

Exploratory Bivariate and Multivariate Geovisualizations of a Social Vulnerability Index

Georgianna Strode
Florida State University
gstrode@fsu.edu

Victor Mesev
Florida State University
vmesev@fsu.edu

Susanne Bleisch
Fachhochschule Nordwestschweiz
susanne.bleisch@fhnw.ch

Kathryn Ziewitz
Florida A & M University
kathryn.ziewitz@famu.edu

Fennis Reed
fennis.reed@gmail.com

John Derek Morgan
University of West Florida
jmorgan3@uwf.edu

In the United States, the Centers for Disease Control and Prevention (CDC) is the national agency that conducts and supports public health research and practice. Among the CDC's many achievements is the development of a social vulnerability index (SVI) to aid planners and emergency responders when identifying vulnerable segments of the population, especially during natural hazard events. The index includes an overall social vulnerability ranking as well as four individual themes: socioeconomic, household composition & disability, ethnicity & language, and housing & transportation. This makes the SVI dataset multivariate, but it is typically viewed via maps that show one theme at a time. This paper explores a suite of cartographic techniques that can represent the SVI beyond the univariate view. Specifically, we recommend three techniques: (1) bivariate mapping to illustrate overall vulnerability and population density, (2) multivariate mapping using cartographic glyphs to disaggregate levels of the four vulnerability themes, and (3) visual analytics using Euler diagrams to depict overlap between the vulnerability themes. The CDC's SVI, and by extension, vulnerability indices in other countries, can be viewed in a variety of cartographic forms that illustrate the location of vulnerable groups of society. Viewing data from various perspectives can facilitate the understanding and analysis of the growing amount and complexity of data.

KEYWORDS: social vulnerability; multivariate mapping; glyphs; Euler diagrams

INTRODUCTION

THE MOTIVATION FOR THIS project was a request from the Sustainability Institute at Florida Agricultural and Mechanical University (FAMU) to locate vulnerable populations within the Florida panhandle. The Sustainability Institute planned to work with pockets of vulnerable populations to strengthen their capacity to anticipate and respond to weather-related hazards. While previous efforts at disaster planning preparedness had been top-down and county-centered, the Institute wanted to identify specific populations of vulnerable residents to select pilot communities for a project to strengthen capacity at the neighborhood scale. At the project's onset, the Institute used social vulnerability index (SVI) data from the Centers for Disease Control and Prevention (CDC) to identify

vulnerable populations, but found the volume of data difficult to interpret. They requested custom maps to aid in the visualization of locations of vulnerable residents, leading to the project we describe in this paper.

Social vulnerability can be broadly defined as the risk and potential loss to society as a result of both naturally occurring and human-induced disasters (Cutter 1996). In the United States, the CDC describes social vulnerability as referring to "the resilience of communities when confronted by external stresses on human health" (CDC 2018). When a disaster strikes, socially vulnerable people are more likely to be adversely affected, require more time to recover, and are at a greater risk of injury and death



© by the author(s). This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0>.

(Wright et al. 1979; Cutter, Boruff, and Shirley 2003; Peacock et al. 2008). Social vulnerability can be effectively addressed by providing public services in a timely manner to reduce human suffering and to minimize economic losses (Flanagan et al. 2011). Information on social vulnerability can aid in resource planning and social equity decisions and, when used in conjunction with other data, can highlight transportation, housing, recreation, educational, nutritional, medical, and other needs.

The CDC has compiled social vulnerability data using information from the 2012–2016 American Community Survey (ACS). Their SVI is freely available to the public in a digital format, which makes for easy input into a GIS. Typically these data are utilized for univariate mapping, where each vulnerability theme is represented by a single map. However, maps illustrating two or more variables can be employed to convey more complex relationships. The SVI data are multivariate in nature and as such are amenable to the use of more complex geovisualization techniques to illustrate patterns of social vulnerability.

Our objectives in this project were threefold. One, to demonstrate how bivariate maps can combine SVI with population density in order to allow readers to distinguish between urban and rural areas. Two, to employ multivariate maps when illustrating multiple types of vulnerability simultaneously, using graduated glyphs. And three, to show how Euler diagrams can reveal overlaps between the four themes of vulnerability. These three geovisualization methods provide a broader perspective on data

interrelationships and can identify pockets of specific vulnerability types.

While mapping more than one variable simultaneously allows the representation of composite trends and relationships, this complexity can come at the cost of visual clutter. To address this, we employ multivariate symbols called *glyphs* to help us fulfill our second objective. Glyphs are commonly used to simultaneously represent several pieces of data for the same location, using one or more visual marks and visual variables. For example, the height of a rectangular symbol might represent one variable, its width another, and its color a third (Ward 2008). One of the better-known types of glyphs is the Chernoff face, where facial features such as nose length or eyebrow orientation indicate data values (Chernoff 1973).

This challenge of making non-spatial multivariate information readable using visual analytics was our third objective. Univariate maps often use analytics such as bar charts to complement the map and further explain spatial data. Much of the data we work with is multivariate, originating from matrices, spreadsheets, sensors, and mathematical computations (Ward 2002). The larger the number of variables, the more difficult it becomes to detect, classify, or measure features and relationships. To work through these difficulties, we explore a more complex visual analytics technique, the Euler diagram, which complements the simple methods for spatial and temporal representation of data that are typically available through GIS.

SOCIAL VULNERABILITY

IN 2018, THE Federal Emergency Management Agency (FEMA) reported 124 major disasters in the United States, far more than the 16 reported in 1988 (FEMA, n.d.). Some have linked this increase with climate change, as extreme weather events become more frequent and intense (Haines et al. 2006; Wolkin et al. 2015). Research by the National Center for Environmental Health and the CDC has indicated that emergencies stemming from storms would be handled more effectively if accurate information on vulnerable populations were available. Of particular importance is defining which vulnerable groups are at risk, where they live and work, and the fastest ways to communicate with them (CDC 2015).

The risk posed by some environmental threats, such as tornadoes, hurricanes, and floods, is a function of their frequency and of societal exposure, determined by an assessment of the number of people and properties located along storm tracks or floodplains. Vulnerability to these threats is also linked with socio-economic indicators, such as the area's proportion of non-English speakers, households earning incomes below the poverty line, and mobile homes, as well as its property values. Indeed, the likelihood of a fatality is 20 times greater for mobile homes than for fixed structures (Brooks and Doswell 2001). While the elderly population of coastal Sarasota County, Florida, faces a considerable threat to property from hurricanes,

their overall risk is low because they tend to be affluent retirees, and are therefore less socially vulnerable. In contrast, elderly populations living further inland in Florida may face fewer environmental threats but tend to be more economically vulnerable (Wang and Yarnal 2012).

Understanding where socially vulnerable populations exist, via SVI data, can help with allocating resources more effectively during the four phases of the disaster cycle: preparedness, mitigation, response, and recovery. For example, mitigation planning requires that officials identify socially vulnerable populations in order to provide increased assistance during a disaster. Effective mitigation decreases human and economic losses by providing social services after a disaster (Flanagan et al. 2011). SVI data can be used to help estimate the volume of supplies needed to sustain victims, the number of emergency personnel needed, the areas where emergency shelters are to be located, and the communities that will need continued post-disaster support. It can also be used to plan evacuation strategies, particularly for people with special needs (CDC 2018).

There are many vulnerability indices, each developed for various purposes using different types of data. For example, Cutter, Boruff, and Shirley’s (2003) social vulnerability index (SoVI®), which was originally designed at the county level, is used to quantify old-age social vulnerability, as well as environmental hazards (Tarling 2017). In contrast, the CDC developed its SVI as a public tool to identify populations requiring more resources, for improving the effectiveness of disaster preparedness (Flanagan et al. 2011). Both indices are publicly available and can be used with basic GIS tools. We chose to work with the CDC’s SVI data, although our methods would work with any index with a limited number of dimensions (in this case fewer or equal to 4).

The CDC’s SVI consists of 15 variables, extracted from the 2012–2016 ACS organized into four themes: socioeconomic status (4 variables), household composition & disability (4 variables), ethnicity status & language (2 variables), and housing & transportation (5 variables). Variable descriptions are listed in Table 1. The 15 census variables are grouped into four themes. Each census tract receives a ranking for each of the four themes and an overall ranking

Census Variable	Description (all variables are estimates)
Socioeconomic Status	
Below poverty	Persons below the poverty line
Unemployment	Civilian (age 16+) unemployed
Income	Per capita income
No high school diploma	Persons (age 25+) with no high school diploma
Household Composition & Disability	
Aged 65 or older	Persons aged 65 or older
Aged 17 or younger	Persons aged 17 and younger
Older than Age 5 with disability	Civilian noninstitutionalized population with a disability
Single-parent households	Single-parent households with children under 18
Ethnicity Status & Language	
Minority	All persons except white, non-Hispanic
Speaks English “less than well”	Persons (age 5+) who speak English “less than well” (i.e., “not well” or “not at all”)
Housing & Transportation	
Multi-unit structures	Housing in structures with 10 or more units
Mobile homes	Mobile homes
Crowding	At household level (occupied housing units), more people than rooms
No vehicle	Households with no vehicle available
Group quarters	Persons in institutionalized group quarters

Table 1. The census variables and themes of the CDC SVI.

from highest to lowest in terms of their applicability for measuring vulnerability.

Socioeconomic status includes income, employment, poverty, and education variables. The poor are less likely to have the income or assets needed to prepare for or recover from a disaster (Morrow 2002). They may potentially lack homeowner's or renter's insurance, thus making property replacement more expensive (Tierney 2006). Unemployed persons may lack health or life insurance (Brodie et al. 2006). Those with higher education are more likely to access and heed information related to disaster preparation and recovery (Tierney 2006).

The **household composition and disability** theme considers the population of those who are at an at-risk age (under 18 or over 65), are single parents, or have a disability. Children often cannot protect themselves in a disaster, and elders living alone or with physical or cognitive challenges are also more vulnerable (Morrow 2002). Single-parent families have more risk, as the caretaker responsibilities cannot be shared between two parents (Flanagan et al. 2011).

