Domestic Violence, Sexual Assault Agency Statistics in North Carolina

By Aleah Howell, Ying Liang, and Debahutee Rout

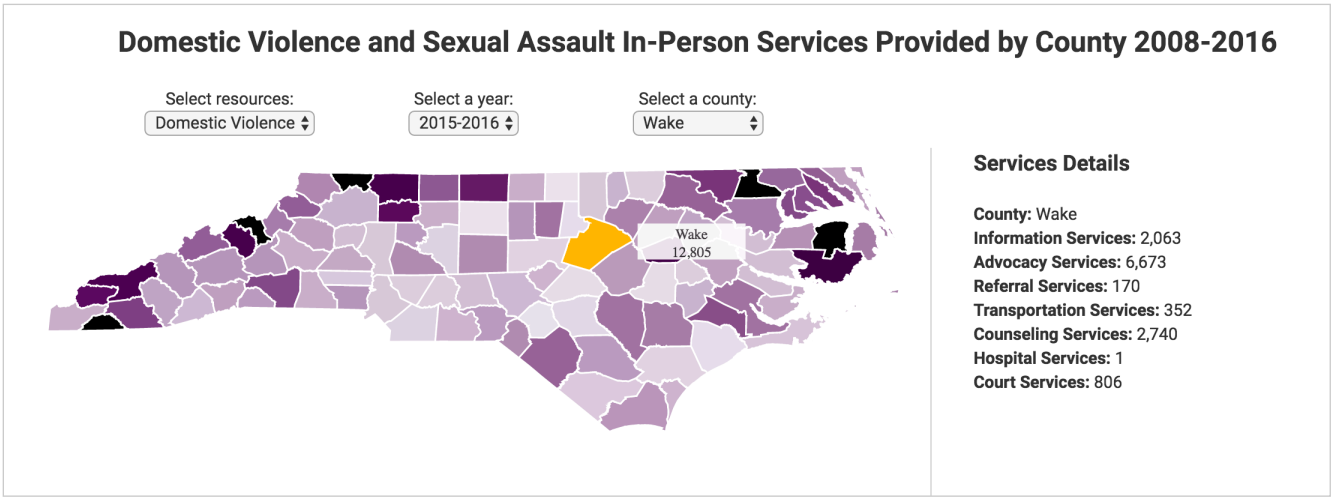

Fig. 1. NC State map county breakdown of Domestic Violence and Sexual Assault Programs. The figure represents the total number of services offered by each county between 2008 and 2016 on an interactive graph. The different shades of purple on the map reflect the density of services each county provides. For any user, the map is intended to serve as a platform to begin browsing data for different counties with respect to sexual assault and domestic violence.

Abstract — The NC Council for Women (CFW) produces the only county-level data on sexual assault and domestic violence services in North Carolina. Domestic violence programs and sexual assault programs are funded by the council. The services that are offered include shelter services, counseling, hospital services, 24-hour hotline servies, transportation, court, and advocacy services. However, no visual analysis of this data currently exists. This project aims to build a visual interface to present and contextualize NC domestic violence and sexual assault agency data. A visualization of the various services offered will allow users to assess how the state fares with respect to domestic violence and sexual assault programs throughout North Carolina's 100 counties. The visualization created with JavaScript, D3, and simple statistics is intended to give a comaparative view across counties. The interactive graphs will allow state officials, service providers, policy makers, and the CFW to better understand the data collected and identify trends in service use. A visual analysis of this data will aid officials in identifying survivors' needs and location; therefore allowing them to focus resources more effectively and provide enhanced help for survivors of domestic and sexual violence. Ideally, visualization of this data will allow new ways for the CFW and the state to act upon the data available to them.

