teradata.

Examining Hire Heroes USA Individual Donors: Geographic Identification & Analysis

Author(s)

- Hope Spenik, Arthur Johne, Jessica Lavrinovich, Michael Calvelli, Ryan Castro
- Virginia Commonwealth University
- lavrinovichjm@mymail.vcu.edu, spenikhk@vcu.edu, johneaa@vcu.edu, castrorj@vcu.edu, calvellima@vcu.edu
- Information Systems
- Undergraduate

Faculty Advisor

Joseph Cipolla, Department of Information Systems, jcipolla@vcu.edu

Summary

As a nonprofit, Hire Heroes USA (HH) highly values its network of individual donors, whose monetary donations help enable the pursuit of HH's mission and vision. The goal of this report is to bring further growth and expansion to HH's network of individual donors through geographical identification and analysis.

Gaining insight on the locations of past and current donors will allow HH to highlight areas in the country which may carry the potential for charitable support. By bringing awareness of the organization and its various donation avenues to these areas of potential, HH can work to broaden its donor network and grow one step closer to being the nation's preferred veteran employment service organization.

Problem and Motivation

Given the set of provided questions for the TUN 2019 Data Challenge, the team chose to orient the focus towards Hire Heroes' development as an organization. HH falls within the top 10% of nonprofits in the U.S., with regard to its annual gross receipts (McKeever, B.S.). For large public charities like HH, the organization's ability to maintain diverse revenue streams is vital to their stability and future operations. With this in mind, the team narrowed the focus to HH's individual donors. The questions used for the basis of this report, as provided by TUN, are listed below:

- Is there a geographic location within the US that most of our individual donors come from?
- Do our social media posts or fundraisers calling for donations hit these areas with little to no donors?

Considered to be the grassroots of a nonprofit, individual donors provide a foundation for the organization, as well as a critical pathway which allows the word of their mission to extend across communities and reach new ones. Ultimately, a nonprofit's success relies on the public's support of their charity, and that support solidifies at the individual level. Gaining a geographical understanding of HH's current network of individual donors will enable HH to extend the reach of their mission even further by highlighting areas of success and areas of potential.

Approach

In order to properly address the problem, the team had to develop an approach to deal with the synthesis of the question at hand, the data provided, and the analysis of the data to provide an answer for the question overall. The team approached the business problem by using a top-down style, making use of the two questions provided by HH. In order to prove the team's case, assumptions were identified, as well as datasets, and data types that could potentially apply to the question.

The team proceeded to work under this top-down approach for the majority of the project while also employing various creative techniques to aid in the development of the solution. The team followed a cyclical process for development. The process started with each team member examining the data sets and types that were identified as being relevant to the question during prior team discussions. The team would regularly convene to share new findings, provide updates to the project, and ideate on how to further progress with the data. The primary creative technique that the team implemented was to regularly present the team's findings to a classroom of peers and the instructor to receive feedback on the team's progress. Through the approaches and techniques described, the team was able to develop a more thorough analysis of provided datasets.

Datasets

In order to pinpoint the location of all individual donors, the team needed to first understand which of the provided datasets were appropriate for this analysis. The datasets used for analysis and visualizations are as follows:

- "SalesForce_Opportunity.csv" (Donations)

- “SalesForce_Contact.csv” (Contacts)
- “Campaign.csv” (Campaigns)

Beginning with the field “Donation_Type_C” in Donations, the team was able to distinguish all donation records associated with Individual Donors. In Donations, “Contact_Owner” contains all associated contact IDs and is consistent with IDs found in Contacts, making it possible for a right join of the two tables. A total of 3,629 individual donors were associated with a unique Contact ID, however, 313 records were found to have unassociated Contact IDs. A join of these two tables made it possible to associate a geographic identity to each individual donor.

Mailing postal codes and mailing states within Contacts were found to be associated with roughly 50%, and 90% of individual donors, respectively. In pursuit of data completeness and accuracy, map visualizations for individual donor locations (Figure 1) were rendered using the mailing state. The Donations dataset also provided data for the remaining visualizations (Figure 2, 3) using fields “LeadSource”, “FiscalYear”, and, “npe01__Payments_Made__c” for determining donation totals.

Assigning geographic data to social media posts and emails was technically feasible through analysis of HH’s various social media outlets (e.g., Twitter, Instagram, etc.) though, was not possible using HH’s provided datasets, as no location data was assigned to digital fundraising campaigns in a structured format. An exhaustive cleansing and preparation of the “Name” field found in Campaigns could carry potential for geographical data associations to be made with digital campaigns, however this exceeded the scope of this report. Additionally, comprehensive analysis of external social media accounts did not fall within the scope of work for this report.

