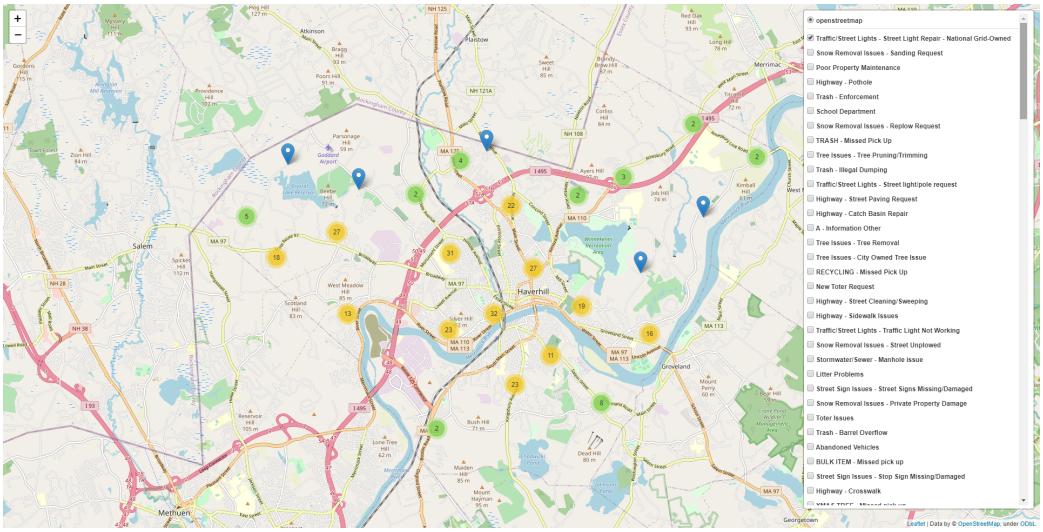

Figure 6: Requests on the map

Also, you can choose to show or hide some types of requests. What's more, you could click on the marker clusters, then zoom in the map. The marker clusters will become separate markers. When you click on the marker, you will see some detailed information of the marker like address and request type as below.

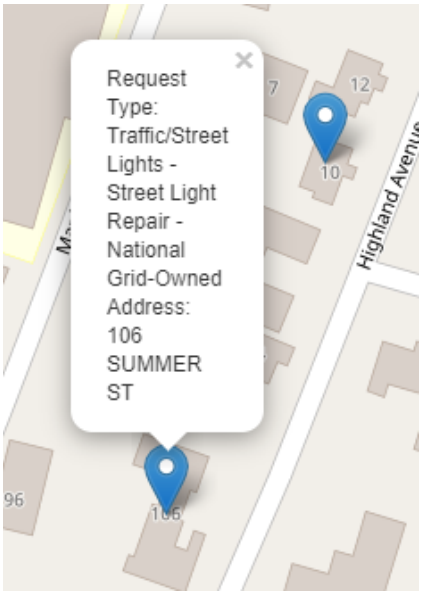

Figure 7: Marker Information

4.4 Map the QAlert data in the CDBG area

We combine the QAlert data and the GIS information of CDBG area. This image shows the combination. Each blue marker in the map indicates a request. When the mouse clicks the marker, the detailed information of the requests will show. The green cycle in the map represents a cluster of requests in the neighbouring area. When we zoom in the map, the green cycle will be scattered into several blue Markers. The right part of the image shows the console of the map. We categorize the requests by the departments, so with the help of the console, we can filter the requests by departments.

