Ben Dodson
16 December 2019
MUSA 620
Prof. N Hand

What the Fine? *A Python Dashboard of Parking Violations and Fines in Philadelphia*

SUMMARY

Coming back to your car only to discover a blue and white envelope affixed to the windshield can be maddening. With fines reaching up to $301 in Philadelphia, it can also be costly. Knowing where parking violations are common and how much they will cost has implications for individuals, developers, and cities. The City of Philadelphia maintains a database of all parking violations that occurred between 2012 and 2017. In its entirety, this dataset consisting of over 9 million entries, averaging about 1.5 million violations per year. With a dataset of this size, most data visualization techniques are unable to process the data without crashing. The solution is to aggregate the data, in this case, temporally by year, and spatially on the census tract level. Additionally, as different uses of the data may be required, instead of providing the data in one calculation, the violations are aggregated into four different operations: mean fine, median fine, total number of violations, and total number of violations normalized to the size of tract. The product is a web-based dashboard which allows users to filter the data by year and view interactive maps of each of the four aggregations. Users can explore how fines are higher and more frequent and some areas, and more dispersed in other areas.
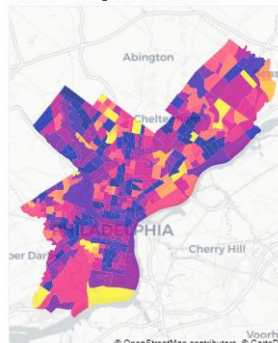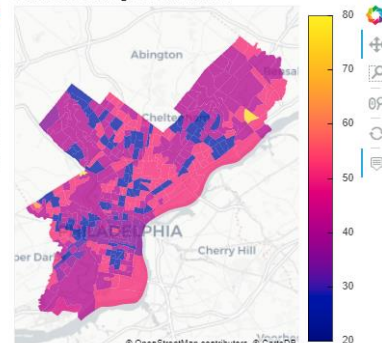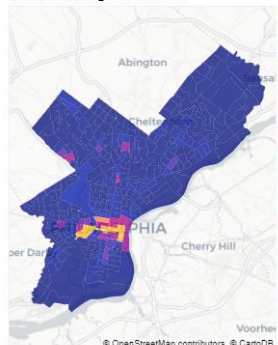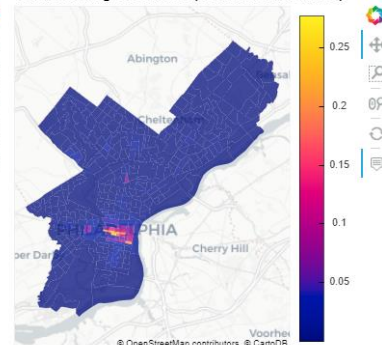
METHODS

The parking violations database contains over 9 million entries from 2012 to 2017. This is too large to visualize with most techniques (namely osmnx) while still providing information for the entire city. Even filtering the data into Center City or by a single year is still too large. The most effective way to aggregate the data is by census tract.

First, the app queries the parking violations data from the parking violations database through the Carto API and immediately converted into a geodataframe using the City of Philadelphia's carto2gpd package. The query removes any entries for which the fine is higher than the maximum fine on the Philadelphia Parking Authority's [fine schedule](#). The query also only brings in only the necessary columns to keep the dataset as small as possible. Using the datetime package, the year is extracted from the issue_datetime column so that the data can be categorized as years. Census tract geometries are also hosted by the City of Philadelphia and brought in using the ArcGIS GeoService API. The violations are joined to the tracts data so that there is a tract number associated with each of the parking violations entries. Then the parking violations data is grouped by tracts and the mean, median, and number of fines for each tract is calculated. Finally, the new dataset is rejoined so there is a polygon associated with each tract. An additional column is created which is total number of fines divided by the land area of the census tract to allow for the total number of violations to be normalized to the size of the tract. For compatibility with the app, the melt function is used so that for each entry there is a tract number, tract geometry, variable name (mean fine, median fine, total violations, or total violations (normalized)) and the corresponding value. The data is exported to a geojson file where it will be reused in the app.

The app pulls in the geojson file (hosted in the GitHub repo) as the dataset for the app. The app is based upon the filtering of code based on the year (a variable determined by the user). This is followed by code for four different plots, one for each aggregation. These plots are created using hvplot's built-in geographical plotting capabilities. For each plot, the data is filtered by the aggregation and user-selected year, then plotted on a map. All four maps are shown side by side. hvplot's robust capabilities allow the apps to be interconnected. Interacting with one plot will be tessellated across the other three plots. All plots are tied to the year slider (built using the param package). Anytime the slider is adjusted, the maps are plots are redrawn with that year's data.

The app comes together in a dashboard built on the panel package. A logo, title, and description are in the first column, and the four plots are spread on the second column. The entire app is made able to be hosted in a server due to the capabilities of the panel platform. The app is hosted using binder.

