Data for Good In Your Neighborhood: A case study on how data can benefit your local community

Timothy NeCamp
University of Michigan
Ann Arbor, MI
tnecamp@umich.edu

Emily Morris
University of Michigan
Ann Arbor, MI
emorrisl@umich.edu

Evan Reynolds
University of Michigan
Ann Arbor, MI
evanlr@umich.edu

Stephen Salerno
University of Michigan
Ann Arbor, MI
salernos@umich.edu

ABSTRACT

Data are beneficial resources for big businesses, national and state governments, and large nonprofits. In this article we describe how the benefits of data can be similarly reaped by local community organizations. We exemplar this benefit through Statistics in the Community (STATCOM) at the University of Michigan. STATCOM is an organization which connects statistics graduate students to local community organizations to provide free assistance around data organization, analysis, and interpretation. We describe three specific projects where STATCOM members partnered with a food assistance program, a crisis center, and a community foundation. For each project, we provide an overview of the community partner, the data available, the statistical methods used, the findings, and discuss how the analyses impacted the partner and the local community.

1. INTRODUCTION

A food assistance program needs to determine the optimal placement of mobile food pantries in low-income and underserved urban areas. A crisis call center hopes to accurately predict the number of calls from runaway, homeless, and high-risk youth in order to appropriately staff its counselors. A local community foundation wants to decide how to best allocate 18 million dollars in funds to improve quality of life services for their senior population.

Each of these organizations strives to provide public good directly to its surrounding communities but is hindered by these needs. Often, local community efforts hope to answer similar questions but do not have the means to do so [12]. What they sometimes do have, however, are data. The food assistance program, Food for Thought Toledo (FFT) [9], has data on levels of poverty, nutritional stress, food deserts, and other key neighborhood characteristics. The crisis center, Ozone House of Ann Arbor [16], has daily call logs for the past 10 years. The local community foundation, Ann Arbor Area Community Foundation (AAACF) [1] has responses to a survey instrument designed to assess quality of life in older adults.

In many academic fields, answering a question is synonymous with collecting and analyzing data. It is then through principled statistical methods that one can utilize the power of these data to make inferences and predictions. For local community organizations, data can be similarly harnessed to evaluate need, allocate resources, broaden impact, and assess efficacy. Unfortunately, the power of data often remains untapped due to a lack of resources and statistical support, especially at smaller local community organizations. Statistics in the Community (STATCOM) at the University of Michigan [17] aims to fix this problem by offering, free of charge, the expertise of statistics graduate students to nonprofit governmental and local community organizations in the areas of data organization, analysis, and interpretation. This includes schools, government agencies, health centers, homeless shelters, advocacy groups, libraries, adult learning centers, and other agencies and organizations.

STATCOM, which is directed and staffed primarily by graduate students, separates itself from other data for good organizations by working with small local community partners. Many other organizations, such as Data Science for Social Good [4], Statistics without Borders [18], DataKind [5], tend to work with either large institutions or international/national organizations (e.g., American Red Cross, UNICEF in Sierra Leone, and The Municipality of Rotterdam). STATCOM is currently undertaking 10 projects across Southeast Michigan and Northern Ohio, including organizations such as The Children's Center of Detroit [19], Genesee County Health Department [11], Detroit College Access Network [6], and Michigan Broadband Cooperative [13]. By bringing statistics into our community, STATCOM directly touches the people it helps, leaving the greatest impact on those in its direct proximity.

In this paper, we describe three specific STATCOM projects working with (1) FFT of Toledo, Ohio, (2) Ozone House of Ann Arbor, Michigan, and (3) the AAACF of Washtenaw County, Michigan. We highlight the problems at hand, the data being used to solve those problems, and interesting results and findings. We also emphasize the community impact and how each community partner is using STATCOM's work.

Bloomberg Data for Good Exchange Conference.
16-Sep-2018, New York City, NY, USA.

2. OPTIMIZING FOOD PANTRY LOCATIONS WITH FOOD FOR THOUGHT TOLEDO

Food for Thought Toledo (FFT) [9] is a non-profit organization that serves families experiencing food insecurity in the Toledo, Ohio metropolitan area. Their work is done through a mobile pantry which travels throughout the region to partner locations once a month. FFT provides aid to approximately 400 families each month [10]. Despite making an impact on the community over the past ten years, FFT still faces challenges positioning themselves in neighborhoods that are easily accessible to those in highest need. Having collected data on the performance of 35 different locations for several years, there was a great opportunity for collaboration with STATCOM, who could provide the statistical support to identify optimal pantry locations.

