**311 Call Data Analysis High Level Overview**

Concept

Our 311 analysis is based on a dataset of 311 call metadata available from the NYC OpenData portal. Currently we use it to look for the prevalence of quality of life indicators: (1) Trash complaints (2) Noise complaints (3) Graffiti complaints. We use this to ascertain whether multifamily properties are in particularly pleasant areas to live and if so to

The intention is to expand this to other types of complaints at some point in the future.

Generating output

The output from the R scripts is a csv file which shows the absolute numbers of calls and proportions of calls relative to (1) population (2) residential units for a variety of geographic areas. These areas include neighborhoods, boroughs, city and 1/8 mile and ¼ mile radii of the property or properties.

The final output is then created in excel. This is done by first creating a pivot table which can isolate a time frame of multiple years. The data in the pivot table is then copy/pasted into another sheet and the results are manually inspected to find which relative proportions of complaints look best compared to other areas of interest. As with all charts created in excel DO NOT copy/paste the chart from excel into a powerpoint. Instead create a chart in powerpoint and then paste the data from the excel file into the powerpoint chart. This avoids the problem of broken links.

Use of script

The dataset contains complaints registered to multiple agencies some of which overlap. Trash, noise and graffiti complaints are each a combination of different complaint codes from different agencies. These are included in the script but can also be ascertained by using a combination of regex statements and data dictionaries which can be obtained from the open data portal.

The relative proportions are created by using a script which utilizes geospatial methods to determine the number of residents and number of residential units within given areas. The script incorporates population changes from year to year. Currently the script only utilizes the most recent count of residential units. This can be fairly easily altered by using the time series pluto rds file.

Use in modeling

It should be noted that there’s a second (well grounded) theory on the volume of 311 complaints. As new residents move in they tend to complain more loudly to the city about issues previous residents had come to terms with. A variety of external analyses seem to indicate this effect. Because it easily geolocated, simple to use and extends back as far as our DOF data it has the potential to provide multiple features for project black box.