Data Analysis for Lucas Cioffi (Greenburgh, NY 2019 Town Supervisor Election)

Table of Contents

Introduction **1**

Visualizing Police (Traffic Accidents and Larceny) Data **1**

Figure 1: Overall Map of Vehicle Accidents and Larcenies 2

Figure 2: Vehicle Larceny (Potential "Hot Spots") 3

Figure 3: Traffic Accident Data 4

Next Steps **4**

Paul Feiner's Campaign Finance Data **5**

Figure 4: Paul Feiner's Campaign Financial Contributions 5

Table 1: Distribution of Feiner's Contributions by State 6

Voting Trend Data **8**

Introduction

In this report, I[1] will keep track of summaries of analysis as well as ongoing work. The main data analytics initiatives are:

1. Visualizing Police (Traffic Accidents and Larceny) Data
2. Analyzing Paul Feiner's Campaign Finance Data
3. Analyzing Voting Trend Data

Visualizing Police (Traffic Accidents and Larceny) Data

Goal: Make police data more transparent and accessible for the residents of Greenburgh. For example, plotting traffic accidents and larceny data on a map and creating other interesting visualizations could raise interest in the campaign and increase demand for even more data transparency.

I plotted Traffic Accidents and Larceny data using Google My Maps. I made the map public for anyone with a link to view, so if you click on the link, you can interact with it a bit by zooming

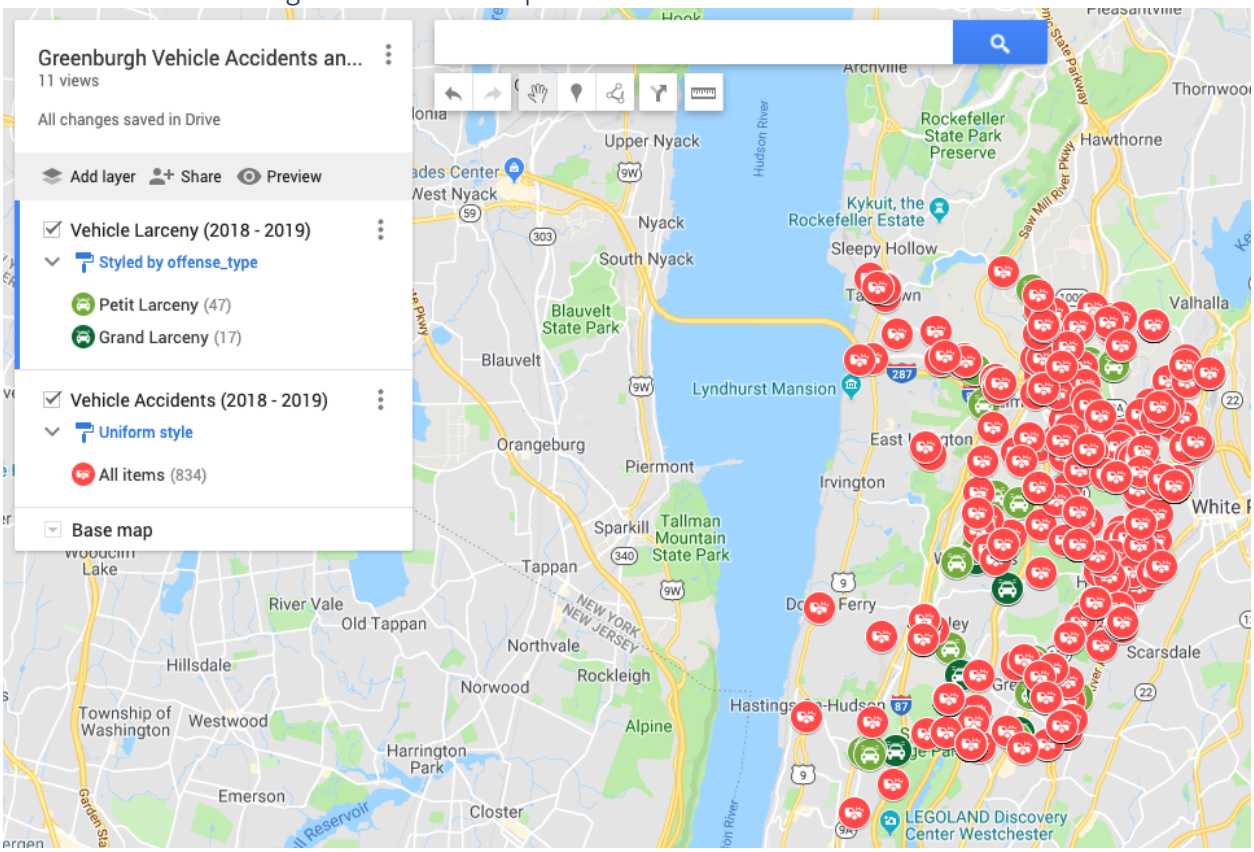
[1] Author: Vincent La; Email: vincela14@gmail.com