Index Terms — sexual assault, domestic violence, interpersonal violence, advocacy, advocacy services, visualization

1 Introduction

Sexual violence, stalking, and intimate partner violence are public health problems known to have a negative impact on millions of people in the United States each year. About 20 percent of women and 2 percent of men have been raped during their lifetimes. According to the CDC, on average, nearly 20 people per minute are physically abused by an intimate partner in the United States. During one year, this equates to more than 10 million women and men. On a typical day, there are more than 20,000 phone calls placed to domestic violence hotlines nationwide. Human service programs fulfilled by the NC Council for Women (CFW) are essential in providing support for survivors of violence. The council compiles semi-annual client services statistical data for domestic violence and sexual assault services utilized by women, men, and children seeking support, information, counselling and advocacy. CFW monitors the delivery of services and produces the only county-level data on client services in North Carolina. However, no visual analysis of this data currently exists.

The visual analytics system that this project has developed will help state government or any interested user identity people affected by violence and recognize what kind of services they are receiving. An analysis of this data will help the users better identify survivors' needs, location and program use. The represented graphical representations may help the CFW to make changes to programs wherever necessary. While the ultimate goal is to stop violence before it begins, survivor resource programs are

- *Aleah Howell is a master's student in the School of Media and Journalism at UNC-Chapel Hill. E-mail: aahowell@live.unc.edu.*
- *Ying Liang is a master's student in the School of Information and Library Science at UNC-Chapel Hill. E-mail: lying910@live.unc.edu*
- *Debahutee Rout is a master's student in the School of Information and Library Science at UNC-Chapel Hill. E-mail: deb1@live.unc.edu*

an important step toward this goal. This project aims to further their purpose by finding trends and helping communities work toward healing and prevention. This project aims to present and contextualize NC domestic violence and agency statistics by asking questions such as:

1. How many people in each county are using domestic violence and sexual assault services?
2. Where are the services used most? Is there a geospatial trend?
3. Does one gender and/or race use services more than others?
4. Has program use changed over time?

2 Exposition

The North Carolina Council for Women is an advocacy division within the North Carolina Department of Administration (DOA), which acts as a business manager for the N.C. state government. In addition to overseeing operations tasks, the department provides advocacy programs (CFW being one of them). CFW provides funding and support to domestic violence, sexual assault and human trafficking programs, the Women and Girls Initiative and the Batterer Intervention Program.

The Council has a repository of data for sexual assault and domestic violence programs starting from the year 2004 to 2016. On the most basic level, the data is divided into domestic violence agency programs and sexual assault agency programs. Each dataset includes data for all N.C. counties with respect to the various different services offered. The data first gives the county population, the total number of hot-line calls, and the total number of clients for each county. The next section of the data shows age, race, and gender data for clients in each county. The following sections provide number of services provided, number of support groups, presentations and trainings, and volunteer hours for each county. For domestic violence agencies, data for shelter services is provided. For sexual assault agencies, data for type of assault and offender relationship is provided.

This project focused on utilizing this repository of agency data. The data was used to create an interactive map of the state counties focusing on domestic violence and sexual assault agency programs. Two additional graphs were created to contextualize the county data by showing aggregate use of programs between 2004 and 2016. The main focus of the project was collating data across timelines (2004-2016) and trying to put them in perspective. Data was selected and aggregated to show the big picture. This project was an effort to discover what data could be used to create a set of extrapolated values. The resulting interface provides information in an interactive mode that allows the user to arrive at new ideas or conclusions.

2 Data

The NC Council for Women requires each state-funded Domestic Violence (DV) and Sexual Assault (SA) grantee to report on client service provision. Each agency fills out a statistical report every six months that involve demographics, services provided and volunteer hours. After 2008, each agency was required to gather data on a monthly basis. The reports are sent to the NC Council for Women, where the data is checked and analyzed. In order to provide yearly data, the NC Council for Women combines two semi-annual reports from all agencies for each fiscal year. The Domestic Violence and Sexual Assault agency statistical year begins each July 1 and concludes the following June 30. Yearly data sets exist from 2004 to 2016.

CFW provides raw data and yearly aggregate data in the form of excel documents on their website. No data visualizations currently exist. A visual analytics approach will reduce the time it takes for viewers (CFW, agency directors, state officials, agency volunteers, etc.) to analyze results and share with their partners; therefore, improving programs and allowing officials to set more accurate program goals.