The **ethnicity and language** theme considers race, ethnicity, and English-language proficiency. Social and economic marginalization, including real estate discrimination, has left certain populations more vulnerable in all stages of a disaster. Lack of English proficiency can hinder disaster-preparedness communication (Morrow 2002).

The **housing and transportation theme** includes housing structure, crowding, and access to vehicles. Housing is tied to personal wealth, and lower-income persons are more likely to live in poorly constructed homes (Tierney 2006). Mobile homes are not designed to withstand severe weather or flooding, and are frequently located away from interstate highways or public transportation (Flanagan et al. 2011). In more densely populated areas, multi-unit housing poses risks for tenants as emergency exits and stairwells can become overcrowded, and large numbers of people exiting buildings into the streets can complicate safe and orderly evacuation plans (Cutter, Boruff, and Shirley 2003; Tierney 2006). Automobile ownership is lower in urban and less wealthy areas, causing difficulties with evacuation plans (Morrow 1997). Meanwhile,

those who do own vehicles can face severe traffic congestion (Flanagan et al. 2011). Persons residing in group quarters (college dormitories, prisons, and nursing homes) face evacuation challenges due to facility understaffing during emergencies and the need for specialized vehicles (Flanagan et al. 2011).

To construct the SVI data, every census tract with non-zero population is ranked, from highest to lowest vulnerability, according to each of the 15 variables. For example, each census tract in a state (or in the United States, if a national assessment is being conducted) is ranked according to its per capita income. For each variable, the tract is then assigned a percentile rank, reflecting the proportion of tracts that its particular score is greater than or equal to. For example, a census tract receiving a percentile ranking of 0.237 (or 23.7%) is more vulnerable than 23.7% of census tracts with regard to that particular census variable.

Next, the percentile ranks for every variable in a theme are summed together for each tract. The tracts are then ranked according to those sums, and given new percentile ranks reflecting their vulnerability according to the theme. Finally, the overall index is calculated by summing the percentile ranks for each of the 15 variables, and then ranking those sums. Each of the five summary datasets (the percentile ranks for the four themes and the overall index) is then classified into quartiles (Flanagan et al. 2011; Flanagan et al. 2018). The CDC's SVI data are provided in three formats: an online viewer on the CDC website, downloadable GIS data at the census tract scale, and premade county maps in PDF format.

There are limitations to this method of calculation. Ranking census tracts according to their vulnerability achieves a sequential order but does not account for the size of gaps in data values. As a result, there is no distinction between large or small data gaps. The four vulnerability themes are calculated from different numbers of census variables, meaning each of those variables has a different level of influence on the outcome. The two variables of the Ethnicity Status & Language theme, for example, each carry more weight than the five of the Housing & Transportation theme. Finally, the quartile classification system cannot adequately distinguish between data values close to the breakpoints. Consequently, data values of .249 and .25 will be placed in two different classes.

MULTIVARIATE REPRESENTATION

ONE OF OUR OBJECTIVES was to represent vulnerability data using multivariate methods. To do this, we employed multidimensional glyphs to convey several datasets simultaneously. Typical attributes of glyphs include shape, size, orientation, position, direction, and magnitude, as well as color, texture, and transparency (Ward 2002). Color and size are more dominant than other attributes and can help direct the user's attention (Borgo et al. 2013). Glyph placement within the represented geographic area is important, as glyphs should not overlap (Ward 2002; McNabb and Laramee 2019). Glyphs, as compared to other symbols, allow patterns involving two or three data dimensions to be visualized clearly, as long as they fall within the capabilities of human visual perception. However, glyphs are constrained by the number of data records that can be visualized, as large data sets may require small glyphs in order to fit them all on the page or screen, which will make pattern recognition more difficult (Ward 2008; Fuchs et al. 2017).

VISUAL ANALYTICS AND EULER DIAGRAMS

Most of the data we work with are multivariate, whether they originate from matrices, spreadsheets, sensors, or mathematical computations. The larger the number of variables, the more difficult it becomes to detect, classify, or measure features and relationships. Meeting this challenge of making non-spatial multivariate information more accessible and understandable was our third objective.

Geovisualization relies upon the human perceptual system to guide analysis. Ward's (2002) mantra "I'll know it when I see it" reflects the idea that we often do not know what we are looking for when presented with a large data set (Shneiderman 1996). Geovisualization techniques facilitate exploratory data analysis, including the identification of visual trends, which may lead to explanatory data analysis if trends are confirmed with statistical relationships (Ward 2002). To aid in geovisual exploration, GIS has incorporated visual analytics as a complement to the simple spatial and temporal representation of data. Common visual analytics such as pie and bar charts are often used to provide a quantitative overview of the data to complement the spatial distributions shown on maps. Some other examples of visual analytics include the analysis of tweets

(Andrienko et al. 2013), box plots (Willmott, Robeson, and Matsuura 2013), and custom isolines (Collins, Penn, and Carpendale 2009). Many of these analytics work well for univariate data but cannot show relationships between two or more types of data. Visual analytics using bivariate or multivariate data have been used in GIS, such as Sankey charts to visualize the relationships between land use and land cover (Strode et al. 2019) and Euler diagrams (Lemon and Pratt 1997).

The visual analytic concept of *set membership* is of key interest to us in our project. We can group data items into sets based on their specific properties and, since members can belong to more than one set, sets can potentially overlap (Alsallakh et al. 2013). Three relationships between sets are possible: containment (one set falls entirely within another), exclusion (sets have no members in common), and intersection (sets have some members in common). Organizing information into sets lets us ask questions about the relationships between members. For example, by sorting persons into sets representing their club membership, we may ask whether certain clubs are exclusive of each other; by sorting products into sets based on their features, we may ask whether a product feature is always in combination with another product feature (Alsallakh et al. 2013).

The Euler diagram is a traditional, and easily understood, method for visualizing set membership. It is a variation of the Venn diagram. While the Venn diagram maintains its size and shape regardless of data quantity, the Euler diagram is more fluid, adjusting the size and shape of the graphic to match data quantities. Euler diagrams quantify areas of overlap using proportional boundaries, with empty data overlaps removed to give a more intuitive understanding of multiple data relationships. Euler diagrams are the most common representation of set-type data, but are limited in the number of sets they can represent (Alsallakh et al. 2013).

In this study, we were interested in using Euler diagrams to model set membership in order to see the relationships between the four vulnerability themes. Do certain vulnerability themes always occur together? Is one theme excluded from the others? Geospatial representation of the four vulnerability themes reveals their spatial locations but cannot reveal the relationships between them.

METHODOLOGY AND RESULTS –

WE BEGAN TO EXPLORE the CDC SVI data by making basic maps of our study area. Figure 1 shows the CDC’s overall SVI ranking for each census tract in Alabama, Georgia, and Florida, while Figure 2 shows each of the four SVI themes separately for the city of Tallahassee, Florida. Univariate maps like these are the typical way that the CDC currently represents the SVI. However, they do not account for population density, which is key to identifying concentrated groups of vulnerable populations. Two areas of equal vulnerability and equal size could have varying population densities; in a disaster, too many or—more importantly—too few emergency responders may be directed to risk areas.

BIVARIATE METHOD SHOWING SVI AND POPULATION DENSITY

To create maps that conveyed both vulnerability level and population density, we decided to create a grid of 1km vector cells, with each cell containing both density and SVI data (Zhou et al. 2017). We would then be able to vary the color of the cells to indicate the level of vulnerability, and their size to indicate population density; Figure 3 features an example. This grid-based representation would alleviate some of the well-documented problems of choropleth maps, such as inconsistently sized areal units and “data cliffs” where values abruptly change at an area’s border (Langford 2013; Mennis 2003; Martin, Lloyd, and Shuttleworth 2011; Jia, Qiu, and Gaughan 2014).

To prepare the data for this representation, we began with a grid system. We chose to use the US National Grid (USNG) because it is a recognized standard (Federal Geographic Data Committee 2001; FEMA, n.d.; US Fire Administration 2015), usable anywhere on Earth (despite its name), scalable at multiples of ten, and a vector system, thereby allowing multiple data values to be associated with each grid cell. The gridded GIS files (US National Grid Information Center, n.d.) include population counts aggregated from cadastral-based dasymetric methods (Strode, Mesev, and Maantay et al. 2018) where possible, and from WorldPop (worldpop.org)

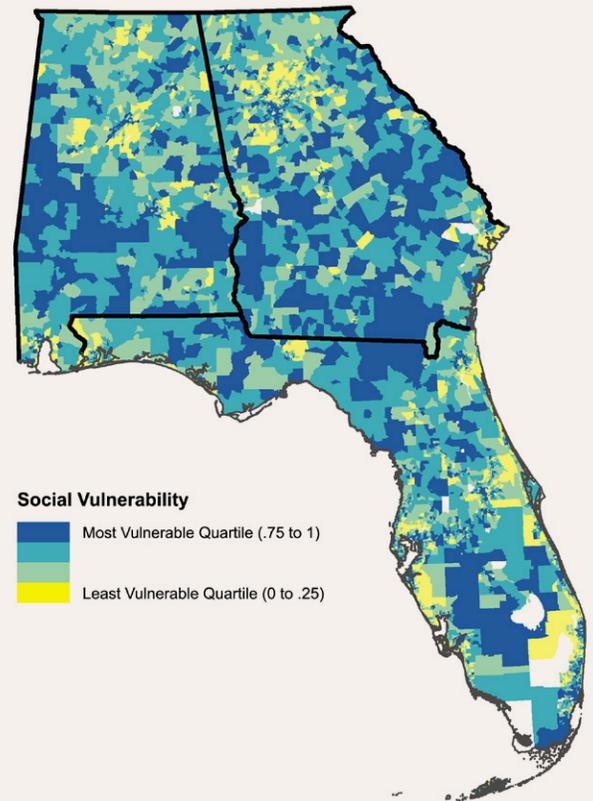


Figure 1. Univariate social vulnerability data, by census tract, for Alabama, Georgia, and Florida.