in/out, looking into the data table, filtering, etc. Note that if interested, this map can be embedded on to a website:

```
<iframe
src="https://www.google.com/maps/d/u/0/embed?mid=1NJjN31BTQHaQD2
Cbp50NgZ8TSQd5Hc-8" width="640" height="480"></iframe>
```

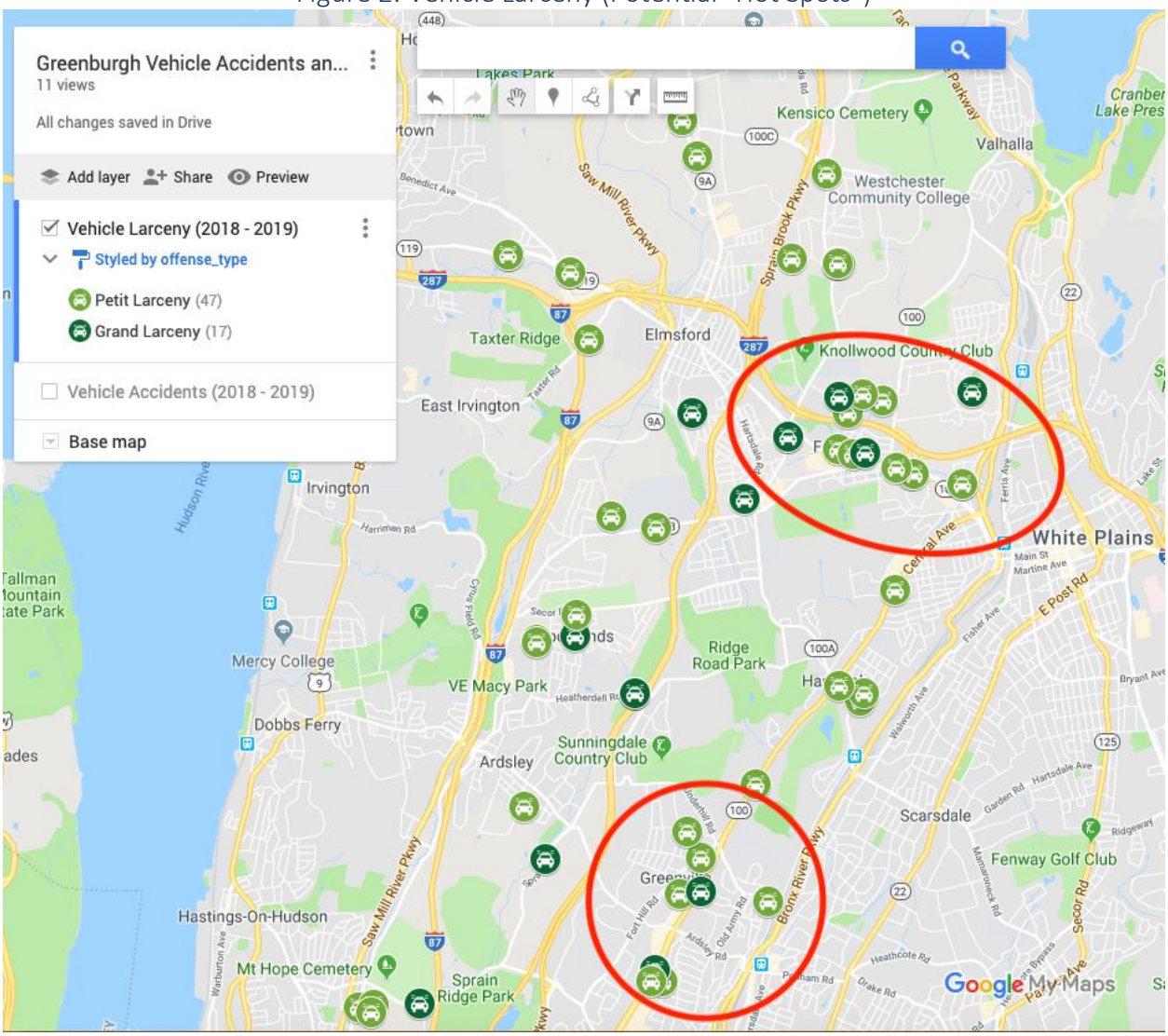
Attached here are some screenshots that seemed interesting.