Despite the lack of digital and social media campaign location data, nearly all “tangible” fundraisers, found in Campaigns, did have associated geographic data in a relatively structured format, in the field, “Program_Location”. However, taking into account the relatively low contribution that HH’s tangible fundraiser events have as a lead source with regard to those with greater influence (Figure 3), we chose not to pursue any further geographical analysis for brick-and-mortar fundraisers.

Tools and Analytics

Tools used for data analysis included Microsoft Excel, Tableau Desktop 2018.3, and Tableau Prep 2019.1. Data collection, transformation/consolidation, and storage were executed in Excel and Tableau Prep. Visualizations were performed within Tableau Desktop. The team examined all datasets provided by HH and selected which sets were relevant to the business problem. Using Excel for the general viewing and observation of datasets, the team then sought to decide which data types were useful. Data manipulation, to the extent of table joins, aggregations, calculations, unions, assignment of data types, and file exports for use in Tableau, was performed using Tableau Prep.

Results

Between 2007 and 2019, Individual Donors accounted for 4.9% of HH’s total donations. Despite accounting for a considerably small portion of donations for HH as a whole, individual donors are traditionally considered to be the grassroots of a nonprofit organization and thus carry great weight in extending the charity’s reach. Individual donors are the inherent foundation

of HH, and likely choose to donate as a result of their resonation with HH's mission. Our recommendations for growing HH's individual donor network are as follows:

- ➢ Prioritize campaigns for states with high donor potential.
  - FL, NY, IL, NC, VA, PA, OH, WA
- ➢ Grow and maintain strong donor states.
  - GA, CA, TX, NJ
- ➢ Re-establish mail as a primary lead source.
- ➢ Encourage data completeness at intake and storage.

Figure 1 denotes the locations of all individual donors and their associated mailing state within the continental U.S. as well as the average amount for a given donation. Creating the distinction between a strong individual donor state and a state with high potential for growth was developed by taking the product of a state's average donation amount and its total individual donor count. "Strong" states indicate high donor populations and high donation averages. "High potential" states indicate either high donation averages and moderate donor populations, or vice versa. North Carolina was initially considered to be a strong state, however, was re-established as carrying high potential following further analysis which had uncovered an outlier resulting in skewed donation averages. Figure 1 does not reflect the removal of this outlier.

Figure 2 indicates that a majority of donations are associated with mail and web as their lead source. Figure 3 depicts the distribution of these donation lead sources from 2007 to 2019 and demonstrates that mail campaigns had been the primary source of individual donations through 2016, with web-based campaigns taking the lead in years following. It is presumable that HH has decided to set greater focus on web-based donation campaigns while reducing their mailing campaigns. This shift to web donations has had a negative effect in the short term and suggests that mail campaigns should still be a strong consideration as a primary campaign source for HH in the future, as they tend to offer a more personal experience for the donor in comparison to websites as the medium.

Figures 2 and 3 also highlight shortcomings with regard to data completeness; the "Unknown Lead Source" category references null values from "LeadSource" in Donations. Initial data analysis was not successful in resolving these null values, which account for 17% of individual donations. Considering the sudden uptick of these empty values the matter should be addressed prior to becoming a potentially greater issue in the long-term.

Contributions and Uniqueness

The team had a cohesive structure among us that not every team could lay claim to; seamless cooperation between us positively benefited the results we found. Going into this analysis, the team realized that in order to be successful we had to play to the strengths of each member. For example, the left-brained Arthur and Ryan used their talents to clean and join data tables utilizing Tableau; and the right-brained Michael, Jessica, and Hope used their eyes for visuals to create graphs and draft a cohesive write-up. This allowed us to utilize the individual resources efficiently to reach new heights with the analysis that otherwise we wouldn't have reached.

A key and surprising insight we found during the analysis was how little donors make up HH's funding. Further insight into where HH derives their funding showed that only around 5% of their funds were collected from individuals. Does this pose a question for further analysis in

the future: if individual donations are relatively insignificant, is targeting them effective use of resources?