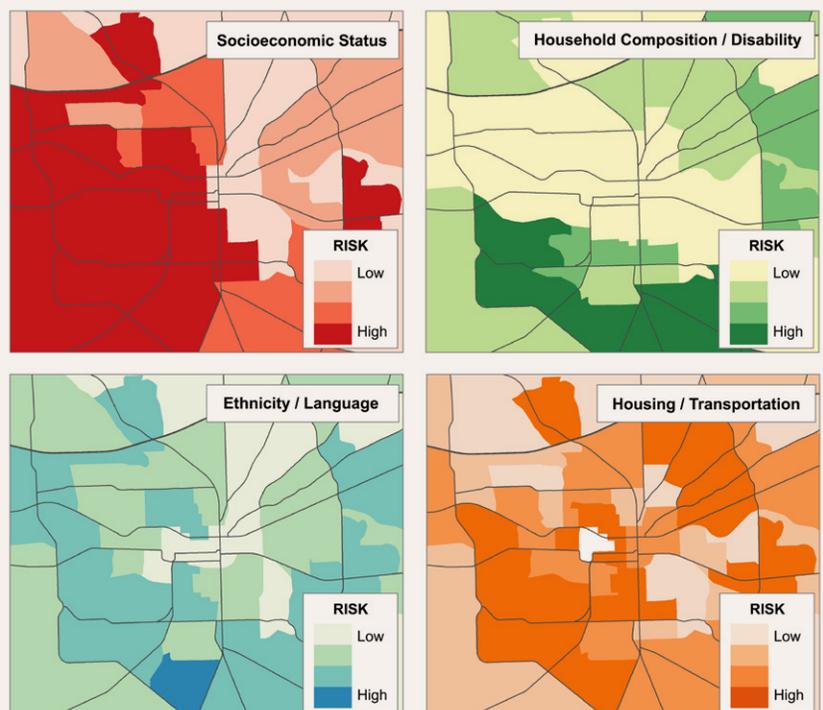


Figure 2. Univariate maps representing the risk quartiles of each of the four themes of social vulnerability (see Table 1) for Tallahassee, Florida.

raster data otherwise. The 1km population grid cells were then overlaid with the census tracts holding SVI data. For cells that intersected more than one census tract, the values were areally interpolated by the amount of land area that falls within each cell. The resulting dataset was a 1km grid, containing values for the overall vulnerability index, for each of the four vulnerability themes, and for the total population. Figure 4 illustrates the results of this process for the Tallahassee area.

The conventional univariate representation in Figure 1 can be contrasted with the bivariate representation in Figure 3. The univariate representation only displays the overall SVI, while the bivariate representation displays the SVI and population density. From a visual perspective, the bivariate map eliminates unpopulated areas and produces clusters of high population density to represent urban areas.

MULTIVARIATE MAPS OF THE FOUR CDC SVI THEMES USING GLYPHS

While bivariate maps allowed us to visualize both population density and SVI, the SVI index is made of four vulnerability themes, and we wanted to represent each of these simultaneously. To do this, we turned to glyphs. The key characteristic of glyphs is that they encode several dimensions of a single data point using one or more visual marks and visual variables. While the Chernoff face is perhaps the best-known example, Fuchs et al. (2017), in their overview of 64 papers that performed controlled studies on different glyph designs, point out that the possibilities for designing glyphs are endless. We adapted our glyph design from Bleisch and Hollenstein (2018), who used square glyphs to simultaneously represent up to four walkability indices in a regular grid.

Each of the four vulnerability indices is represented by a colored square. In making color choices, we sought four contrasting colors similar to Brewer's (1994) diverging/diverging scheme. We chose highly saturated versions of pink, green, blue, and orange, to enhance their contrast with underlying map colors. Pink and orange squares were placed diagonally to each other, as

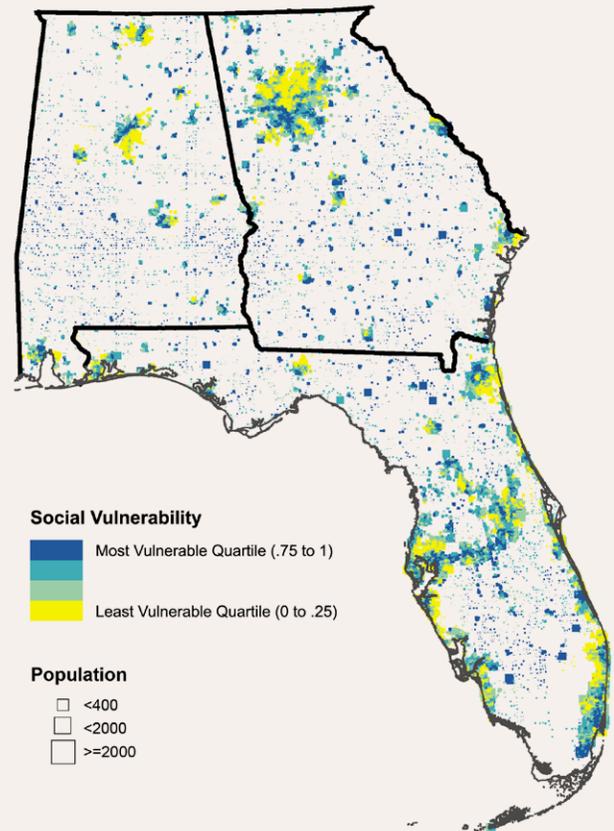


Figure 3. Bivariate map of Alabama, Georgia, and Florida showing the social vulnerability index and population density. Contrast with Figure 1.

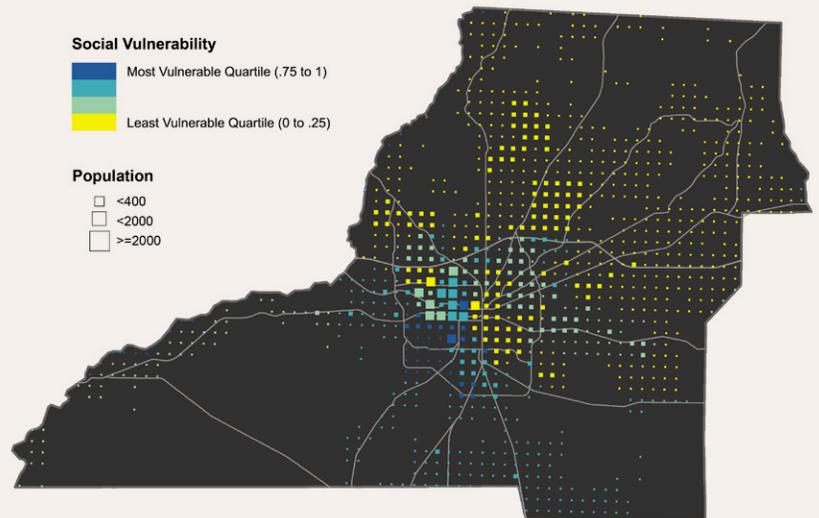


Figure 4. Bivariate map showing the social vulnerability index (represented by symbol color) and population density (represented by symbol size) for Leon County, Florida. The city of Tallahassee is represented by larger cells, while rural areas have smaller cells and unpopulated areas have no cells.

were blue and green squares, to increase the contrast with neighboring glyph cells. We then represent vulnerability scores for each of the four themes by resizing the squares (Figure 5). The lowest risk is represented by the smallest glyph and the highest risk by the largest glyph.

To ensure that the glyphs never overlap, and that they display at any scale, we used a 100m grid system within the 1km grid to construct the glyphs, as shown in Figure 5. Glyphs were drawn by adding color to some of the 100m grid cells according to the vulnerability theme and quartile. For example, 1km grid cells with a vulnerability theme ranked in quartile 1 (lowest) would add color to one 100m grid cell (1x1). If the theme was ranked in the second quartile, four cells would be required (2x2). Themes ranked in the third quartile require nine cells (3x3), and the highest quartile uses 16 cells (4x4).

Placement of the glyph colors, and thus themes, needed to be consistent. Starting from the center of the 1km grid cell, socioeconomic data are shown in the upper left, household composition/disability in the upper right, ethnicity & language in the lower left, and housing & transportation in the lower right. Multiple small cells of the same color appear as one unified square symbol. The final glyphs are composed of four contiguous squares with consistent colors and varying square sizes. Figures 6 through 10 illustrate examples of our glyph system. Figure 6 shows an overview of the variety of combinations of vulnerability levels that can be represented. Figure 7 depicts variation across a larger, hypothetical area. Figure 8 illustrates the four vulnerability themes in Leon County, while Figure 9 combines the glyphs with a population density basemap.

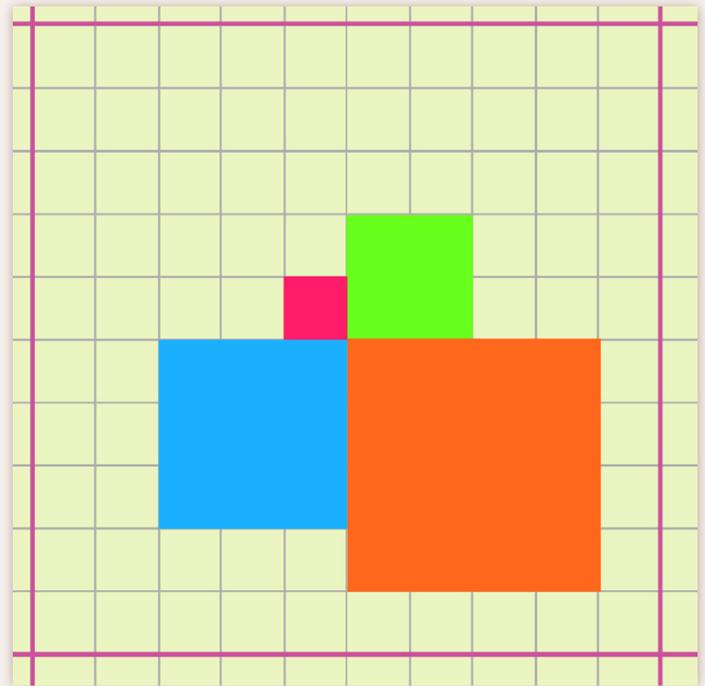


Figure 5. 1km grid (pink outline) over a 100m grid (gray). CDC SVI theme with the lowest vulnerability is represented by the smallest glyph (pink), and highest vulnerability by largest glyph (orange).