Figure 1: Overall Map of Vehicle Accidents and Larcenies



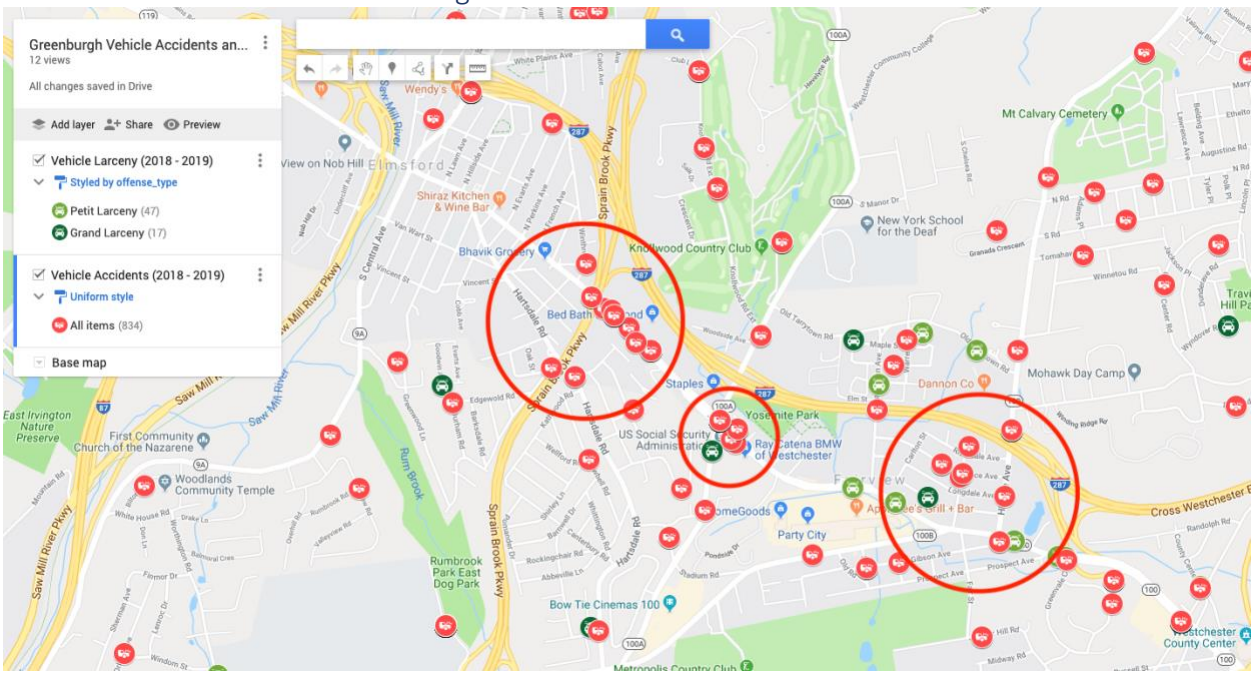
Zooming in a bit, just eyeballing the map there seem to be a couple “hotspots” for larcenies. One area that stood out to me was Fairview along Tarrytown Road. The other was Greenville along Central Park Ave.

Figure 2: Vehicle Larceny (Potential "Hot Spots")



Adding back in Vehicle Accidents data, the general patterns seem to be that accidents are more frequently located along major roads and intersections with freeways. This generally makes sense to me, and I'm curious if you have other thoughts on other interesting insights you can take away from this data?