3 The Visualizations

Data from the CFW is well organized with domestic violence programs and sexual assault program having their own separate excel sheets with the county name, county population, county total number of calls, and county total clients for that program. However, some of the earlier datasets from 2004-2006 also included data sheets for Displaced Homemaker programs and Abuser Treatment programs. Data for these programs was not offered consistently throughout the yearly reports, therefore this project ignores those sheets and focuses only on domestic violence and sexual assault program data sheets. After group brainstorming sessions, specific groups and columns of data were selected and then put together and stored in separate sheets (one for each year).

The first two graphs created for this project show aggregate data for domestic violence and sexual assault program services in North Carolina. The purpose of these graphs is to contextualize what services are being provided and how many people are using them.

The first graph is a multiline-stacked bar graph. This graph shows aggregate data for all services provided by North Carolina agencies over time between 2004 and 2016. This graph shows the total number of people using hotline services, total number of people enrolled in support groups, and the total number of people enrolled in in-person services for domestic violence or sexual assault agency programs. There are two dropdown menus on this graph. The first allows the user to choose between DV and SA programs. The second dropdown menu allows the user to see each service individually or stacked together. As the user hovers over each bar in the graph, a tooltip shows the number of people calling or enrolled in the different services. The x-axis denotes the year and the y-axis denotes the total number of services in thousands of services. The y-axis maximum value changes depending on if the user is viewing DV or SA. The graph is plotted to show the numbers across the time period of 2004-2016.

The bars and their height convey a visual of how many people have enrolled for or accessed the different services. The user can compare services within a program as well as between DV and DA programs. With the help of this graph the CFW or any interested user can make a comparative study of how many people have been enrolled in services across years and determine if there has been an increase or decrease in numbers for any specific service. This visualization can help policy makers decide which service needs more attention and budgeting than the rest and look into probable causes as to why one service is more popular or even less popular than the other. Figure 2 shows the graph programmed to show

different services accessed by people enrolled in a DV or SA program. Figure 3 shows the graph after "in-person services" was triggered in the dropdown.

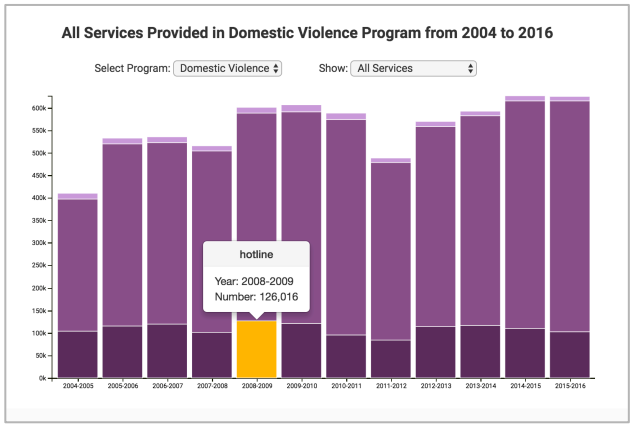

Fig. 2. Stacked Bar graph of all services provided in Domestic Violence and Sexual Assault programs from 2004 to 2016.

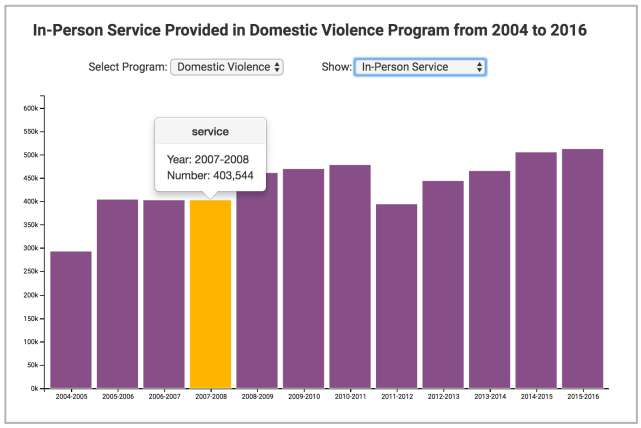

Fig. 3. Bar graph shows in-person domestic violence services only.