Unpopulated areas are removed from all maps to reduce clutter. Finally, Figure 10 juxtaposes the four univariate maps of Figure 2 with a multivariate glyph map of the same data.

DATA FILTERING

The CDC publishes SVI maps using a quartile classification, where data are divided into four categories according

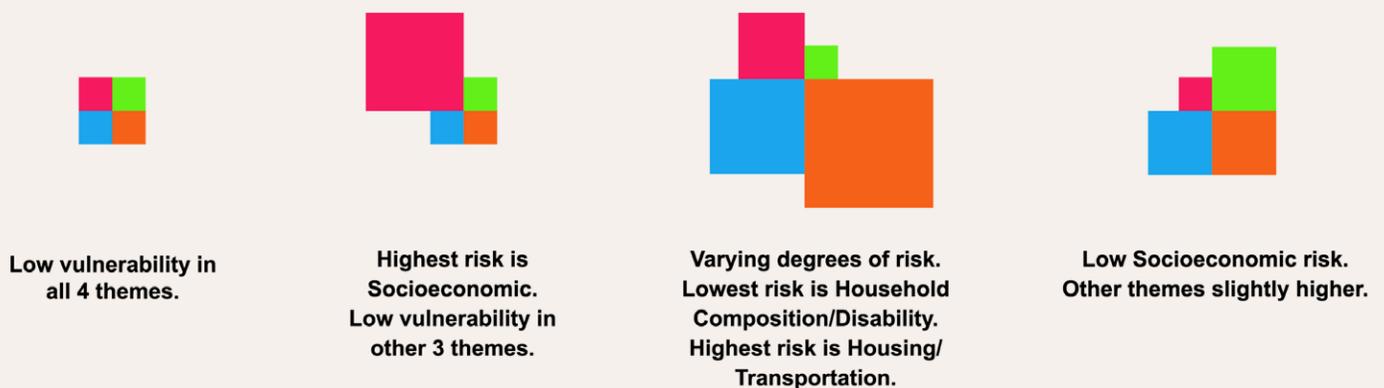


Figure 6. Glyphs of combinations of vulnerability themes (color) and their degrees of risk (classed square sizes).



Figure 7. Multivariate representation of the four themes using glyphs to show type (color) and degree (classed square size) of vulnerability.

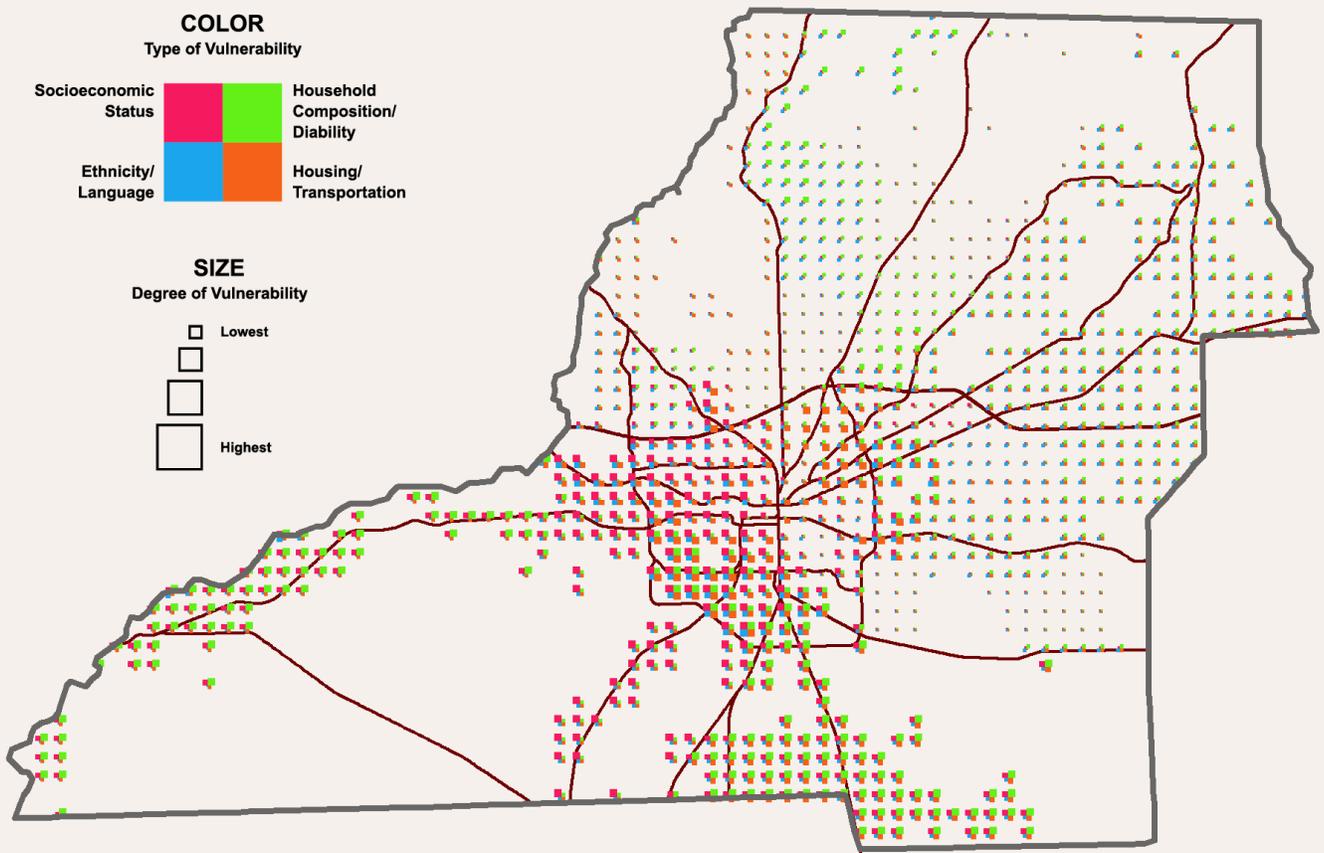


Figure 8. Multivariate map of social vulnerability for Leon County, Florida.

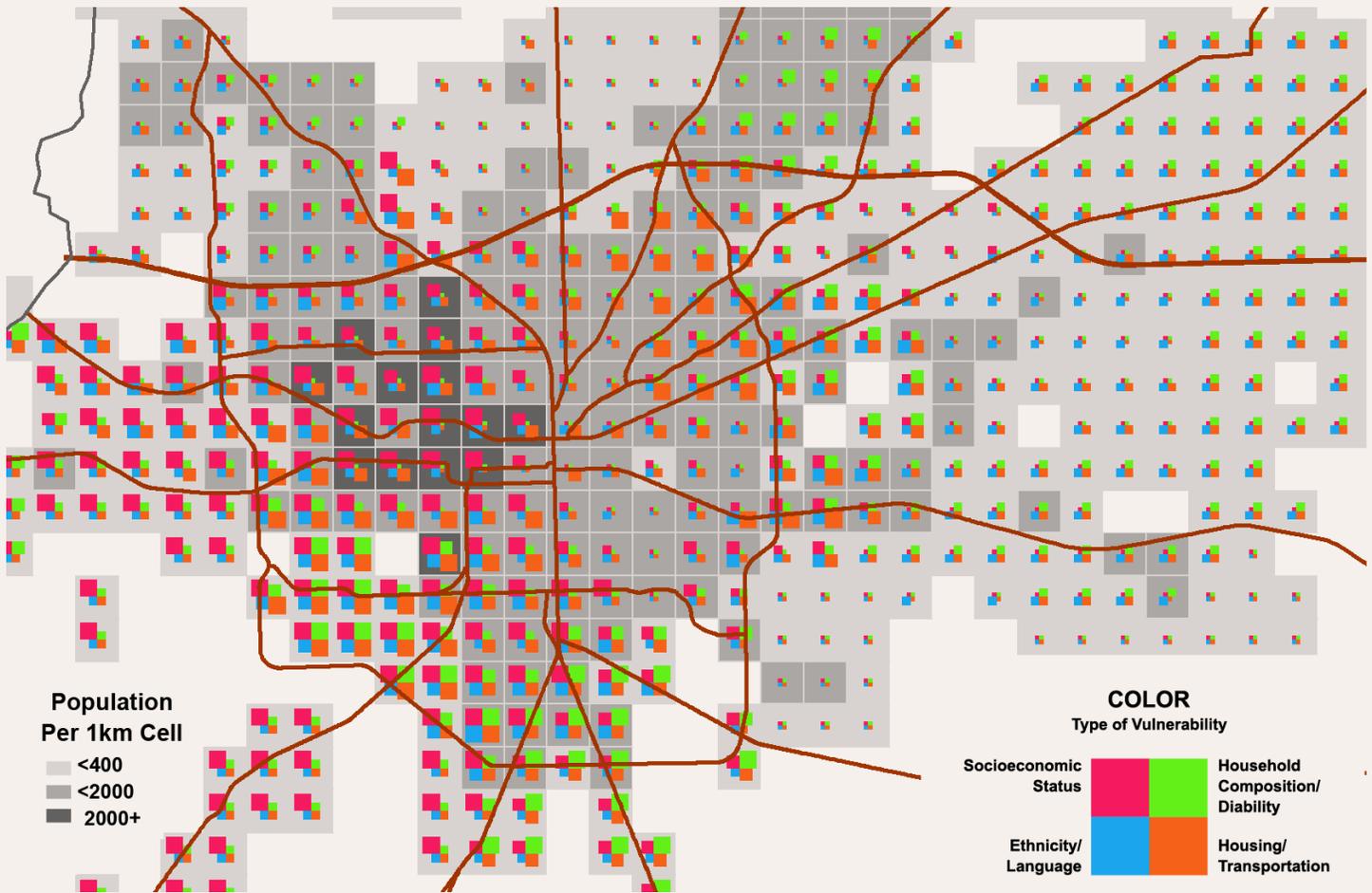


Figure 9. SVI data with population density of Tallahassee area.