Figure 3: Traffic Accident Data



Next Steps

1. I'm curious to see what your thoughts are here and if this is what you were going for? The Google Maps product is super light weight so it was relatively easy to put this map up after cleaning the data and geocoding it. However, it's pretty minimal in terms of features so more complex designs would probably require actually developing code and building a visualization. I did make the map public and there seems to be an option to embed it on a website so let me know if this works for you?
2. If we were to want to develop this more, there's a few additional features that I think could be interesting:
    a. Dashboard aggregating the data in a particular geographic area. In other words could have a summary table that aggregates the number of accidents and larcenies in a particular area based on the current zoom level of the user.
    b. A report that identifies particularly problematic areas that could be used to generate policy proposals and solutions. Currently, I just eyeballed and highlighted a couple areas that stood out to me. With more context on the ground, maybe we can also use this map to identify fixes to improve road safety?
3. Note that I was only able to geocode **about 50% of the traffic accidents** data set. Since the traffic accidents data set only includes the "Roadway" (and additionally the Roadway/Intersection only in some cases), the location in the report is very imprecise. Thus, the geocoder that I ran this through only picked up latitude/longitudes for about 50% of the reports. If you could get the police department to provide the State Plane Coordinate System (SCPS) for the traffic accidents data as they did for the larcenies data set, that could be helpful for a more complete and precise visualization.