DISCUSSION

The mean fines for census tracts across Philadelphia are see less variation in Center City and University City than in the less urban areas of the city. However, the fines are not necessary less in some tracts as many tracts in North and Northeast Philadelphia can see higher fines than in Philadelphia. Using the median fine aggregation, the data fluctuates less with a hotspot in the Rittenhouse Square area.
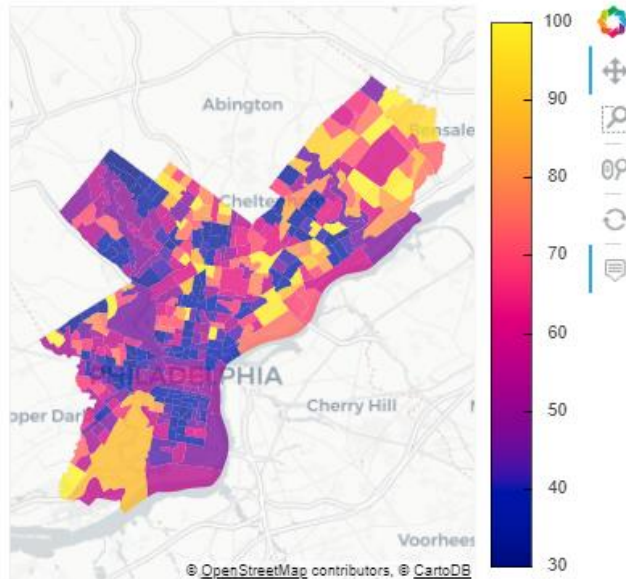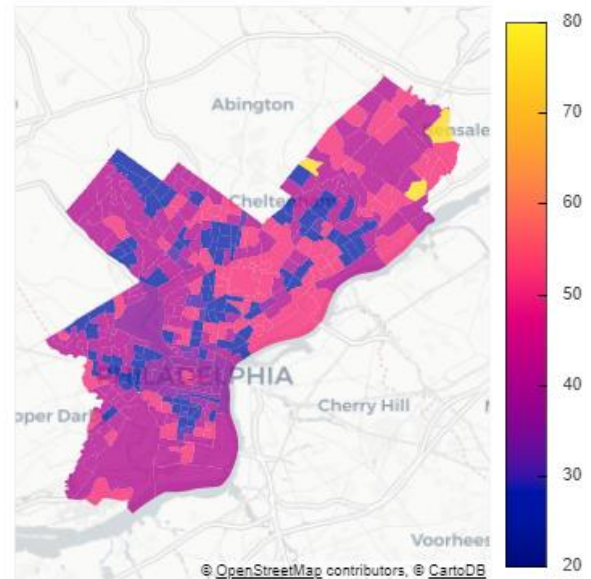

Just looking at the total number of parking violations, there is definitely a hotspot for parking violations in Center City and University City. Even when the total number of violations is normalized to the size of the census tract, a hotpsot still remains. In 2017, in the area around Market St, west of Broad St, there were over 71,000 parking violations with a median fine of $41.
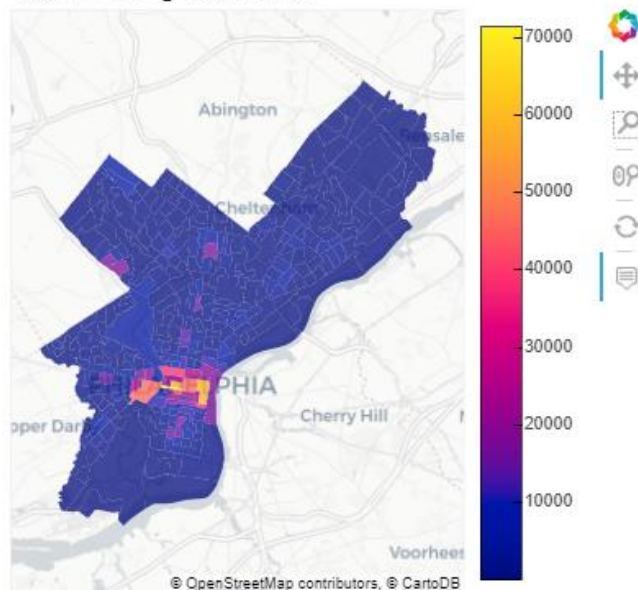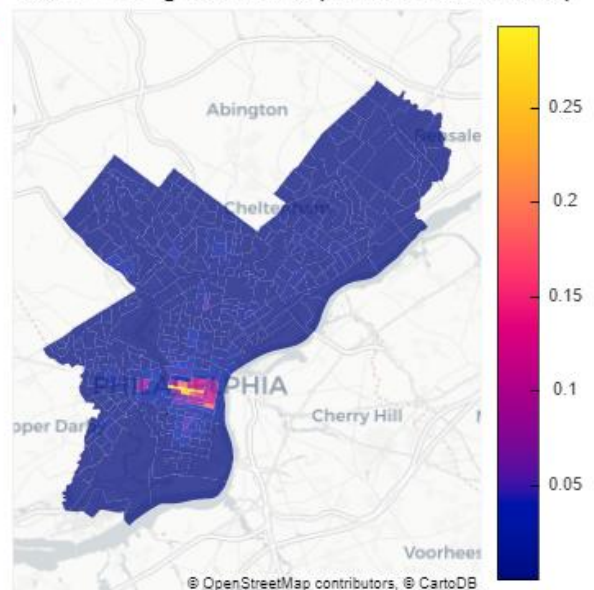
This dashboard could be useful for developers to understand where parking violations are high. In areas where parking violations are high, the cost of private parking may be too high or the availability too low that it is causing a high number of people to illegally park in public spaces and risk a fine. It could also be useful to the Philadelphia Parking Authority which does most of the parking enforcement to see where to deploy parking supervisors. Knowing where most of the violations occur, and the average violation price would allow parking supervisors to be redistributed to areas where parking violation revenue could potentially be the highest. If done correctly and transparently, could create for more equitable assessment of parking violations because the higher the violation, the more danger is being caused by a vehicle being present. For example, the highest fine assessed is $301 for counterfeit handicap placards.

Other data visualization packages may have proved helpful to use with this data but due to the sheer size of the data, a package like osmnx cannot handle this many data entries. Large cities, like Philadelphia, with a large amount of data, would benefit from the development of packages that are able to process datasets with several million rows.

Once parking violations data from 2018 is released, by simply running the data code again, new data for the app will be generated.