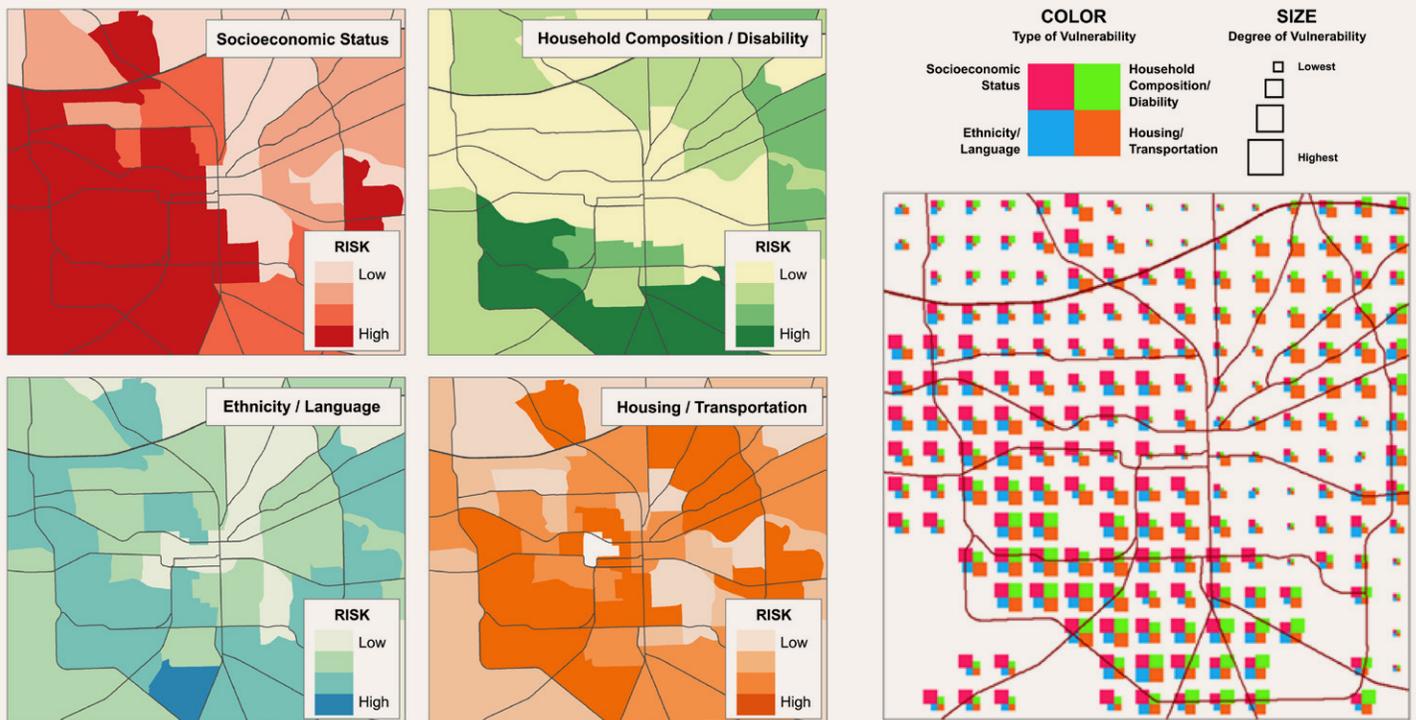


Figure 10. Left: Univariate maps of the four vulnerability themes. Right: Multivariate map of the same area using glyphs to depict the same data.

to the vulnerability score. While our maps above also use the quartile classification, we decided, additionally, to examine the value of filtering our data. By removing some data from the map, we could highlight areas with the highest vulnerabilities. Such a map would be useful to, for example, emergency planners, who may wish to view a map in which data below a threshold (e.g., 95th percentile) are removed, thereby focusing their attention on areas of interest. The maps in Figure 11 show 1km grid cells filtered by percentile: at least one of the cell's four vulnerability indices must meet the percentile threshold to remain on the map. The maps illustrate areas of high vulnerability, ranging from the 95th vulnerability percentile (top 5%) to the 70th percentile (top 30%).

VISUAL ANALYTICS USING EULER DIAGRAMS

As we have only four CDC SVI vulnerability themes, we can use Euler diagrams to represent them. Figure 12 demonstrates how, in the highest areas of vulnerability in Leon County, populations overlap across multiple themes.

We can see that the highest number of persons are vulnerable due to socioeconomic factors, and that this population partially overlaps with those vulnerable due to housing & transportation. At the 85th percentile, vulnerability due to household composition & disability becomes apparent, but has no overlap with the socioeconomic or housing & transportation themes until the 80th percentile. The 85th percentile diagram reveals a second group of vulnerable people separate from the previous groups. The fourth theme, ethnicity & language, is not evident until the 75th percentile and overlaps with the other three themes. Ethnicity & language does not occur in isolation from other themes. Figures 13 through 15 show the pairing of the maps and Euler diagrams to fully convey information on higher vulnerabilities.

INSIGHTS ENABLED USING MULTIVARIATE VISUALIZATION

These methods were developed and adapted using feedback from the Florida Agricultural and Mechanical University's Sustainability Institute and planners at the

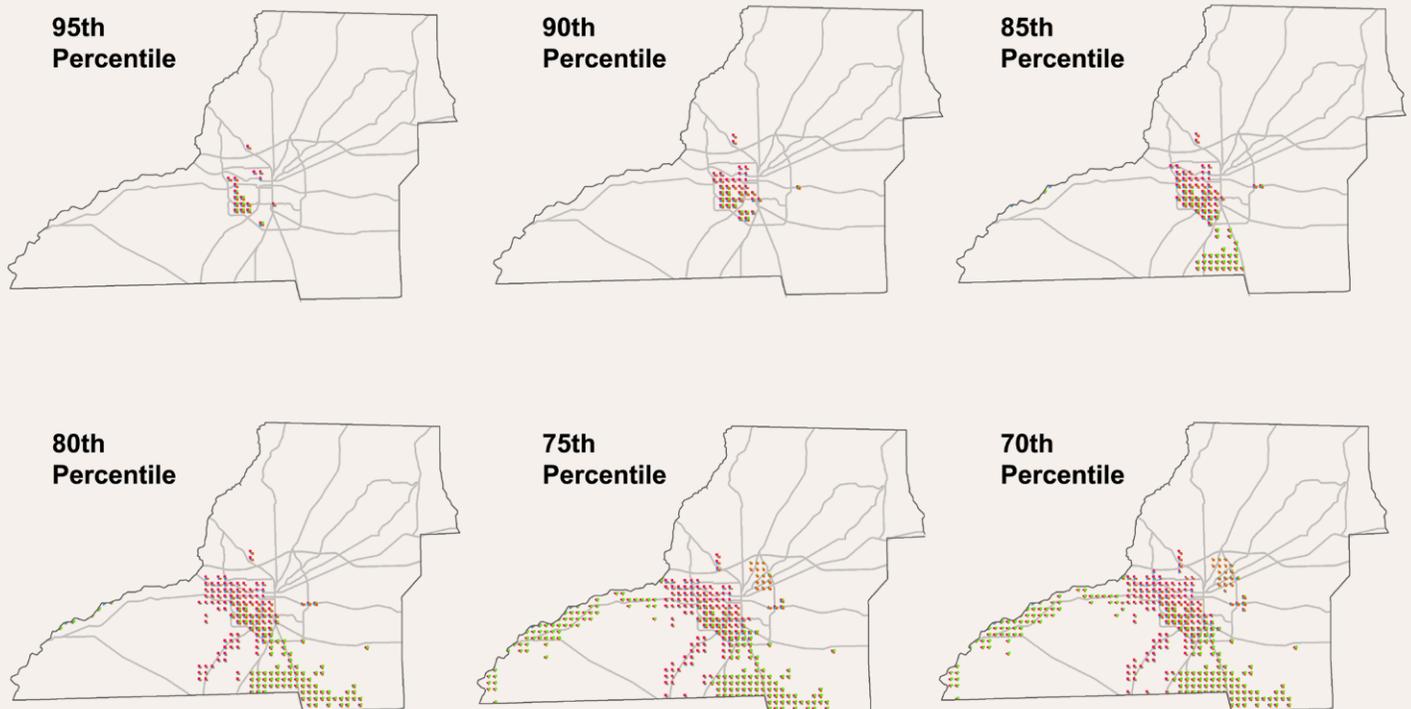


Figure 11. Leon County data filtered by percentile. At least one vulnerability index must be above the threshold indicated next to the map. Figures 13, 14, and 15 show larger images.