When receiving food from an FFT pantry, families were asked to report an address, as well as other demographic data such as how many people were in their household. FFT gathered this information over the past four years and across the 35 mobile pantry locations. In addition, the group periodically surveyed their customers. These surveys indicated that about 80% of families in need drove or received a ride to the pantry, and the average travel time was about 13 minutes. We can observe in Figure 1 how far some families travel to a pantry location, further demonstrating the need for a different allocation of pantries. Several health characteristics and food preferences were also surveyed to better gauge the needs of those being served. This information was used to assess the characteristics of a good pantry location. Defining the 'best' locations required close collaboration between FFT and the five-member team of STATCOM students. Many factors influence the definition of an optimal location, which requires balancing easy access to the mobile pantry, need for their services, and available partner organizations. Communication with FFT and understanding their priorities was especially crucial for the analysis to be useful.

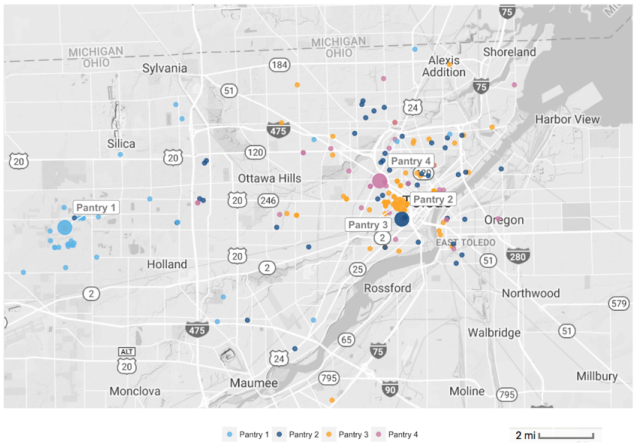

Figure 1: Map of pantry locations and the households who visited during a week in December 2017. The map illustrates the distance traveled by families to each of these four pantries (numbered in the order that mobile pantries were available that week).

2.1 Analysis Methods

Clustering analysis was used to group and rank areas with similar characteristics into need categories; neighborhoods were ranked based on their need, as defined by regional poverty, unemployment, and health characteristics. Obtaining this ranking was useful for FFT location planning. The team was able to further their analysis by also optimizing the selected sites. The distances each household traveled to obtain a pantry's services were recorded and an optimization model was developed to minimize the total distance traveled, constrained to higher priority neighborhoods and other limiting factors specified by FFT collaborators. A location modeling approach was used to model the network of demand nodes, using all the past FFT partners as candidate locations. The objective function minimized accruing unmet demand with constraints related to the number of visits per month and the needs of the households covered by a candidate location. The model was implemented with Cplex software.

2.2 Results

The optimization model generated the optimal locations, and the order with which they should be visited, in order to best allocate resources to the households across the city within a given month. Figure 2 shows the map of pantry locations used in December 2017 and the locations recommended from this analysis. Many of the selected locations were well attended when used as a partner site previously, so it is unsurprising to see them selected as a high priority pantry location. Two current pantry sites were selected as optimal locations for the future, indicated by the overlapping points on the map.

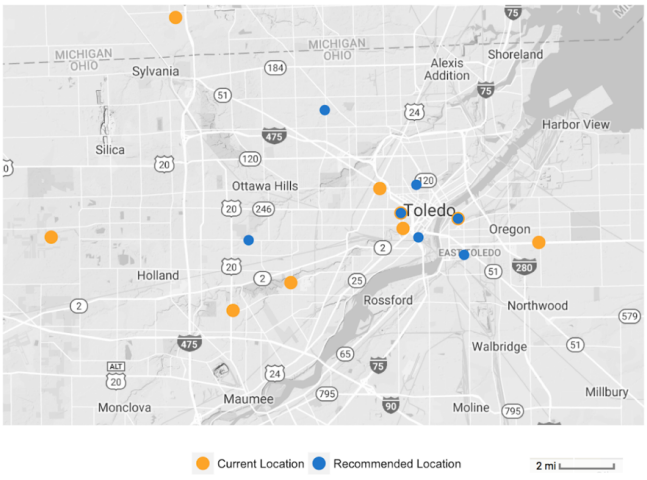

Figure 2: Map of current FFT locations and recommended locations selected by the optimization model.

2.3 Community Impact

FFT plans to schedule pantry locations according to those recommended in this analysis by selecting existing partners and past partners in close proximity to optimal pantry locations. By using these recommended locations, we hope that those with the highest need throughout the greater Toledo area are better served.

