Geographic Analysis of Nonprofit Data: A How-To Case Study

Jana Fredricks
All nonprofits can benefit from knowing where their patrons are coming from. Nonprofits regularly gather address data on their donors and attendees, and geographic analysis is one useful way to visualize and make the most of that information. MapTogether’s Illustrated Guide to Nonprofit Geographic Information Systems (GIS) provides an excellent and accessible guide to geographic analytical tools (also known as GIS, or geographic information systems). As MapTogether explains, geographic analysis places information in a “where context,” and deepens our understanding of where someone (or something) has been, where they are, or where we can expect them to be in the future.

AMT Lab contributors have explored how geographic analysis can help increase programmatic effectiveness, but there are many other ways nonprofits may leverage their data with geographic analysis. Placing people (donors, funders, grantees, audiences, attendees, visitors) in a “where context” can elucidate and inform marketing, development, and programming strategy. Additionally, the map-based outputs of geographic analysis can be used to easily communicate gaps in current organizational processes and strategic opportunities for improvement.

Unsurprisingly, time, cost, and expertise are the primary barriers to the utilization of geographic analysis. We tend to underestimate the amount of time it takes to properly prepare to undertake this type of analysis, particularly when it is new to the organization. In reality, about 90% of the work happens before the software is even involved. As data maintenance is streamlined and the geographic analysis tools become familiar, conducting these analyses becomes quicker and easier.

This document is a preparatory guide for those embarking on geographic analysis for the first time. This case study explores how one organization used this type of analysis to inform decisions about donor recruitment and retention. This guide aggregates useful resources, highlights important considerations, and outlines the process of geographic analysis step-by-step using a succinct case study.

THE CASE

The following case is built around the donor data of a mid-sized, Pittsburgh-based arts organization. The organization was interested in understanding what strategic insights they could glean from placing their donor data in a “where” context. First, they wanted to visualize the geographic distribution of their donors throughout the city of Pittsburgh. Then, they were interested in targeting current low, mid, and high-level donors to cultivate donors for larger gifts moving forward. In addition to donor cultivation, the organization wanted to identify geographical trends in donor retention, and gain insight into the location of current donors, LYBUNTS (those that gave last year but unfortunately not this year), and SYBUNTS (those who gave some year but not this year). Finally, they wanted to know what (if any) population characteristics correspond with

---


areas of high donor concentration, which could inform future donor cultivation strategies.

RESEARCH QUESTIONS

This case study travels step by step to answer their research questions. (Check out some GIS-specific tips for framing research questions in this article). The case organization wanted to know how geographic analysis can inform future retention and recruitment strategies. Together we defined the following questions:

(1) Where do the case organization’s donors live?
Understanding the geographic distribution of donors and the characteristics of highly concentrated areas could help the organization identify recruitment and retention opportunities.

(2) Where do the organization’s low, mid and high-level donors live?
Donors live in geographic pockets and mapping low, mid, and high-level donors can identify high-ROI targets for cultivation and recruitment efforts.3 External trends such as neighborhood gentrification could influence the geographic patterns of their donor base, or perhaps the organization’s presence in a specific area has decreased, resulting in higher attrition. This question could reveal promising targets for recruitment or outreach strategies.

(3) Where do the organization’s LYBUNTs (those that gave last year but unfortunately not this year), SYBUNTS (those that gave some year but unfortunately not this year), and current donors live?
Segmentation based on donation recency (current donors, SYBUNTS, and LYBUNTS) could help the organization identify areas where donors once were but are no longer.

(4) What are the demographic characteristics of areas with high donor concentration?
By cross-analyzing geographic attributes such population, median age, or per capita income with traditional donor metrics such giving level and frequency, the case organization could deepen their understanding of the lifestyle and defining features of their donors.

SOFTWARE

After weighing the tradeoffs between cost, time, data compatibility and deliverable types, I decided to use a business intelligence software called Tableau to conduct my analysis. To inform this selection, I investigated a variety of software options to see determine which matched the resources and goals of this particular case organization.

There are low, mid, high-cost software available for geographic analysis. One of the more expensive options is ArcGIS Pro, a highly

---

specialized geographic information system. It costs $3000 per year per user, and generally exceeds the needs of most nonprofit arts organizations. At the other end of the spectrum, Google Earth and Tableau Public are free alternatives. Both have basic mapping capabilities, but data privacy should be explored prior to uploading sensitive data.