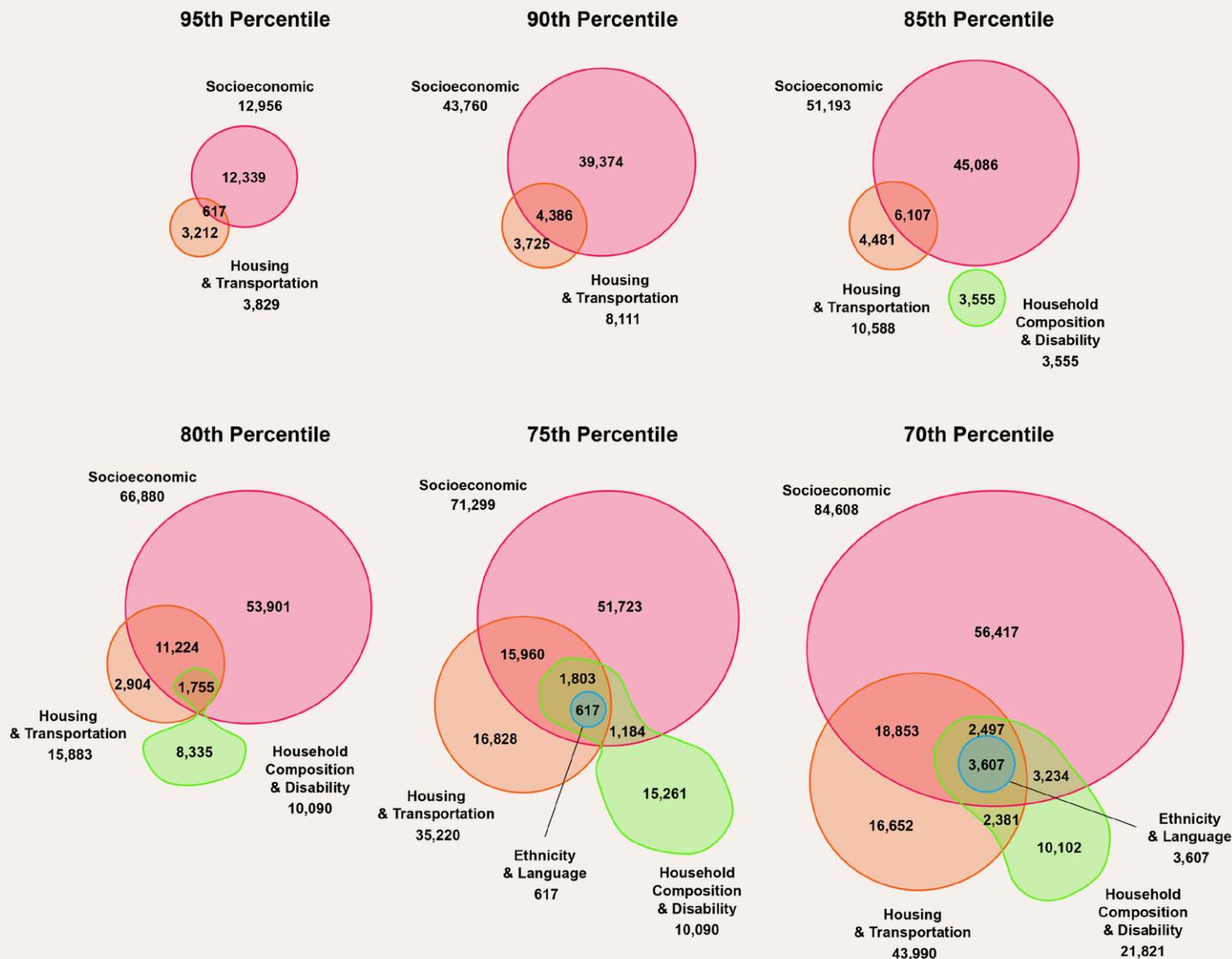


Figure 12. Euler diagrams of population counts within CDC social vulnerability themes at varying degrees of vulnerability for Leon County.

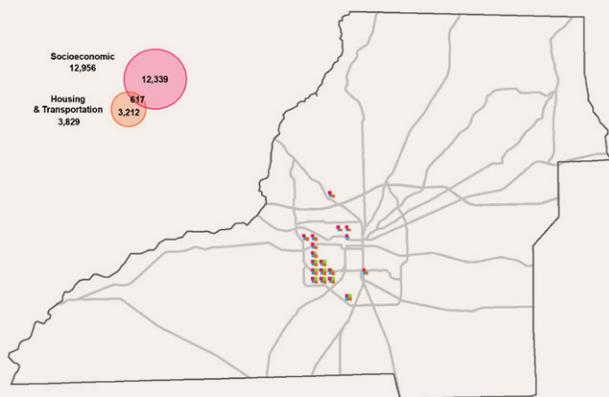


Figure 13. Euler diagram and filtered glyphs of areas with at least one social vulnerability theme ranking in the 95th percentile.

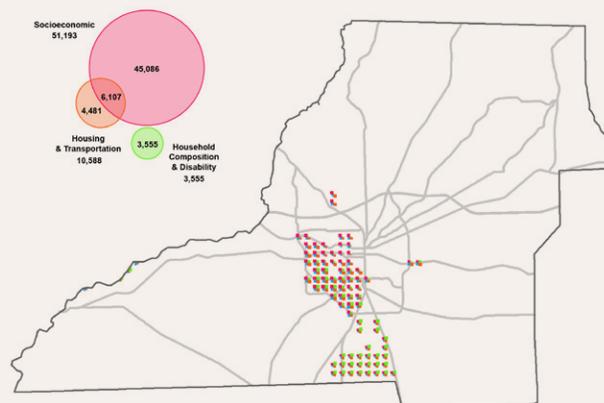


Figure 14. Euler diagram and filtered glyphs of areas with at least one social vulnerability theme ranking in the 85th percentile.

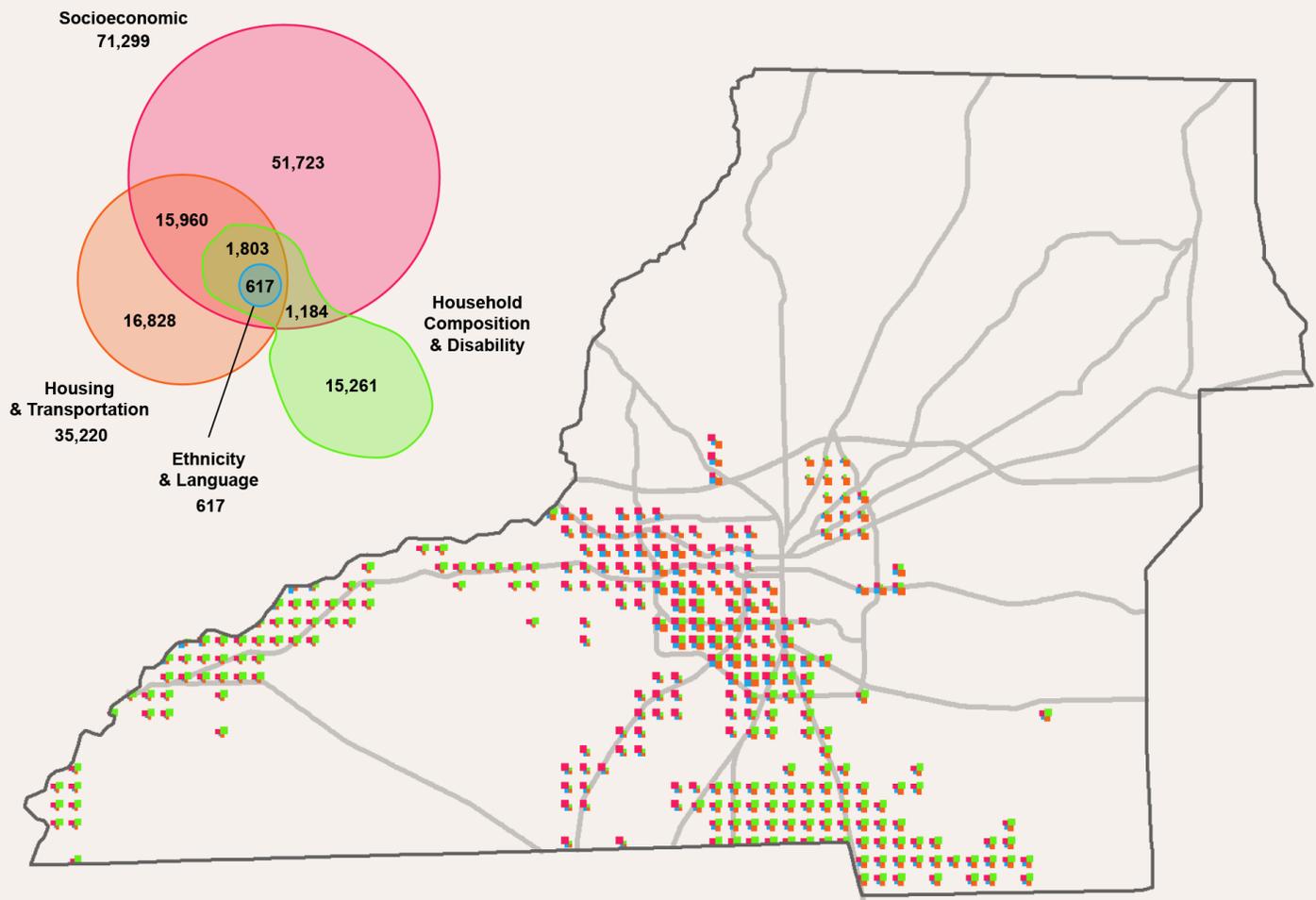


Figure 15. Euler diagram and filtered glyphs of areas with at least one social vulnerability theme ranking in the 75th percentile.

City of Tallahassee. After this project’s completion, we solicited additional feedback from experts in other fields as well as final feedback from the original planning group.

Visual Design

“This multivariate mapping technique for the visual display of social vulnerability index indicators is a brilliant display of design techniques to simplify a very complex issue. By combining color and size as indicators of type of vulnerability and degree of risk (respectively), it is possible to look at just one map and easily interpret the SVI for a particular area. The combination of these design elements makes the overall map easier to interpret than looking at the four univariate vulnerability indicator maps. The multivariate mapping technique is a wonderful example of how design techniques can simplify and aid data visualization.” (*Professor of Visual Communication*)

Planning

The SVI data can be used in all phases of disaster management, including mitigation and planning. The following statements are from planners at the City of Tallahassee:

“The [multivariate] map draws my attention to the west side of the city. In particular, the southwest part of the city appears to have multiple layers of vulnerability confirming what we know about the socio-economic demographics of that area. As practitioners, it alerts me about the complexity of building resilience for vulnerable people and clues me in to the customized approach necessary, from neighborhood to neighborhood.” (*City Planner 1*; see Figure 16)

“I can see where we need to locate our next resource center. This [multivariate] map makes the location stand out.” (*City Planner 2*)

Emergency Management

Puerto Rico has experienced several disasters in recent years: Hurricane Maria in 2017 and earthquakes in 2019, presenting challenges for the recovery process. The following comments relate to Figures 17, 18, and 19, which show both univariate and multivariate views of vulnerability indices for Puerto Rico.