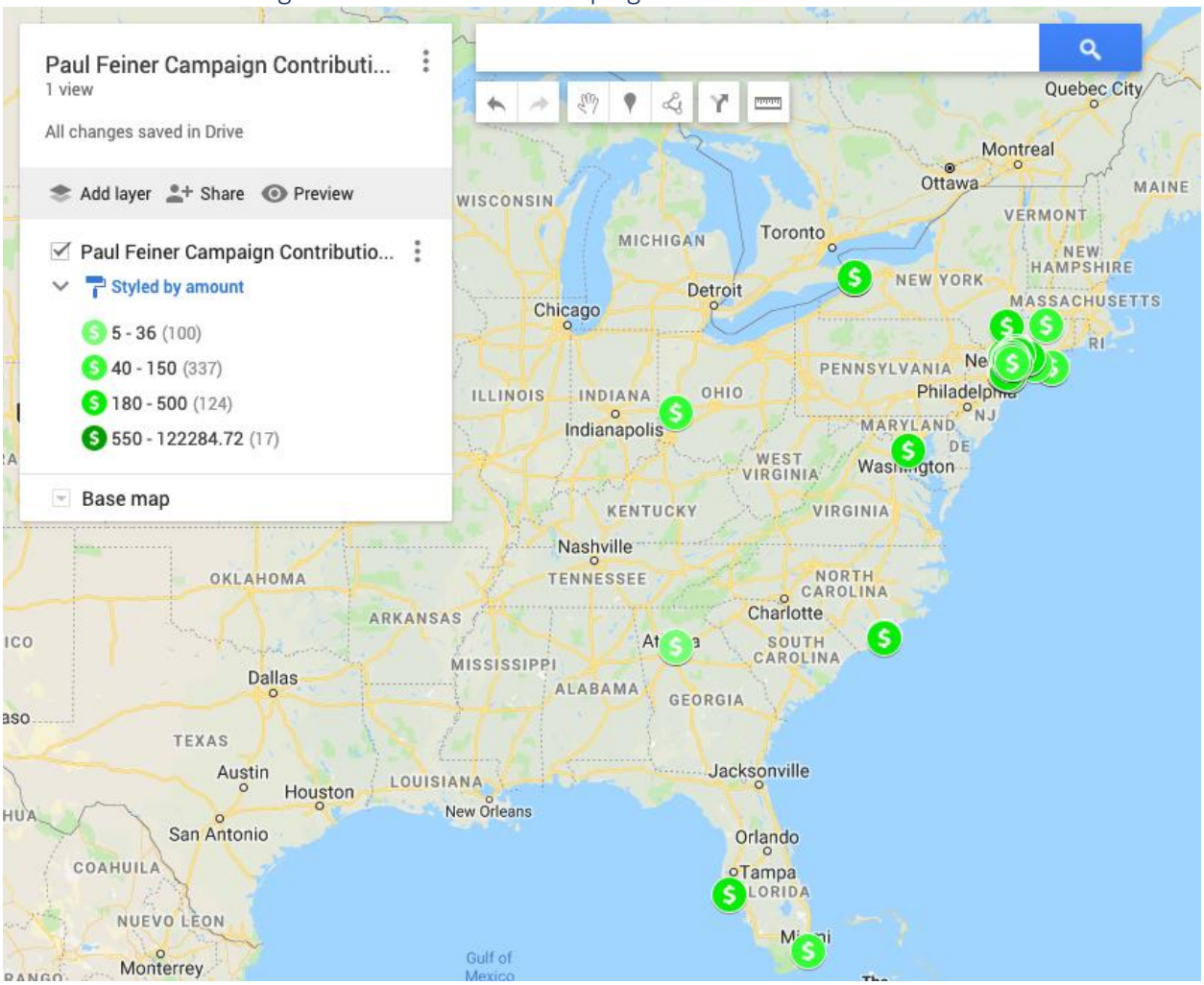
Paul Feiner's Campaign Finance Data

Goal: Analyze Paul Feiner's campaign finance data. Uncover any potential unscrupulous behaviors (e.g. cross reference against list of individuals who cannot donate due to conflicts of interest).

Similar to the police data above, I plotted Mr. Feiner's campaign finance data on Google Maps. Note, I did not make this map public since it seemed like this visualization was more for your team's internal use than for the public. However, I could make it public if you thought it was interesting.

Note, the map below is incomplete. To geocode the campaign finance contributions, I used OpenStreetMaps. Due to rate limiting (and I'm just using a free account), I think there's a limit to how many things I can geocode at a certain time. I got through about 40% of the data that you sent over so this isn't complete, but I think good enough to draw interesting insights.

Figure 4: Paul Feiner's Campaign Financial Contributions



Just looking at an overview of Mr. Feiner's campaign contributions, most of the contributions are in fact located in the New York area. However, there are a few contributions from out of state from all over the eastern part of the US.

Table 1: Distribution of Feiner's Contributions by State

State	Percentage of Contributions
NY	95%
Non-NY	5%

Zooming into New York specifically which is where most of the action takes place, most of the contributions are from Greenburgh, NY (roughly 75%). Other contributions are coming from surrounding areas like White Plains and NYC.