Tableau, a business intelligence software with mapping capabilities, provides a happy middle ground. Though it’s mapping capabilities are far more limited than ArcGIS Pro, it has many benefits. A Tableau Creator subscription costs $70/user/month (billed annually), and is likely to meet the needs of most arts nonprofits. Tableau Foundation began offering free licenses to nonprofits in 2015. Most small to mid-sized arts nonprofits with an operating budget of less than $5 million are eligible to apply.

Time is the second important consideration I made when evaluating which software to use. These tools can be labor intensive to learn, and I wanted to be sure it would not be too advanced to learn in a relatively brief timeframe. Luckily many of the tools I considered have extensive training resources available, including some specifically tailored to nonprofit needs. ArcGIS, for example, provides advanced analytical tools that serve a very specific purpose: to process and manage geographical data. Because the case organization did not need such a specialized tool, I opted for a more versatile software that could be used for other analytical and data visualization needs at the organization.

I took advantage of Tableau’s training packages and starter kits. After about three hours I had a very basic competency in the software. Within 20 hours of using and experimenting, I had a very operational understanding of Tableau’s essential tools and capabilities.

Another consideration I made before purchasing a subscription to Tableau was the compatibility of the organization’s data type, and the file format I wanted to work with. I wanted to ensure the organization’s data was compatible with the software I chose – for example, Tableau is Salesforce compatible and allows a direct link between the CRM and the software. Additionally, I confirmed that .kmz and shapefiles (the most common spatial file formats) could be imported into Tableau. In the end I did not import spatial files into Tableau, but it was important to consider for future geographic analyses.

Finally, in order to choose the right software, I considered the outputs of the analysis (my deliverables). My outputs were map-based, as are outputs of most geographic analyses – but maps come in many shapes and sizes. I considered the sharability of my analysis; I planned to share it with organizational stakeholders such as staff, board, or potential donors, and wanted to be sure the software I chose allowed for easy sharing.

Some analysis tools are better suited to data sharing than others. For example, Tableau Public and ArcGIS Online provide cloud-based platforms that can be embedded into a website or accessed via URL. If I chose cloud-based options, however, I wanted to be sure the safety of the case organization’s sensitive donor data was secure. Tableau Desktop and ArcGIS Pro, as well as Google Earth and many other software provide local analysis tools. Files created with this software are only accessible by devices that have access to the software. Tableau Desktop, however, allowed me to

---


export .png’s, .jpg’s, and .pdf’s of my maps, so I could easily share my deliverables with organizational staff. (ArcGIS offers a similar feature).

**DATA**

Once I decided to move forward with Tableau Desktop, it was time to turn to the organization’s data. This section explores the challenges I faced when gathering and processing data. First I took stock of the data I did and did not have.

**The Data I Had**

The donor data used in my analysis was exported from the organization’s CRM system and was already quite clean and appropriately formatted. If the organization’s data was siloed between departments, or housed in a spreadsheet that required manual entry, my clean-up job would have been more involved. **Tech Soup** published a nonprofit guide to Collecting and Reporting on Data that provides valuable data entry and management checklists for all nonprofit data (not just donor data).  

I ran a report in Patron Manager (the organization’s CRM) and exported it into a .csv file. The report included the following fields: Account ID (important unique identifier), street address, city, state, zipcode, account type, lifetime donation amount, lifetime donation frequency, date of last donation, and most recent donation amount. Patron Manager allowed for the easy aggregation of fields like lifetime donation amount and frequency. Once exported, only minor clean-up was required. I created a five-digit zip code field. This was a simple left string extraction from the original zip-code field, which had some zip code extensions like 15221-1422. While this was okay, I needed to create a uniform field in order to aggregate my analysis to the zip code level. Additionally, I created an average donation size field, a simple calculation of lifetime giving amount divided by lifetime giving frequency.