The third graph (see figure 4) is a multiline chart that shows the total number of clients enrolled in domestic violence programs and the sexual assault programs. This graph shows in-person service data, excluding hotline services. The x-axis denotes the year and the y-axis denotes the number of clients. Hovering over the dotted data points for each year renders a tooltip that shows the total number of clients. The graph provides a dropdown menu for the user to view data for sexual assault program, domestic violence program, or both. The visualization can be used for comparative study of how many people used sexual assault services and how many used domestic violence services. It also allows the user to see which year the data was above or below the consecutive year figures.

Since the one of the project's goals was to capitalize on the availability of county-level data, it was decided that a geospatial representation of the data would be fitting. A North Carolina map was created (see Figure 1). The map shows all 100 N.C. counties. The first drop down box on the map provides an option for selecting an agency program (sexual assault or domestic violence), the second drop-down provides an option of selecting a year, and the third drop-down box provides the option of selecting a specific county on the map (which then highlights the county the user is searching for).

Hovering over each county triggers a tool tip that displays the county name and the total number of services provided in that county. We normalized the data for population by dividing the total number of services in each county by county population. This provided an average number of services provided for each individual in that county. The color scale of the map(s) were rendered based on this normalized data. Hovering over a county on the map also renders service details i.e. the types of different services implemented and the number of people enrolled in each service. These services include: information services, advocacy services, referral services, transportation services, counselling services, hospital services, court services, and an 'other' category that was not specified.

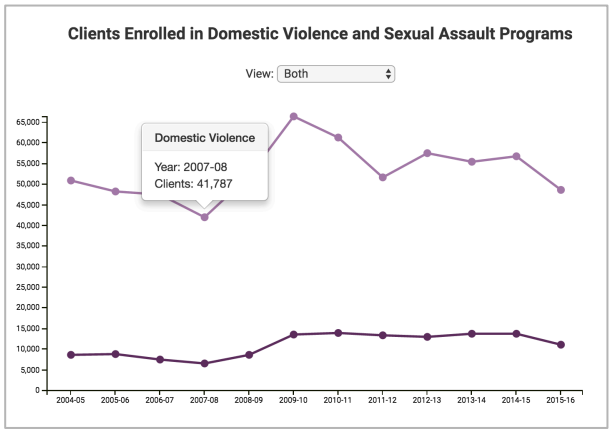

Fig. 4. Multi-line chart of all in-person services provided in DV and SA programs between 2004 and 2016.

4 Discussion

During its inception, the team knew that handling such a large amount of data within the required deadline would not be a feasible task and would limit approaches to make a success interactive. Meetings were scheduled to determine which data to focus on and analyzed and which data to leave out. Even though we understood that all data categories give a wholesome picture to program services, the inclusion of all categories of data was beyond the scope of our abilities with the time restraint.

After many discussions, sketches (see Figures 5-7) and initial prototyping, we decided to narrow down our focus to the two programs funded by the Council for Women i.e. domestic violence and sexual assault and the services they provided.

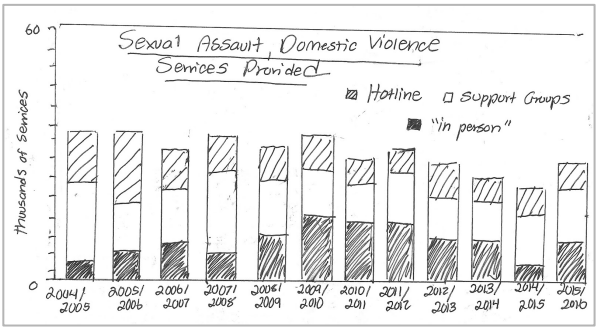

Fig. 5. Original brainstorm sketch for stacked bar chart showing total number of DV and SA services over time.