“As an emergency responder, I use SVI data almost daily in a univariate form. If I had been supporting the recent earthquakes in Puerto Rico, I would be alerted to the poverty in the southwest. The high housing & transportation vulnerabilities indicate pre-existing resource limitations. High percentages of household composition & disability would cause me to question if there is a high elder population. Those factors would lead me to believe that those municipalities are going to have very few resources to recover quickly from this latest disaster. The univariate maps show single vulnerabilities, but assessment of multiple vulnerabilities requires much flipping between maps, taking time and mental capacity. In my opinion, a multivariate SVI map would be a valuable complement to the univariates.” (*Disaster Recovery Specialist working for a government agency*)

DISCUSSION

SOCIAL STATISTICS, INCLUDING the CDC’s SVI, are typically mapped using univariate graphic representations, without population counts, and with no supporting charts or graphs to aid in the understanding of social trends, including areas of social risk and vulnerability. They do not convey the *number* of people who are predisposed to higher risks from social and environmental hazards through pre-existing vulnerabilities. On the other hand, bivariate and multivariate geovisualizations facilitate the

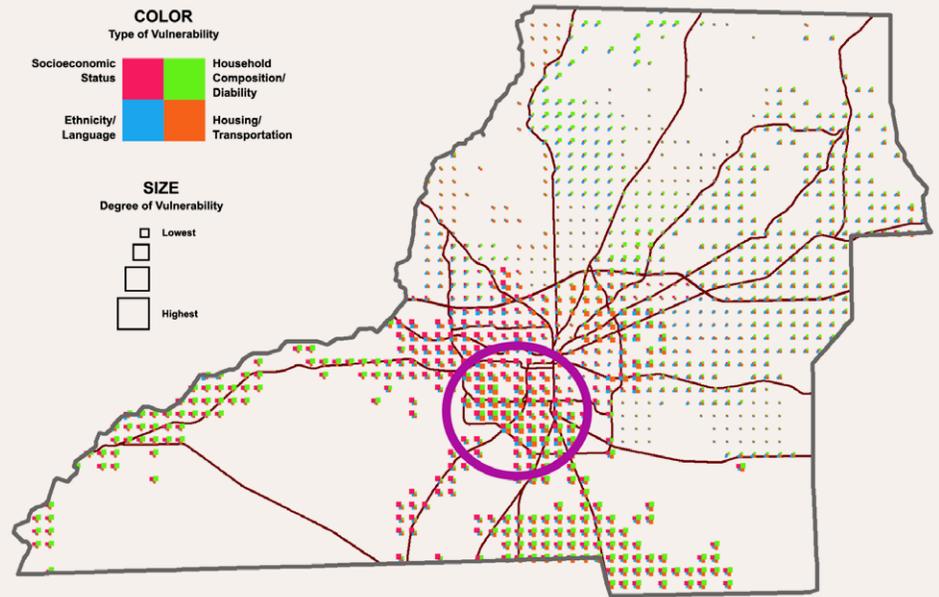


Figure 16. Vulnerability data for Leon County, Florida. The purple circle identifies the most vulnerable areas.

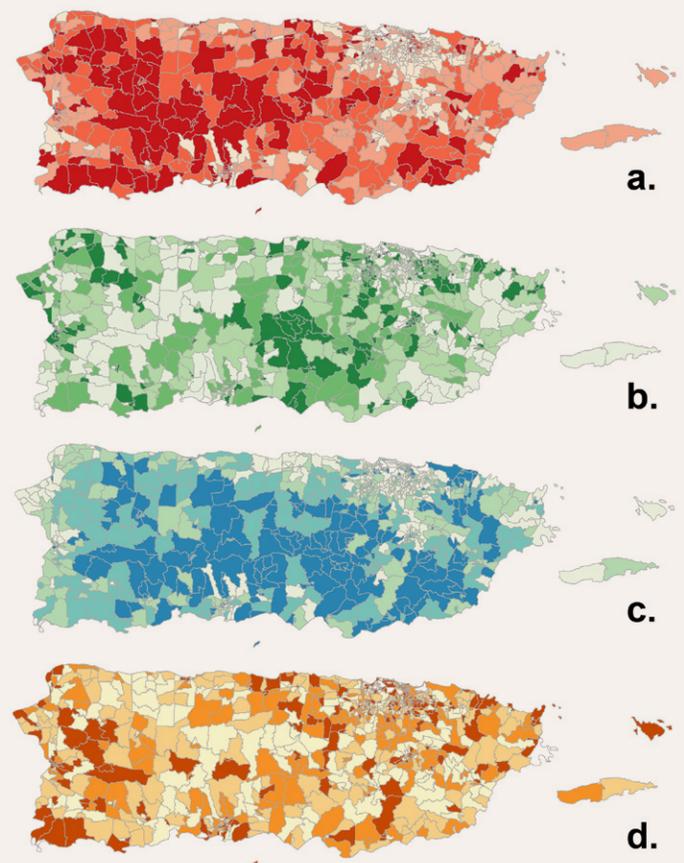


Figure 17. Univariate maps showing the four themes of 2016 social vulnerability for Puerto Rico. Darker colors represent higher vulnerability. a. Socioeconomic Status; b. Household Composition & Disability; c. Ethnicity & Language; d. Housing & Transportation.

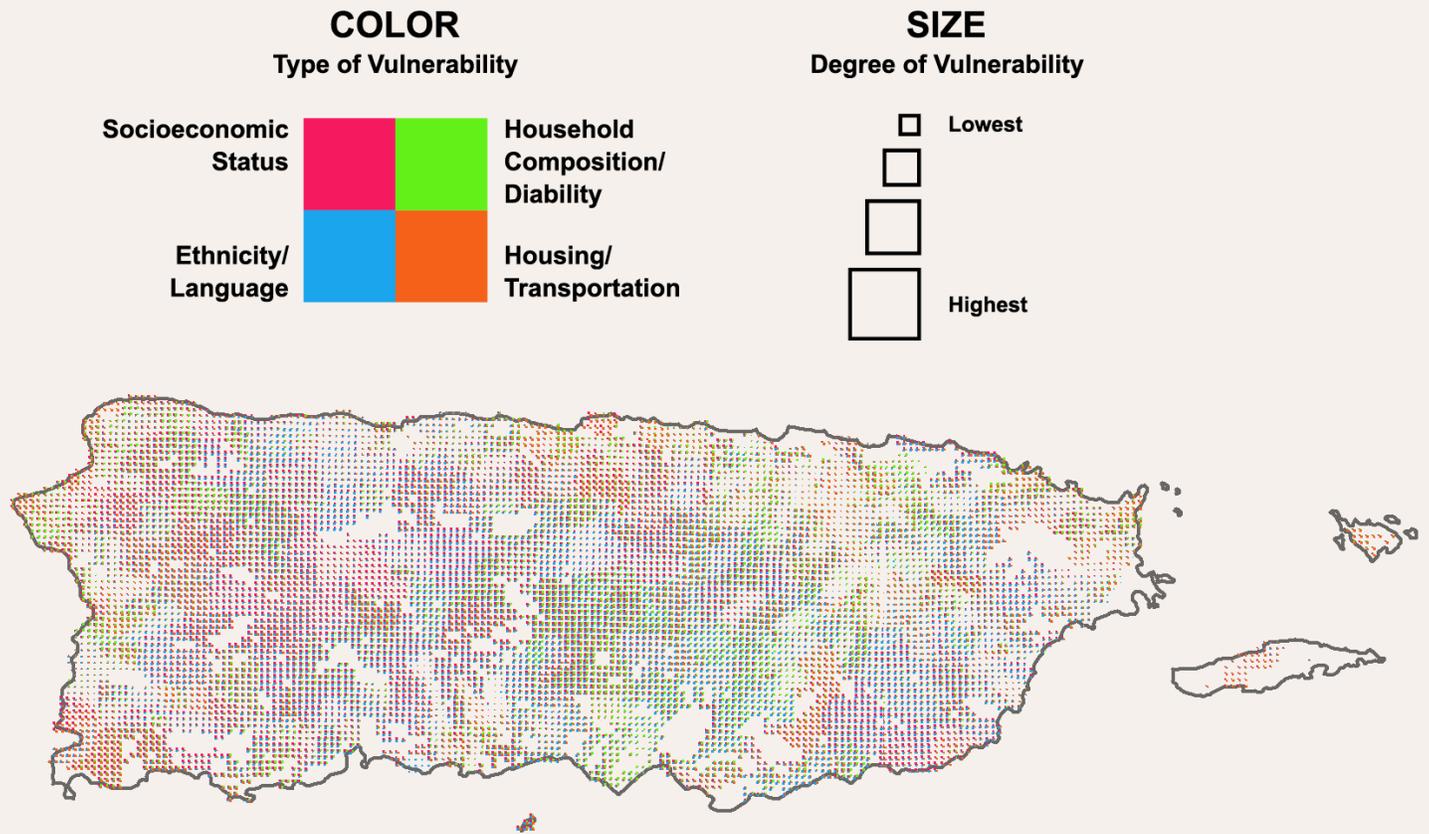


Figure 18. Multivariate 2016 social vulnerability map of Puerto Rico.

comparison of more than one social variable, and enable the correlation of overall vulnerability with population density. Using the designs we have demonstrated above, maps featuring as many as four variables (the CDC’s SVI vulnerability themes) can be produced, showing their quartile rankings. The glyphs illustrate the magnitude of the four vulnerability themes within a single point symbol, while Euler diagrams convey relationships across multiple themes and their associated population counts. Further, data filtering can be applied to focus on areas or themes of interest and to remove potentially distracting data (i.e., highlighting areas of higher vulnerability and removing areas of low vulnerability). These maps and visual analytics can be combined as needed to convey a fuller picture of an area’s social vulnerabilities in future planning projects.