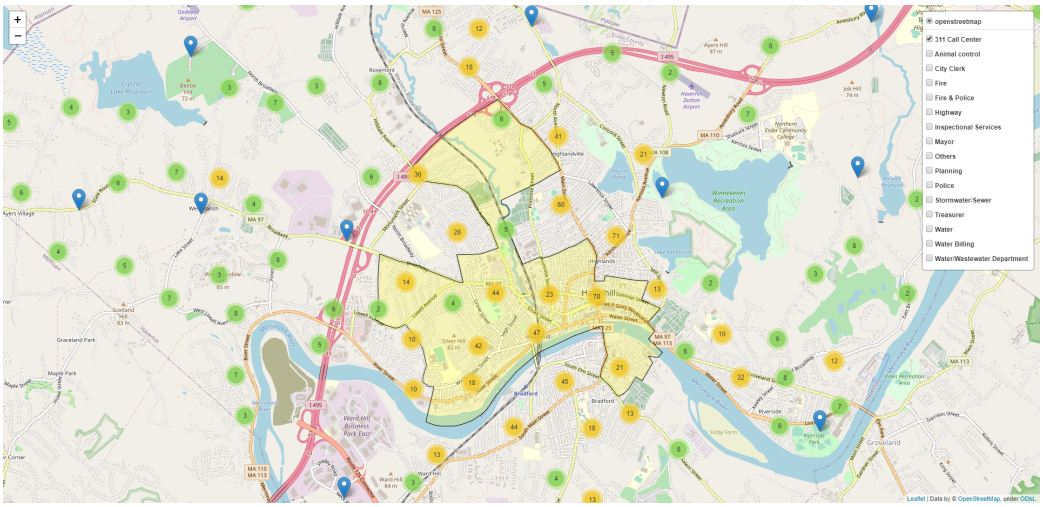

Figure 8: Map the requests on the CDBG area

4.5 Map the QAlert data in the precincts and wards

This image shows the combination the the QAlert data and the GIS information of the precincts and wards. We use the same strategy to compose the map.

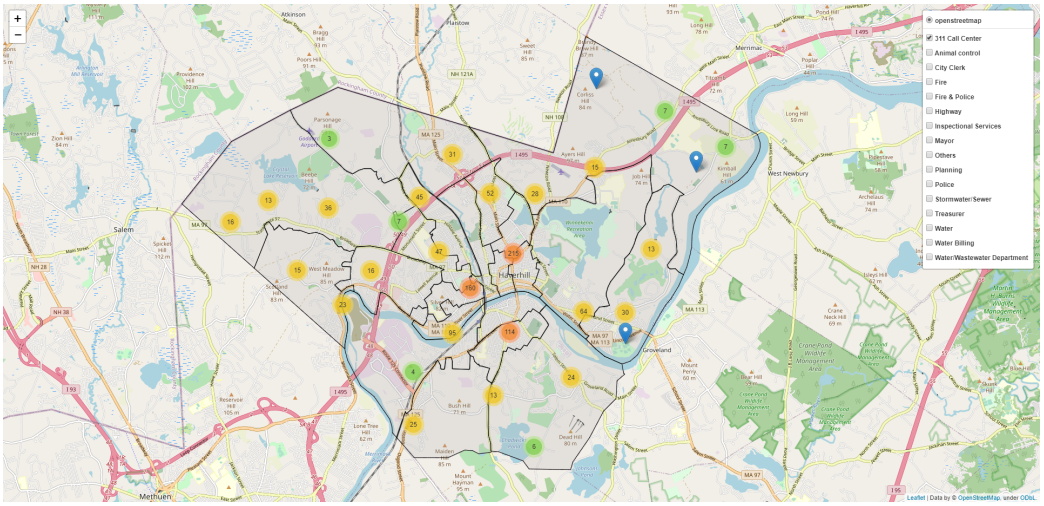

Figure 9: Map the requests on the Wards and Precincts

4.6 Visualize the refuse routes

We also use geopandas and folium to visualize the refuse routes. The image is the visualization of the trash routes. The red curves represent the routes executed in the red week, the blue curves represent the routes executed in the blue week and the black curve represents every friday's route.