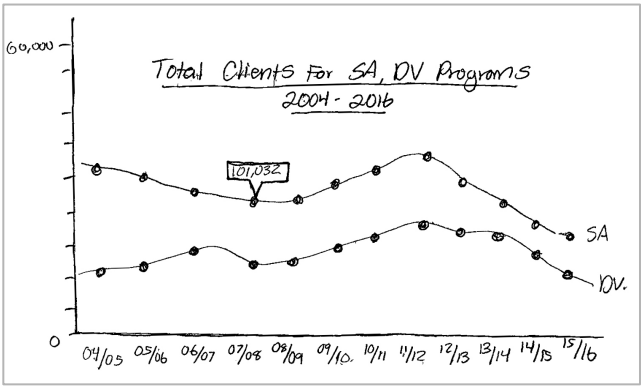

Fig. 6. Original brainstorm sketch for multi-line chart showing total number of clients for DV and SA agency programs.

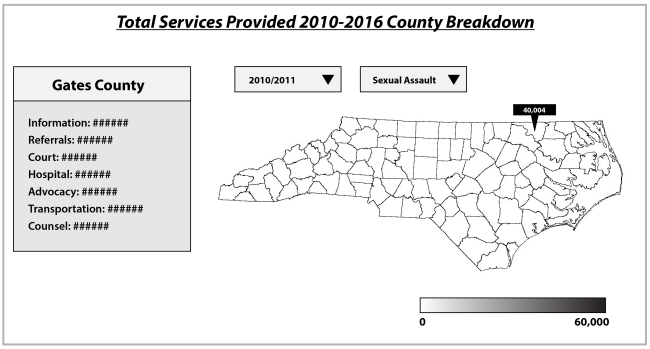

Fig. 7. Wireframe sketch created in Adobe Illustrator to brainstorm the map interface and functionality.

These figures represent the sketches we created in the process of brainstorming in our weekly meetings. With the dataset being narrowed down to two major topics, the next task was to focus on what data to use within the two main datasets (DV and SA). The challenge was how to portray that data in a succinct yet thorough manner that could show some basic comparisons between the two programs as well as within a program, across a span of 12 years. In the original data sheets, each program is discussed with respect to age, race, gender, services provided, number of support groups, shelter services etc. However, some data categories are inconsistent between programs. For example, across programs the age field is divided into different age ranges. The age category was divided into <25, 25-34, 35-44, 45-54 for domestic violence versus being divided into 0-12, 13-17, 18-25, 26-40 for sexual assault. Also, domestic violence programs provide data on shelter services while sexual assault programs provide information on type of assault.

Within the 13 weeks provided to complete the visual prototype and considering the programming skill level of our three-person team, we had to cut back on our original goals. In the end, the team decided to focus on visualizing (1) Overall, how many services were provided 2004 and 2016, (2) Overall, how many people are using these services 2004 and 2016 and (3) Which counties are using more or less services than others, and what services are they providing. The main variables we chose to change were (1) year, (2) program, DV or SA, (3) type of service and (4) county. Our team believed it was important to contextualize the data before diving deeper into specifics such as age, gender, and race. Hence, the first two graphs provide a temporal analysis of aggregate services and clients.

Tooltips and colors were basic design elements that we chose to aid our visualizations. The different shades of color on the map was an effort to give the user basic information quickly without having to interact. Tooltips are an effective shortcut for the user to get information without being required to click multiple times. The project utilized three different types of graphs: stacked bar graph, multiline chart, and an interactive map. Each graph used drop downs for the user to filter out information. A purple color pallet was chosen because purple represents domestic violence awareness, just as teal represents sexual assault awareness. The NC Council for Women also uses purple frequently on their website.

4.1 Teamwork and Team Meetings

With the task at hand, the team met once a week to discuss progress and allocate tasks. We analyzed the data, brainstormed visualizations, created solutions to the initial problems, and programmed visualizations for the final prototype. Each team member was assigned a graph to program based on programming experience. Throughout the project, each weekly meeting started with taking stock of our progress and what we needed to achieve for the next week's meeting. Each team member worked with the others and gave feedback/input until the whole team came up with a unanimous decision. With constant collaboration and teamwork we developed three different graphs that organized and represented the data. Each individual member carried out the responsibility of completing one graph representation and after the visualization was complete, all data was put together for the presentation and final report.