Appendix: Data Visualizations

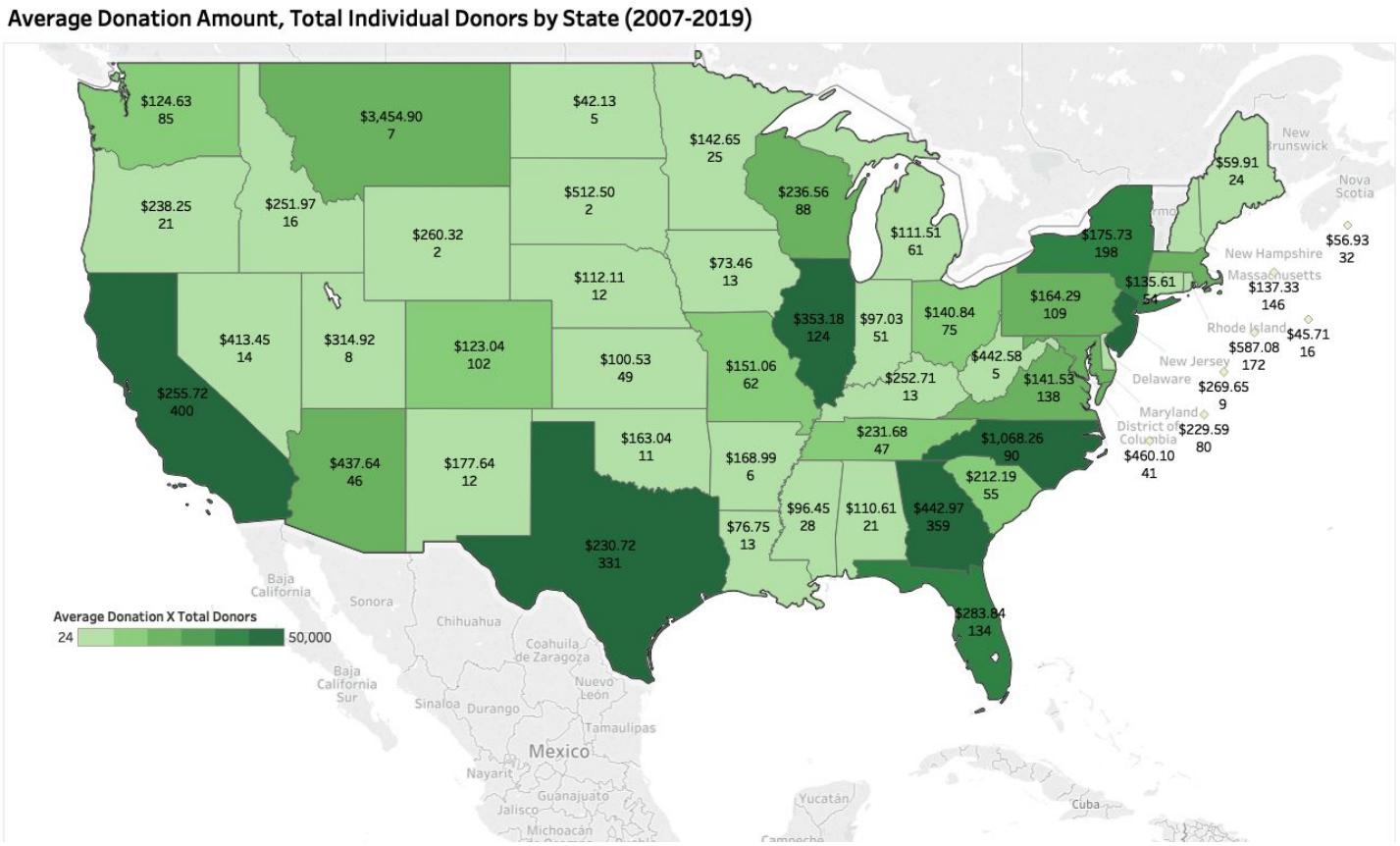

Figure 1

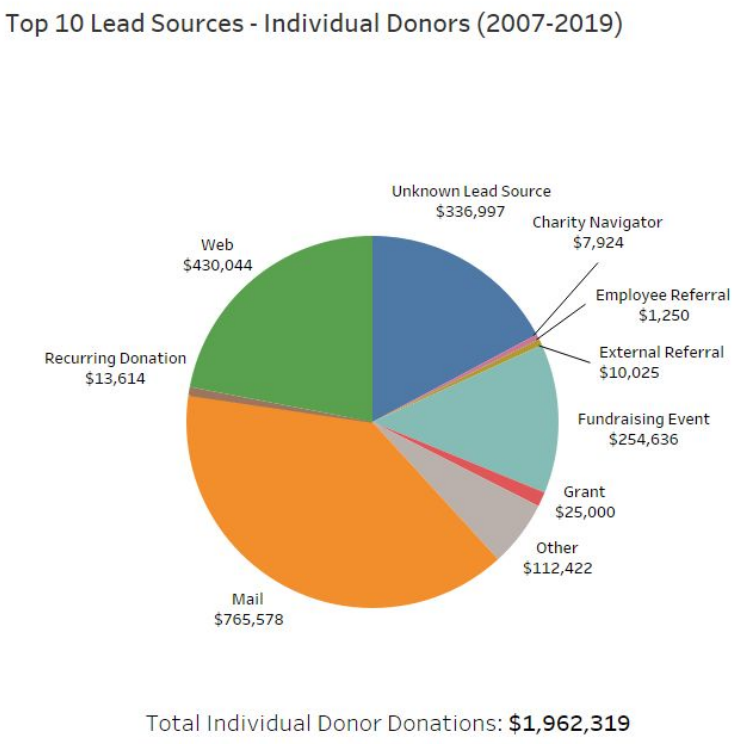
Top 10 Lead Sources - Individual Donors (2007-2019)
Unknown Lead Source
$336,997
Charity Navigator
$7,924
Web
$430,044
Employee Referral
$1,250
External Referral
$10,025
Recurring Donation
$13,614
Fundraising Event
$254,636
Grant
$25,000
Other
$112,422
Mail
$765,578
Total Individual Donor Donations: $1,962,319