As an ongoing project, STATCOM plans to continue its partnership with FFT after they have implemented the recommended pantry locations to evaluate the impact of these changes. Not only would FFT hope to see an increase in the number of users, they would also hope to see customers traveling shorter distances. Currently many families report traveling to more than one pantry location and traveling beyond the nearby pantry sites, demonstrating that their needs are not being met by the current availability of the pantries in their area. By optimizing these locations, more families in need would have a convenient source of food throughout the month. STATCOM hopes to provide a platform which allows collaborators to continue to use the results from the analysis in the future. For FFT, STATCOM intends to provide a way to evaluate the performance of new pantry locations throughout the implementation process.

3. UNDERSTANDING CRISIS CALL PATTERNS WITH OZONE HOUSE OF ANN ARBOR

Ozone House of Ann Arbor, MI [16] has provided critical support, intervention, and assistance to homeless, runaway, and high-risk youth since 1969. With an estimated 1,300 homeless youth residing in the greater Ann Arbor area each year, this organization is an important resource for the youth and their families, especially high-risk youth who have experienced violence, trauma, or contemplated suicide [15]. The doors at Ozone House are always open, to aid in times of crisis [14].

Ozone House operates a 24/7 crisis line to provide support for high-risk youth. To best serve the Ann Arbor youth in need, staffing and resource allocation is an important consideration. Understanding trends in the crisis calls that Ozone House receives can help staff prepare for what types of calls are expected. Further, determining when there are periods of high and low predicted call volumes can ultimately better allow staff to be prepared for a rush of calls or a slow afternoon. With the goal of always being there for the youth who need help, adequately staffing the crisis line is an important step.

In this project, a team of five STATCOM students worked with Ozone House to analyze data from the past ten years, amounting to over 16,000 crisis calls. For each call, Ozone House staff collected the date, time, call duration, type of call, age of the caller, and a five-point anxiety score measured pre- and post-call. From the date collected, the day of week, month, and year were extracted.

3.1 Analysis Methods

The STATCOM team provided a descriptive analysis to understand the days in which calls are more frequent, the types of calls that are most frequent, the average age of the caller, and the average call duration for each call type. Using data visualization techniques, the team discovered interesting temporal trends that highlighted high call volume trends on a weekly, monthly and yearly bases. Finally, the team developed a time-series model that can be used to predict the call volume on any given date in the future.

3.2 Results

Data visualization identified Mondays as the day of the week with the highest frequency of crisis calls. Interestingly, there were no obvious trends within a month or across months. There was a moderate linear increase in call frequency from 2007 to 2016. There were also statistically significant improvements in caller anxiety scores for each of the following call categories: abuse, homelessness, suicide, violence, and other. For each type of call, the STATCOM team calculated the average call duration and reported the distribution. The predictive time-series was built using date and time information. Currently, the model is being tested to assess model performance.

3.3 Community Impact

A primary goal of STATCOM's work is to ensure that projects achieve long-lasting public good. Since projects consist of partnerships with organizations in its shared community, STATCOM especially has the desire to make sure the results can continue to be used in the future, with and without hands-on guidance from STATCOM volunteers. Therefore, creating usable prediction tools and statistical models is an important last step in this community partnership. Currently, STATCOM volunteers are building an R Shiny Application to allow Ozone House to predict the future number of crisis calls on a given day. Volunteers are also helping to set up a database in which volunteers can enter new data as it is received. Then, using the application, retrieve real-time summary statistics, graphical displays of the data, and predict future call volume. By optimizing their staffing based on these predictions, Ozone House will be able to better allocate resources and ensure it can meet the complex needs of high-risk youth in Ann Arbor.

4. ASSESSING QUALITY OF LIFE OF OLDER ADULTS WITH THE ANN ARBOR AREA COMMUNITY FOUNDATION

In another project, STATCOM partnered with the Ann Arbor Area Community Foundation (AAACF) [1]. AAACF is a philanthropic organization which aims to improve the quality of life of citizens within Washtenaw County, Michigan. Recently, AAACF received over 18 million dollars to help improve the quality of life of older adults who are aging in place within the county, especially those with lower life expectancy and lower socioeconomic status [2]. In order to both identify these at-risk populations and understand their needs, AAACF partnered with a team of four STATCOM students to develop, administer, and analyze a survey assessing the quality of life of older adults (60 years of age and older) residing in Washtenaw County.

4.1 Survey Development

To assess the quality of life (QOL) of the seniors in Washtenaw County, STATCOM volunteers developed a survey which utilized the brief Older People's Quality of Life Questionnaire [3], an accepted measure of QOL for older adults. To obtain an understanding of useful services, a qualitative question was included: 'What services do you need to stay happy, healthy, and independent?'

