Deliverable 1

1.Current Work

1.1. Data Description: We have four datasets as shown below

a. Haverhill-request.csv

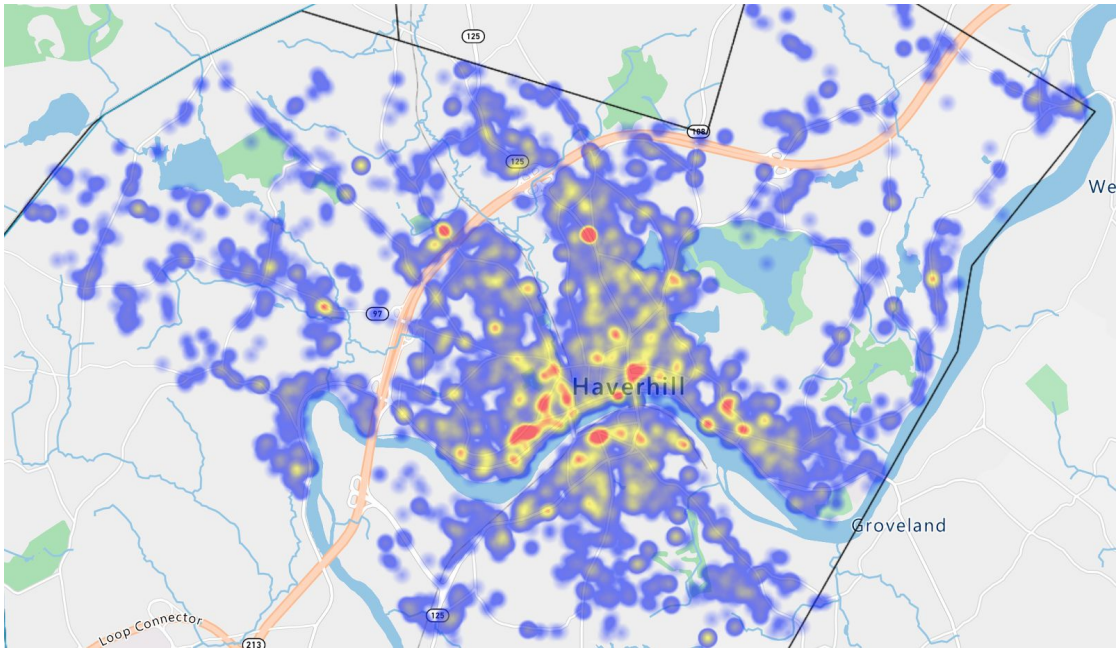
This csv file contains the specific descriptions of every request for the past two years, including Request ID, Creating Date, Address, Longitude, Latitude, and etc. However, nearly 70% of the database information is useless for our project. First, the requests without specific addresses can be removed. Besides, if the **Request Type** field = Information Other, or Transfer Other, these requests can be removed, since they are just some trivial information requests, e.g. what day is Easter, when is City Hall open, etc.

b. Geojson

Geojson is a JSON based geospatial data exchange format. It defines several types of JSON objects and their combination methods to represent data about geographic elements, attributes and their spatial scope.

We got three datasets this time: one is Hav_Precincts_Wards.json that contains the coordinates of all wards, which is in "Polygon" type, in Haverhill. By using this dataset, we can draw the area of each ward. One is Hav_CDBG_Area.json that contains the coordinates of Community Development Block Grant which is the center of Haverhill. We can focus on this specific area with this dataset. Another is Hav_Refuse_Routes.json that records the garbage trucks.

2. Analysis proposal



The above heat map is the preliminary analysis we have performed. We have incorporated the most recent data from Haverhill's 3-1-1 call system, and computed the density of the effective calls made. The examination of this heatmap will serve as a practice for us to understand the distribution of the 311 calls, because we will research in-depth the precincts and wards in Haverhill, and perform similar analysis on each of these sub-sections. Furthermore, we are exploring the technique of creating heatmaps, and will incorporate skills in R to make heatmaps in a more clear and professional manner in the future. We will also make use of the types of calls made, and create heatmaps separately for easier visualization for the decision makers to divide labor accordingly to our visual analysis. Finally, we will create a master heatmap, which has distinct colors for each type of data processed, with the wards and precincts clearly drawn out on the map, to showcase the processing of data, which hopefully will help the government of Haverhill respond to the 3-1-1 calls faster in a cost-efficient way.

3. Future Work

- Our current priority is to perform data cleaning on the most recent version of the 3-1-1 requests CSV file. From our preliminary examination, this CSV file contains more than 150,000 data entries, with 33 attributes for each data entry, causing this a great obstacle for us to divide the types of data into categories, and further manipulate them within each wards / precincts. Therefore, we will be using coding exercises introduced by Professor Galletti, and clean data with the description above, shown in the "Current Work" section.
- Draw the boundaries in the specific region in order to have a clear perspective and plot different wards/precincts with different colors. Then, merge these regions into one map.
- Mark requests on the map to make clients see requests clearly.