After this I conducted a simple preliminary analysis to gain insights regarding data distribution. For example, I knew I wanted to segment the donors based on low, med, and high level giving. I looked at the distribution of average gift size (see Figure 1) to understand where my segments should end and begin. I found that 82% of donors made average gifts of less than $100. 10% of donors average gift was between $100 and $300, and the remaining 8% of donors made average gifts of over $300. I used this distribution to segment donors into low-level donors (average gifts of less than $100), mid-level donors ($100-$300), and high-level donors (greater than $300).

Then I looked at account type. I found that 98% of the dataset was composed of Individual and Household donor accounts, and 2% of Business and Foundation accounts. Given that this analysis aimed to address trends in individual donors, I made a mental note to filter the dataset to eliminate all non-individual account types such as businesses and foundations.

---

The Data I Didn’t Have

Once I was familiar (maybe too familiar) with the data I had, revisited my research questions. Can the data sufficiently answer my questions? The data I just cleaned is sufficient to answer my first three research questions (the geographic distribution of donors, the distribution of low, mid, and donors, as well as the distribution of LYBUNT’s and SYBUNT’s). However, the data is insufficient to answer my fourth research question, which assesses the population characteristics of areas with high donor concentration. The .csv I exported from the CRM does not contain any data on population characteristics. I needed to turn to external data sources. Fortunately, there are excellent public sources of geographic data, such as the Census Bureau, and at the city, county, and state levels of government as well. There are also robust private datasets compiled by organizations like Reference USA, which provide a variety of in-depth data via many public libraries.

Tableau has a very handy native feature of its mapping function. I easily overlaid common population features such as median household income and median age onto the maps I created using Tableau’s in-house population characteristic feature. Ultimately, it was not necessary to seek additional data from external sources for this analysis.

Less common population features such as consumer behavior and other non-Census features, however, are not native to Tableau and would have required data from an external source. If I were interested in more obscure population characteristics or behaviors, I would need to download that data and join it to the donor data, a process explored in great detail in Tableau’s resource for new mappers.7 I could


https://onlinehelp.tableau.com/current/pro/desktop/en-us/buillexamples_maps.htm#Step1Connect
easily download an excel file from the US Census Bureau. Were I to do this, I would first need to identify the geographic focus of my analysis, and decide if I was interested in block groups, census tracts, zipcodes, or some other boundary. The case organization already tracks the zipcodes of their patrons, and so my analysis focused on zip code-level attributes that could be easily joined with the donor data I already had.

**GEOCODING**

Geocoding is the process of using a street address to assign longitude and latitude coordinates. This can be an involved process, particularly if high accuracy is required. In the case of many arts organizations, accuracy is not as a high a priority as it would be for other sectors. For example, mapping a donor’s address a few blocks away from its actual location is low-risk for an arts organization. This would not be the case for an emergency response service like the fire department, where high accuracy is imperative. Because accuracy is flexible for many arts nonprofit analyses, there are a variety of mid and low cost geocoding services that can be used to identify x and y coordinates of donor addresses.

The ArcGIS online editor subscription for $200 per user per year provides access to Esri’s World Geocoding Service. This service has additional costs, but can provide accurate, world-wide geocoding based on frequently updated street-level spatial files. Texas A&M offers a [free geocoding service](http://geoservices.tamu.edu/Services/Geocode/) for national addresses. Data Privacy is a concern with each of these cloud-based geocoding methods. Texas A&M, however, offers an offline service wherein I could send my data to them and they geocode it offline and return a geocoded file.

Because my analysis focuses on aggregating data to the zip-code level, geocoding was not necessary. If I wanted to examine donors in relation to particular streets or neighborhoods, I would need to geocode addresses. However, in this case I am looking at the zip-code level, which does not require geocoding.

**DATA ANALYSIS**

*Import and Play*

Now that I have gathered and processed my data it is time to import it to Tableau. It was necessary to do some back and forth between the software and the data source to get the data prepared for analysis. This was simply a matter of removing data entry errors from the data source, such as dashes or other special characters in the cell values, or changing a field’s data type from text to numerical.

Once I imported my data I began playing around with the different functions of the software. To undertake geographic analysis in Tableau I needed to first assign my data a geographic role. The data source’s geographic fields (zip-code, city, state, or country) field, needed to be assigned to their appropriate geographic role’s in Tableau. Then I started dragging and dropping to visualize variables I was interested in.