4.2 Challenges and Reflection

During the course of this work, we faced couple of hurdles, some of which we had anticipated as a team and some of which were a surprise. There were hurdles that each member faced while programming the graphs. Our team began this project as beginners at D3. Looking back at class exercises, researching similar work done by other developers, studying D3, and seeking help from peers clarified many of our questions and helped us solve our challenges.

The first challenge we faced was simply the amount of data we had. Including specific data for each of North Carolina's 100 counties, there were thousands of data elements for each year, with the whole dataset spanning 12 years. Every number felt important as we considered the big picture. To solve this problem, we analyzed the data and chose the variables we believed to be the most important for our visual prototype.

Another challenge we faced was learning the proper CSV format for loading tabular data into html. The cross origin resource sharing was also a challenge.

Creating the North Carolina map was one of the most challenging parts of this project. We first tried rendering the N.C. map with Datamaps, which are customizable SVG map visualizations in a JavaScript file using D3. However, adapting Datamaps to create a North Carolina map proved to be more difficult than initially thought. Datamaps has pre-written JavaScript and D3 code that renders world maps and continent maps, including a map of the United States. We found it difficult to interpret and adapt another developer's code for our purposes. Therefore, we created the North Carolina map from scratch based on other tutorials such as Mike Bostock's "Let's Make A Map". Another challenge with creating the map was creating and reading the correct type of json file. In order to render North Carolina, we needed a json file that rendered North Carolina's counties. We ran into many issues with converting the Geogson (geometry) file of N.C. to a Topojson (line arcs) file. Topojson is an extension of geojson that encodes typology. Overall, creating the North Carolina map was a significant learning curve.

Another challenge was the scale of the y-axis on the stacked bar chart. In the stacked bar chart, the height of each stacked bar is the sum of components. However, the components of each bar are hard to compare, as they do not have an equal point on the y-axis. The stacked bar visualizes two programs, but one (domestic violence) has much larger numbers than the other (sexual assault). The increase or decrease of the data is very similar between programs over the years. However, the two programs have different y-axis scales that fit the data so the bars have similar height. Therefore, if the user doesn't notice the change in y-axis, they will think the two programs have similar numbers.

We decided to implement a changing y-axis because if we had used a fixed y-axis, it would have been very difficult for the user to view sexual assault data. The bars would have been so small in comparison to the domestic violence data that the viewer would not have been able to discern between the components of the stacked bars or been able to make a clear comparison. Therefore, the bar chart is more about showing the trend in data over time between the two programs rather than which program has more or less services — because sexual assault very obviously has fewer services provided overall.

5 Conclusions and Future Work

This project was a valuable learning experience. Learning and understanding D3 to create dynamic visualizations was a challenging experience. We learned how to select significant data from a vast dataset and use JavaScript to create visualizations. The project was a challenge for our team who had little to no knowledge of how to use D3 and simple statistics. After completing the prototype, our team is more familiar with D3, data joins, functions, and incorporating SVG formats into a program.

D3 can make it easier for the developer and user to interpret data and transform it into meaningful visualizations. Although this project only scratches the surface of what D3 can do, this project has allowed us to understand how we can use JavaScript to transform large data sets into aesthetically pleasing and functional interactive visualizations.

Considering the time constraints of the project and the programming skill level of our three-person team, we had to cut back on our original ambitious goals for this prototype. Moving forward, this project would benefit from additional visualizations. Specifically, within the North Carolina map visualization, each county could have breakout visualizations that show age, race, and gender distributions for domestic violence and sexual assault programs. It would also be valuable to further investigate minimum and maximum y-axis values on graphs that show such polarized datasets. User testing of the prototype would also be valuable to help questions such as: Is data presentation clear? What information is missing? What are the main take-aways from viewing the data?