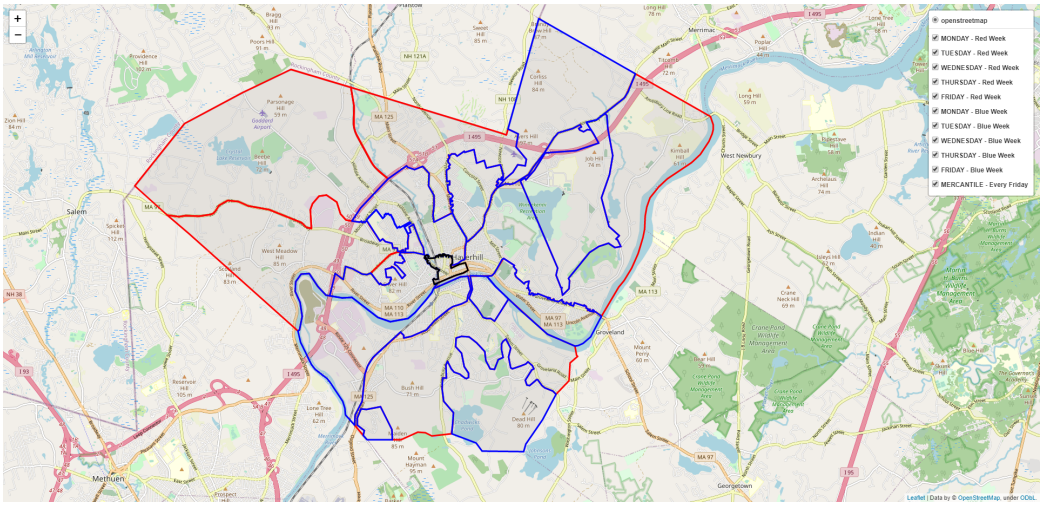

Figure 10: Refuse routes

4.7 Design and implement a GUI software

We also have designed and implemented a GUI software that can help the client conveniently get the above results with the new dataset.

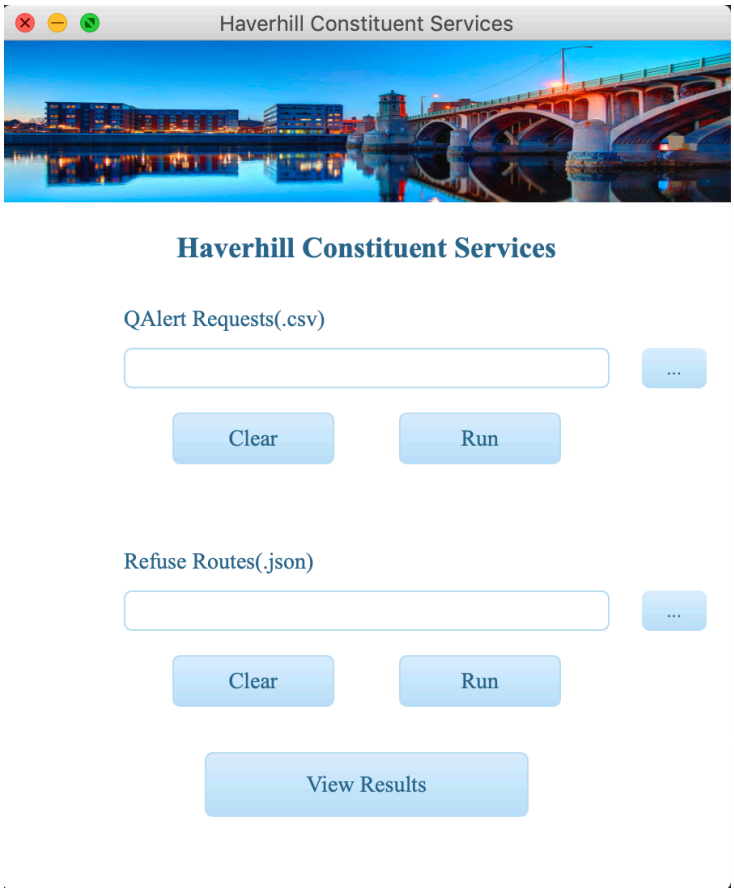

Figure 11: Software interface

Here are some operating instructions:
(1) Click the tool button and select the input QAlert csv file.