I was interested in zip-codes, so I dragged an dropped my zip code field into the “Marks” interface. I was interested in visualizing how many donors live in each zip code, so I right clicked to choose a count of zip-code, rather than a sum. From there I wanted my data to display as a percentage of the total, a simple

---

Table calculation I found by right clicking on the count by zip code label and choosing table calculation. See Figures 2 and 3.

Figure 2: All Time Donors, Screen Shot of Tableau

![Tableau screen shot of all time donors](image1.png)

Figure 3: Screen shot of software table menu

![Software table menu](image2.png)

As seen in Figure 2, Tableau’s default visualization is points, which originate at the center of the polygon metric (in this case, the polygon is zipcodes). Using the Marks dropdown menu, I could easily change the visualization to “Maps” instead of “Circle”. The result was a choropleth map, like the one seen in Figure 4, based on donor density per state. My biggest critique of the mapping capabilities in Tableau is that the symbology (the way data is visualized geographically) is largely un-editable. I was not able to greatly influence the color or distribution categorization. As a result, the visualizations lacked some of the sensitivity or specificity one might hope for. One solution I found when looking at very small geographic areas, like a particular city or county, is to use graduated point symbols rather than choropleth maps (which are based on polygons).
Graduated point symbols are easy to create using Tableau. Once my map was visualized using points (like the map shown in Figures 2 above, and Figure 5 below), I dragged the measure of interest and dropped it on the Color and Size boxes in the Mark sidebar. Figure 5 shows donor density using graduated point symbols. The larger and darker the point, the higher the concentration of donors. Each point in Figure 5 represents donor density in a single zip code.

**Explore and Analyze**

Now that I imported my data and got my bearings, I returned again to my research questions, and developed visualizations that answer those questions.

**Question 1: Where do donors live?**

This is visualized in Figure 6 using the graduated point method from above. It is evident that zip code 15212, where the organization is located, has a high donor concentration: over 10% of total donors. Additionally, 15217 also has a donor concentration of over 10%, and 15206 of over 7%. It is likely that the organization’s
Development Director is already aware of areas of very high donor concentration. What perhaps is more interesting, are the mid-sized circles which are home to 3 or 4% of total donors.

Figure 6: Screenshot answering where do donors live in a density map

---

**Question 2: Where do low, mid and high-level donors live?**

Now that I have identified where donors live, it is time to segment them based on areas of interest. This question concerns the average gift field I created in excel after exporting the data from Patron Manager. The distribution of average gift suggested the following categories: low-level donors (less than $100), mid-level donors ($100 - $300), and high-level donors (over $300). I used these categories to create groups of my data, narrated by Figures 7, 8, and 9. Then, I filtered the visualization using my new groups for Low, Mid, and High-Level donors and used graduated point symbols to create maps for each category.

Figure 7: Screenshot of software menu for creating groups of data
Figure 8: Screenshot step 2 of creating groups in Tableau

Figure 9: Screen shot step 3 of creating groups in Tableau

Figure 10: Screenshot of low-level donor density map of Pittsburgh, PA
Figures 10, 11, and 12 provide some interesting insights. First, it is evident that there are more low-level donors than mid and high-level donors, and more mid-level than high-level donors. Unsurprisingly, the three zip codes previously identified house the greatest concentration of low, mid, and high-level donors. Low-level donors are more dispersed around and near the city of Pittsburgh, while mid and high-level donors are concentrated in the city proper. This map reveals zip codes that may be ripe targets for cultivation efforts, moving low-level donors up to mid, and mid to high-level.

Beyond these areas of notably high donor concentration, there are some interesting insights to be gleaned from mid-concentration areas. For example, in Figure 10, it is evident that some low-level donors (between 4% and 8%) reside in several zipcodes along Pittsburgh’s Northern perimeter, just south of the northern river. Mid and high-level donors are minimally represented in these zipcodes. These may be important areas to target for more significant contributions, or perhaps there are reasons the organization should be aware of that these donors are not able to give at a higher level.
Question 3: Where do LYBUNTs (those that gave last year but unfortunately not this year), SYBUNTs (those that gave some year but unfortunately not this year), and current donors live?

Using the same methodology as above, I created three new map visualizations for each of these categories. The table in Figure 13 shows the distribution of current donors, LYBUNTS, and SYBUNTS, and Figures 14, 15, and 16 show the corresponding maps.