Figure 2

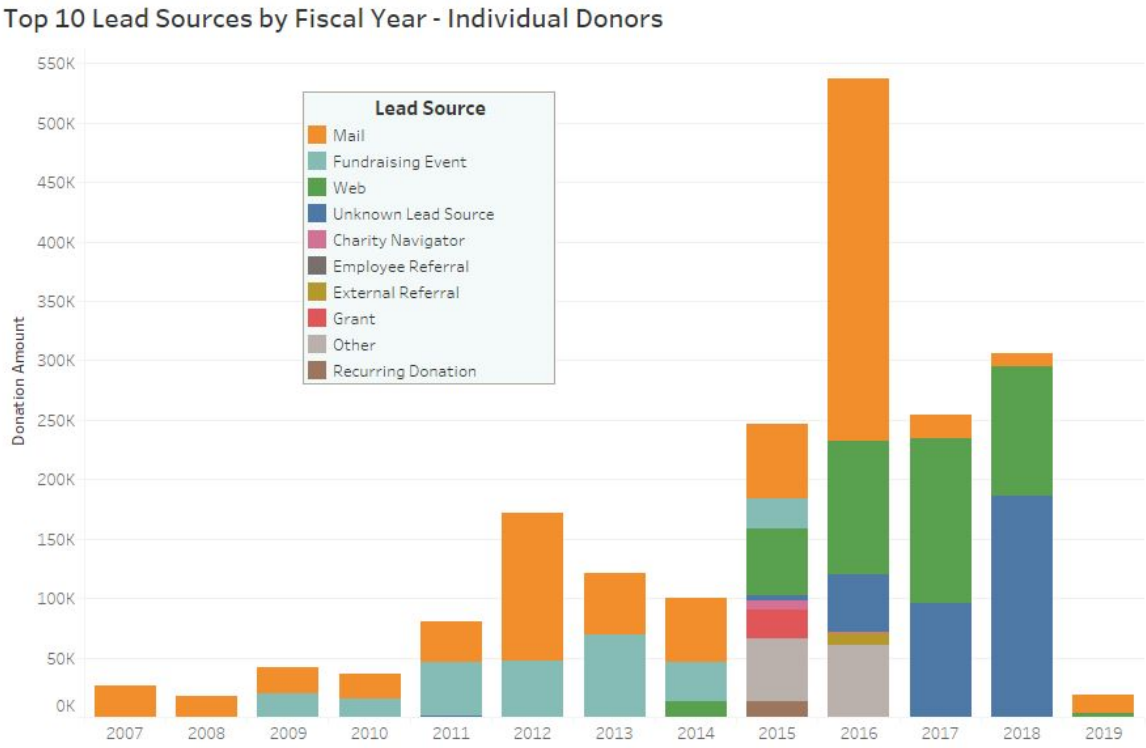
Top 10 Lead Sources by Fiscal Year - Individual Donors
Donation Amount
550K
500K
450K
400K
350K
300K
250K
200K
150K
100K
50K
0K
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
Lead Source
Mail
Fundraising Event
Web
Unknown Lead Source
Charity Navigator
Employee Referral
External Referral
Grant
Other
Recurring Donation

Figure 3

1) References

- Hire Heroes USA. (n.d.). Retrieved from https://www.hireheroesusa.org/
- 2019 Data Challenge. (n.d.). Retrieved from https://www.teradatauniversitynetwork.com/Community/Student-Competitions/2019/2019-Data-Challenge/Submission-Guidelines
- McKeever, B. S. (2018, November). The Nonprofit Sector in Brief 2018: Public Charities, Giving, and Volunteering. Retrieved April 29, 2019, from https://nccs.urban.org/publication/nonprofit-sector-brief-2018#finances-01