Demographic characteristics were also collected to better understand the variation in QOL across sub-populations.

Since AAACF intended to use the funds to help at-risk older adults, they were particularly interested in three subpopulations of older adults: (1) financially vulnerable, (2) geographically vulnerable, and (3) living alone. Financially vulnerable older adults were identified through 3 questions about financial security. Geographically vulnerable older adults were identified by those living in two zip codes of Washtenaw County containing more older adults with known lower life expectancy. Older adults living alone were identified based on their response to a question of living alone status.

4.2 Survey Administration

Due to limited resources for administering the survey, STATCOM partnered with the Community Technical Assistance Collaborative (CTAC) [7], a student organization in the School of Social Work which focuses on data collection and evaluation. Since a goal of the study was to understand QOL for at-risk older adults, the STATCOM team aimed to administer a large number of surveys to that sub-population. To do this, STATCOM partnered with other Washtenaw County community organizations (such as Meals on Wheels, Catholic Social Services, and various senior centers) which provide services to older adults. Through these partnerships, a convenience sample of 629 surveys was collected.

4.3 Survey Results

In order to determine which of the vulnerabilities were associated with significant differences in QOL, the STATCOM team analyzed the survey responses by regressing a composite QOL score (ranging from 14 to 70) on demographic variables and vulnerability indicators. The survey responses showed that financially vulnerable older adults and geographically vulnerable older adults had significantly lower QOL when compared to their non-vulnerable counterparts, with an estimated average drop in QOL score of -4.11 and -7.07 for financial and geographic vulnerability, respectively (p-value < 0.0001 for both). Older adults who were living alone did not have a statistically significant reduction in QOL (p-value $= 0.60$).

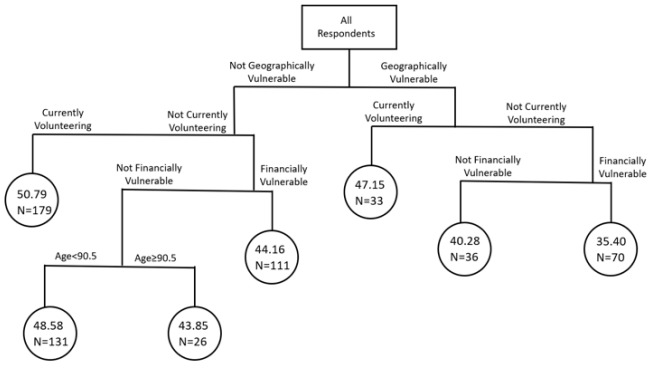

Figure 3: Regression Tree for Quality of Life. The value in each node represents the average QOL for that group. N indicates the sample size.

STATCOM volunteers then performed a Classification and Regression Tree (CART) analysis to develop a decision tree for the AAACF. The resulting tree is displayed in Figure 3. In the trees, the value of each node represents the average QOL score for that group and 'N' indicates the sample size of that group. The community partners appreciated the tree visualization because it illustrates the hierarchical relationships among vulnerabilities. Tree-based modeling approaches also help inform future policy decisions by prioritizing older adult subpopulations in highest need. The tree shows that respondents living in vulnerable zip codes and not currently volunteering have a low quality of life. Specifically, those who live in vulnerable zip codes, are not volunteering, and are financially vulnerable comprised the lowest average QOL group. Conversely, those who do not live in vulnerable zip codes and who volunteer have the highest average QOL. In addition to the vulnerabilities defined a priori, there is also a strong association between higher QOL and older adults who volunteer.

Lastly, the STATCOM team classified the responses to the qualitative survey question in a word cloud to visualize the results (Figure 4). The word cloud emphasizes the services most highly valued by respondents. As shown, responses surrounding transportation and household services had the highest frequency as necessities for a happy, healthy, and independent life. Healthcare, money, and social services were also emphasized as important.

Figure 4: Word cloud of topics of responses to 'What services do you need to stay happy, healthy, and independent?'

4.4 Community Impact

The STATCOM team presented the survey results to the AAACF, administrators of the Glacier Hills Legacy Fund [2], and at various town hall meetings throughout Washtenaw County. In addition, the team created community reports for all of the cities and municipalities that were surveyed. This information is currently being used by the AAACF to allocate funding for services for varying older adult populations within the county, and the community reports have helped inform decision making for local governments in smaller regions of Washtenaw County.