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Percentage of Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYBUNTS</td>
<td>Donors that gave some year but not in 2018</td>
<td>30%</td>
</tr>
<tr>
<td>LYBUNTS</td>
<td>Donors that gave in 2017 but not 2018</td>
<td>27%</td>
</tr>
<tr>
<td>Current Donors</td>
<td>Donors that gave in 2018</td>
<td>43%</td>
</tr>
</tbody>
</table>

Figure 14: Screenshot of ‘some year but not this year’ donors density map
These figures show the slow dispersion over donors over time. Figure 14 shows relatively high SYBUNT concentration in the city proper. Figure 15 shows LYBUNT concentration disburse into the perimeters of the visualization, and Figure 16 shows that current donors are concentrated in areas of high overall donor concentration, but also fan out into the perimeters of the map.

While there are more current donors than either LYBUNTS or SYBUNTS, the organization is losing more donors than they are gaining in areas of high overall donor concentration. This is visualized by the opacity of the hues of the red circles. The deeper hues of the graduated points in Figure 14 indicate the majority of SYBUNTS residing in 15212, 15217, and 15206. Over time, donors in these zip codes have decreased. Zipcode 15212, where the organization is located, is home to only 31 current donors, 44 LYBUNTS, and 101 SYBUNTS. Some of the LYBUNTS may be planning to give during the final quarter of 2018, which this analysis does not account for. The data suggests, however, that the organization is losing more donors than they are gaining in this zip code. Similar though less pronounced trends can be seen in 15217 and 15206. Areas of milder concentration may also provide valuable
insights to the organization. This information can be used to target outreach and retention strategies.

**Question 4: What are the demographic characteristics of areas with high donor concentration?**

This question addresses the geographic attributes of high-concentration donor areas. Though it is impossible to draw finite conclusions, this context provided additional insights into the demographic characteristics of the organization’s donors. Tableau offers a convenient way to visualization the most basic demographic and population characteristics. By choosing the map drop down menu and selecting Map Layers, Tableau provides native map layers of these attributes with a simple click of the button (see Figure 17). I used these layers to create zip-code level choropleth maps using characteristics of population, race, and median age. I then overlaid these choropleth maps with graduated point symbols representing donor concentration (see Figures 18, 19, and 20).

Figure 17: Screenshot of software menu for data layers in Tableau
Figure 18: Screenshot of density map showing donors by race

Figure 19: Screenshot of density map showing donors density population

Figure 20: Screenshot of donors age density map
This feature in Tableau is simple to execute, but also somewhat challenging to read. There is no way to edit the bins of the native population attributes, and color schema are very limited. Nonetheless it produces useful, incredibly simple visualizations. For example, it appears in Figure 19 that some areas of high donor concentration do correspond with areas of the city that are densely populated. Figure 20 shows that areas of high donor concentration correspond with a median age of less than 41. Figure 18 shows areas of high donor concentration also have higher white population. One of the downsides is that race attributions, like white population, are shown as numbers rather than percentages of total population. There is no way to change this, but it is necessary to note.

CONCLUSION

This analysis used geographic tools to examine data in a “where context”, and will inform the case organization’s donor cultivation, retention, and recruitment strategies. Moving forward, they may choose to invest their resources into marketing and outreach efforts in areas of low or medium donor concentration. Understanding the distribution of donors will allow them to tailor events to be focused on recruiting new supporters or cultivating existing supporters. Similarly, the organization may target areas of high donor concentration for fundraising events. Areas that abut zip codes with a concentration of high-level donors are prime targets for cultivation efforts. The case organization may examine the slow dispersion of their donors across Allegheny county, and employ evaluation tools such as surveys and focus groups to deduce why this dispersion is occurring. Placing donor data in a “where context” will allow the case organization to develop targeted strategies that increase ROI and increase resource efficiency.

This analysis can be used across departments to generate strategies for marketing and developing initiatives, as well to assess programming reach and the achievement of outcomes. Visualizing data illuminates the overlooked, and is invaluable internally and externally, as organizations seek to share their story with potential funders, program participants, and boards.
BIBLIOGRAPHY


Liu, Peixiao. n.d. “Getting Your Arts Organization Started with Location Based Social Marketing,” 12.