Figure 5: Contributions in NY State

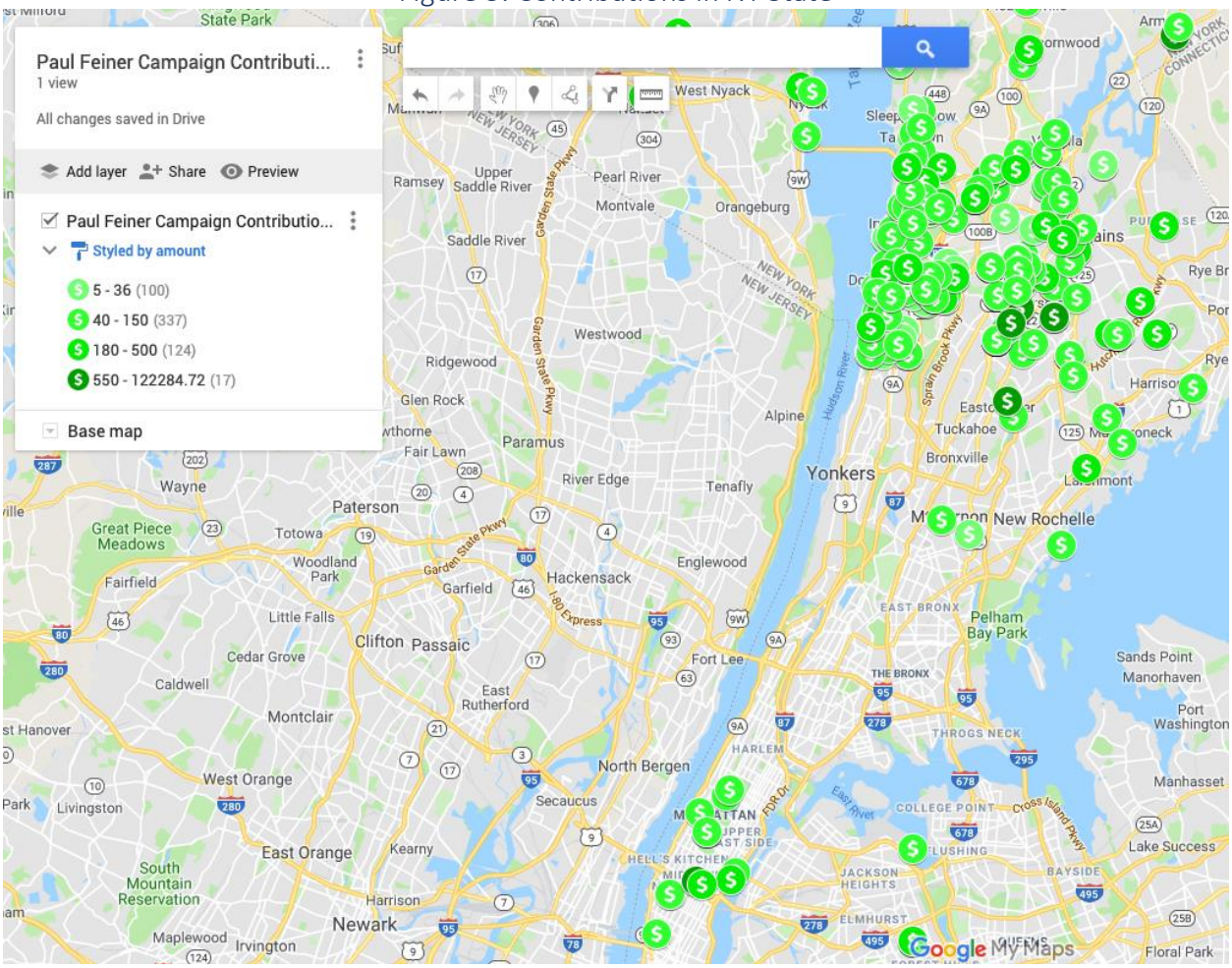
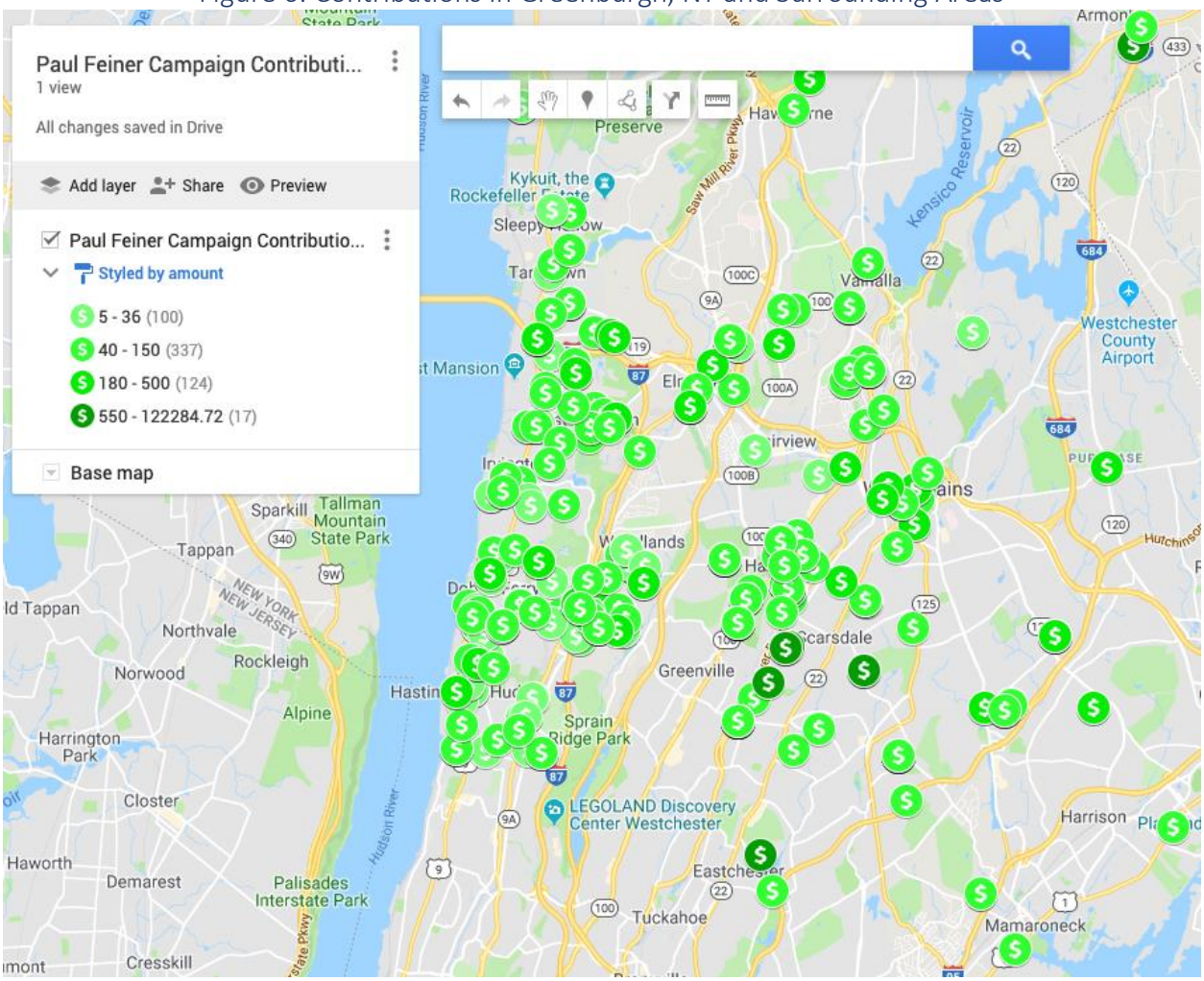