5. CONCLUSION

In this article we described how data for good work is both necessary and valuable to help small, local community organizations. We showed how STATCOM at the University of Michigan uses the expertise of statistics graduate students to help such community partners use their data to better allocate resources, understand their target populations, and improve their operational efficiency. We highlighted three particular projects STATCOM is currently involved with. Though the impact of this work may be smaller and more localized than partnerships with larger national and international organizations, the impact can be much more direct; we have shown how the results of these collaborations are being used by community partners to inform their decision making, and, in turn, help members of the local community in highest need.

In addition to benefiting the community partners associated with each project, STATCOM's current work also helps open doors to future collaborations. Many of STATCOM's projects come through long-standing partnerships with service-based organizations within the University of Michigan (e.g., the Ginsberg Center [8] and CTAC [7]) and the surrounding community, or through referrals from community partners with whom STATCOM has previously worked with. This word-of-mouth style introduction to new projects provides a strong foothold for expanding STATCOM's efforts.

Students who participate in STATCOM also receive many benefits. Students learn how to work with real data and use their statistical knowledge to answer complex questions. In addition, the ability to effectively communicate statistical concepts and results to others is a crucial skill that will benefit students throughout their careers. Lastly, through work with local community partners, STATCOM students receive fulfillment from using their unique skills to positively impact their community and those in need.

Since its inception in 2006, STATCOM at the University of Michigan has been involved in over 50 projects with community-based organizations. Ten of these projects are currently ongoing. The data provided by these current partnerships is rich and can be used to answer many other questions that may help the organizations meet their goals. In addition to analyzing current data, STATCOM aims to give each organization tools and skills to use future data to further improve their decisions and evaluate the impact of those decisions. This includes any code upon request, a formal report of the analysis methods, key findings, implications, and most recently, a web-based application to analyze new data in real-time. By providing these resources, STATCOM ensures its impact will continue beyond its current collaborations, allowing the needs of the community to be better met now and in the future

6. REFERENCES

[1] Ann Arbor Area Community Foundation. *Ann Arbor Area Community Foundation*, 2018. https://www.aaacf.org.

[2] Ann Arbor Area Community Foundation. *Glacier Hills and Trinity Health Senior Communities Donate 18M to Transform Support for Vulnerable Seniors in Washtenaw County*, 2018. https://www.aaacf.org/About/News-Updates/All-News/ArticleId/22/glacier-hills-and-trinity-health-senior-communities-donate-18m-to-transform-support-for-vulnerable-seniors-in-washtenaw-county.

[3] A. Bowling, M. Hankins, G. Windle, C. Bilotta, and R. Grant. A short measure of quality of life in older age: The performance of the brief older people's quality of life questionnaire (opqol-brief). *Archives of Gerontology and Geriatrics*, 56(1):181 – 187, 2013.

[4] Data Science For Social Good. *Data Science For Social Good Projects*, 2018. https://dssg.uchicago.edu/projects/.

[5] DataKind. *DataKind Projects*, 2018. https://www.datakind.org/projects.

[6] Detroit College Access Network. *Detroit College Access Network*, 2018. https://www.detroitcan.org.

[7] Edward Ginsberg Center. *Community Technical Assistance Collaborative (CTAC)*, 2018. https://ginsberg.umich.edu/ctac.

[8] Edward Ginsberg Center. *Unlocking the University of Michigan in Service to the Public Good*, 2018. https://ginsberg.umich.edu.

[9] food for thought. *Food Assistance for the Toledo Area*, 2018. https://feedtoledo.org.

[10] food for thought. *Food For Thought 2017 Annual Report*, 2018. https://issuu.com/feedtoledo/docs/foodforthought_2018annualreport_web.

[11] Genesee County Health Department. *Genesee County Health Department*, 2018. https://gchd.us.

[12] M. P. Johnson. Data, analytics and community-based organizations: Transforming data to decisions for community development. *ISJLP*, 11:49, 2015.

[13] Michigan Broadband Cooperative. *Michigan Broadband Cooperative Who We Are*, 2018. http://www.mbcoop.org/who-we-are/.

[14] Ozone House. *Mission and History*, 2018. http://ozonehouse.org/about-us/mission-history/.

[15] Ozone House. *The Need*, 2018. http://ozonehouse.org/need-impact/need/.

[16] Ozone House. *Ozone House*, 2018. http://ozonehouse.org.

[17] STATCOM. *STATCOM-University of Michigan*, 2018. https://sph.umich.edu/biostat/statcom/.

[18] Statistics Without Borders. *Get Involved Looking to get involved with Statistics Without Borders?*, 2018. http://community.amstat.org/statisticswithoutborders/getinvolved.

[19] The Children's Center. *The Children's Center- Detroit*, 2018. http://www.mhweb.org/wayne/childctr.htm.