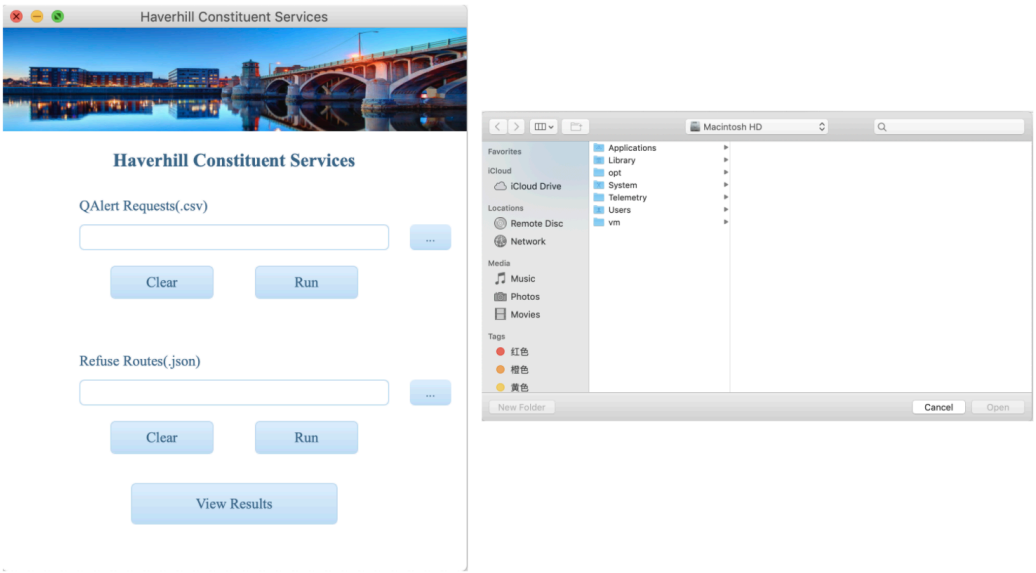

Figure 12: Software interface

(2) Then click the ”run” button, the software will automatically process the input file and generate the target HTML files.
(3) Click the ”View Results” button and view the generated HTML files.

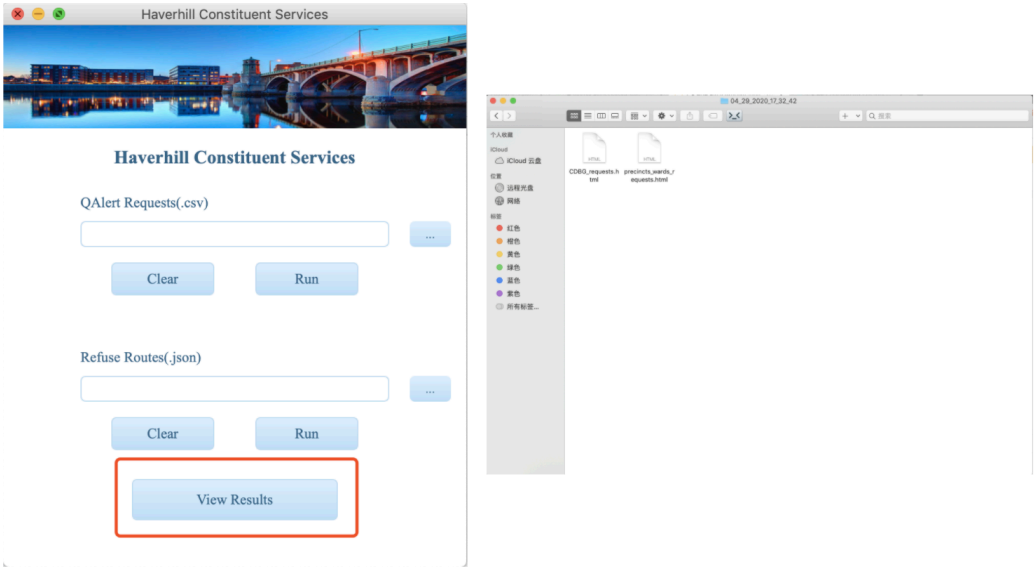

Figure 13: Software interface

5 Limitations

- When we try to package our python files to an exe program, there are something wrong with the package so the GUI software can only run in the python environment temporarily.
- If the structure of the QAlert database changes, the result of our program might be wrong.

6 Conclusion

- We reach the goals of the project.
- We successfully merge the city's 311 customer service data (QAlert) with the precinct and ward information in the GIS format to gain further insights on the customers.
- We visualize the refuse routes to optimize the trash collection scheme.
- We design and implemented a GUI software that can help the client conveniently get the above results with the new dataset.