Figure 6: Contributions in Greenburgh, NY and Surrounding Areas



In the next table, I look at the distribution of contributions (weighting each contribution equally not by dollars) by zip code.

- 75% of contributions are coming from zip codes within Greenburgh, NY.
- Among the 25% of contributions outside of Greenburgh, NY, I notice mostly surrounding cities within Westchester county (White Plains, Scarsdale). Some are from NYC and some are from out of state.

Table 2: Distribution of Contributions by Zip Codes

Zip Code	Percentage of Contributions
Within Greenburgh, NY	**75%**
10530	20%
10583	15%
10706	6%
10603	6%
10607	5%
10533	4%
10522	4%
10591	4%
10502	4%
Other	7%
Out of Area Zips	**25%**
10601 (White Plains)	2%
10605 (White Plains)	2%
10536 (Katonah)	1%
10577 (Harrison)	1%
10025 (NYC)	1%
Other	…

Campaign Contributors' Employers

After geocoding the dataset, I started to see if we could lookup who the employers are of the people within this data set. I started by rank ordering the contributors by donation size and manually looking people up. These insights were based off some manual google searching so take it with a grain of salt; just adding here in case this is interesting.

1. Donald W. Cannon (of White Plains, NY), who donated $2,000 in June of 2019 and who is the top individual contributor seems to be the CFO of "Another Nine LLC"?
2. Peter Guggenheimer (of Scarsdale, NY) who donated $1,000 in June of 2019 seems to be the founder of "Guggenheimer Architects"?
3. Many of the other contributors in the data set seem to be much more elderly and not necessarily working currently.

Next Steps

1. Again, would be interested in hearing your feedback if these are the types of analysis/visualizations that are useful for you and your team. Happy to dive deeper on any of these fronts but just wanted to hear from you first.
2. On contributors' employers, I haven't yet done the work to cross reference against federal contributions. We could do that next and cross reference against list of prohibited people from donating due to conflicts of interest (which you said would take a couple weeks to get that list?)

3. I think there are more sophisticated features within the visualization that could be helpful. As with the police data maybe some dashboarding with some summary views?
4. I also think it could be useful here to generate some reports of contributions by geographic area. Eyeballing the map is useful, but reports of campaign contributions by geographic area could also be another way to look at the "approval rate" of Feiner within an area (most concentrated contributions could mean higher approval rate). Could be another interesting data point to decide where to concentrate resources.

Voting Trend Data

Goal: Analyze voting trend data to identify areas to focus campaign resources on. Also look to Census Data to see if there are any trends by socioeconomic status and/or demographics.

Next Steps

1. Haven't done any work on this yet, but am thinking of building off the Google Spreadsheet you put together and adding in demographic and socioeconomic data so that you can also get a sense of the relationship between these variables and approval ratings of Feiner. What do you think?