For policy makers, planners, public health researchers, and public officials who use maps to develop and implement hazard response or resilience-building strategies, the bi- and multivariate techniques offer several advantages over univariate mapping. First, by distributing population to grid cells and removing those cells in unpopulated areas, the inaccurate impression that population characteristics are spread evenly over a census tract is replaced by a

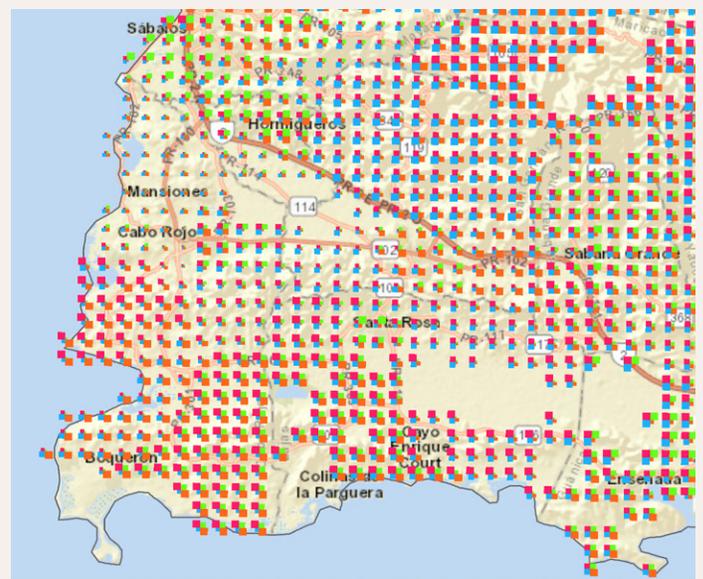


Figure 19. Multivariate map of 2016 Puerto Rico, southwest corner, at close zoom.

more nuanced and accurate representation. Combined with use of glyphs that simultaneously assess various vulnerability themes, the resulting maps convey a wealth of information in one representation. Euler diagrams provide at-a-glance population estimates, while the ability to

filter vulnerability to various percentiles allows for well-informed planning to meet the needs of the most vulnerable people. For example, officials seeking to create “resilience hubs” in new or existing civic structures could readily see where resources are most needed. Local officials, even those who believe they have deep familiarity with the locality’s population, may see patterns and pockets of vulnerability they were unaware of. By communicating information precisely and succinctly these maps will be critical for emergency response teams, which can be composed of personnel unfamiliar with the areas they are assisting. The methods described and illustrated here together form a promising platform for use in better understanding and responding to the geographically uneven characteristics of social vulnerability.

LIMITATIONS

While our method can be considered an innovative demonstration of multivariate mapping, it is not without limitations. Data collected at the census tract level can mask patterns within the tract; it is possible that isolated or rural populations may be overlooked and these may sometimes contain crucial minority enclaves. When census tracts are resampled to a 1km grid system, the finer scale can falsely imply that the data are collected or analyzed at this higher-resolution scale. The quartile classification

CONCLUSIONS

RISK MANAGERS REQUIRE an integrated understanding of vulnerability data, which are multidimensional. Yet visualization of this particular type of data has not yet been thoroughly discussed in the cartographic literature. In this paper, we have demonstrated a useful method of presenting social vulnerability data via quadrant glyphs on a map, to better represent multiple variables along with population size and geographical location. The bivariate and multivariate mapping strategies, as well as the visual analytics shown here, offer an improvement over traditional univariate mapping. A member of an emergency response team told us that she used SVI data daily when her team was deployed to unfamiliar areas. She stated that “This method would clearly be easier than looking through four separate maps.”

With this method, the viewer can target a specific area and immediately see the extent to which vulnerability

system classifies data at a class border without providing any indication that data may be very close to another class. Finally, the data size of the 100-meter vector grids can be challenging, but this can be resolved by removing any unshaded grids or converting to a raster layer.

FUTURE WORK

This method can be easily adapted in a number of different ways. The data can be included in an online interactive environment that could benefit those needing information quickly, such as emergency management and response teams. The Euler diagrams and maps could be hyperlinked, so that clicking on a segment of the Euler diagram could highlight the relevant glyphs, and vice versa. The data itself could be modified for specific purposes, with CDC SVI data being replaced by other relevant indices. For example, groups focused on elderly populations could remove SVI data not pertinent to the elderly and replace them with more relevant measures from external sources, such as whether food or medical resources are available within the community. Users could be given the ability to filter data on the fly, reducing map clutter and showing only targeted areas of interest. Rather than showing the glyphs, cells could be filled with bar charts or other creative designs to convey data meaning.

types vary. The resulting maps are more complex but also less cluttered, as unpopulated areas are removed, and, in the case of the bivariate maps, populated areas are scaled to distinguish between sparsely and densely populated areas. The multivariate glyphs convey the four vulnerability themes and their quartile categories simultaneously, and eliminate the need for four separate univariate maps. Euler diagrams complement the spatial data by conveying the quantity of population affected by multiple types of vulnerability and show the overlap between multiple themes as an area can be affected by more than one type of vulnerability. Furthermore, a data filtering method can be combined with these techniques to isolate and highlight the areas with the highest vulnerabilities.

In addition to being helpful for planners and responders, these maps and diagrams can be a useful means for engaging the community. They can be shared at neighborhood

meetings and residents asked for their input on the strengths of the neighborhood as well as its vulnerabilities. Discussions could be held on causes of vulnerability and means to address or mitigate them. The maps and other visuals can aid advocacy and provide political leverage for neighborhoods with fewer resources. From a planning perspective, these maps and diagrams can be useful in identifying areas in need of resources and infrastructure improvements. They can provide data for decision makers to allocate resources more objectively based on need.

In this article we have presented and described a method for mapping bivariate and multivariate data using the CDC's social vulnerability data. The next step is the evaluation of this method for policy, planning, emergency response, and research beyond ease of viewing and better understanding interrelationships of the variables. We invite readers to apply these techniques to their areas of interest and report on the outcomes.

ACKNOWLEDGMENTS

The authors would like to thank Laura Blackstone, William Tolone, Jim Klassen, Abena Ojetayo, Behnam Shadravan, Bruce Strouble, Brittany Gress, Adam Jacobs, Juliann Cortese, Amy Griffin, Daniel Huffman, and three anonymous reviewers for their thoughtful comments and suggestions.

REFERENCES

- Alsallakh, Bilal, Wolfgang Aigner, Silvia Miksch, and Helwig Hauser. 2013. "Radial Sets: Interactive Visual Analysis of Large Overlapping Sets." *IEEE Transactions on Visualization and Computer Graphics*, 19 (12): 2496–2505. <https://doi.org/10.1109/tvcg.2013.184>.
- Andrienko, Gennady, Natalia Andrienko, Harald Bosch, Thomas Ertl, Georg Fuchs, Piotr Jankowski, and Dennis Thom. 2013. "Thematic Patterns in Georeferenced Tweets through Space-time Visual Analytics." *Computing in Science & Engineering* 15 (3): 72–82. <https://doi.org/10.1109/mcse.2013.70>.
- Bleisch, Susanne, and Daria Hollenstein. 2018. "Exploring Multivariate Representation of Indices along Linear Geographic Features." *Proceedings of the ICA 1*. <https://doi.org/10.5194/ica-proc-1-11-2018>.
- Borgo, Rita, Johannes Kehrer, David H. S. Chung, Eamonn Maguire, Robert S. Laramee, Helwig Hauser, Matthew Ward, and Min Chen. 2013. "Glyph-Based Visualization: Foundations, Design Guidelines, Techniques, and Applications." In *EUROGRAPHICS 2013 - State of the Art Reports*, edited by M. Sbert and L. Szirmay-Kalos, 39–63. Aire-la-Ville, Switzerland: The Eurographics Association <http://doi.org/10.2312/conf/EG2013/stars/039-063>.
- Brewer, Cynthia A. 1994. "Color Use Guidelines for Mapping and Visualization." In *Modern Cartography. Vol. 2: Visualization in Modern Cartography*, edited by Alan M. MacEachren and D. R. Fraser Taylor, 123–147. New York: Elsevier Science, Inc.
- Brodie, Mollyann, Erin Weltzien, Drew Altman, Robert J. Blendon, and John M. Benson. 2006. "Experiences of Hurricane Katrina Evacuees in Houston Shelters: Implications for Future Planning." *American Journal of Public Health* 96 (8): 1402–1408. <https://doi.org/10.2105/ajph.2005.084475>.
- Brooks, Harold E., and Charles A. Doswell III. 2001. "Normalized Damage from Major Tornadoes in the United States: 1890–1999." *Weather Forecasting* 16: 168–176. [https://doi.org/10.1175/1520-0434\(2001\)016%3C0168:ndfmti%3E2.0.co;2](https://doi.org/10.1175/1520-0434(2001)016%3C0168:ndfmti%3E2.0.co;2).
- Centers for Disease Control and Prevention (CDC). 2015. *Planning for an Emergency: Strategies for Identifying and Engaging At-Risk Groups. A Guidance Document for Emergency Managers*. Atlanta, GA: CDC.
- . 2018. *CDC's Social Vulnerability Index Fact Sheet*. Accessed May 1, 2019. <https://svi.cdc.gov/factsheet.html>.

- Chernoff, Hermann. 1973. "The Use of Faces to Represent Points in k-Dimensional Space Graphically." *Journal of the American Statistical Association* 68 (342): 361–368. <https://doi.org/10.1080/01621459.1973.10482434>.
- Collins, Christopher, Gerald Penn, and Sheelagh Carpendale. 2009. "Bubble Sets: Revealing Set Relations with Isocontours over Existing Visualization." *IEEE Transactions on Visualization and Computer Graphics* 15 (6): 1009–1016. <https://doi.org/10.1109/TVCG.2009.122>.
- Cutter, Susan L. 1996. "Vulnerability to Environmental Hazards." *Progress in Human Geography* 20 (4): 529–539. <https://doi.org/10.1177/030913259602000407>.
- Cutter, Susan L., Bryan J. Boruff, and W. Lynn Shirley. 2003. "Social Vulnerability to Environmental Hazards." *Social Science Quarterly* 84 (2): 242–261. <https://doi.org/10.1111/1540-6237.8402002>.
- Federal Emergency Management Agency (FEMA). n.d. *Disasters*. Accessed March 13, 2020. <http://www.fema.gov/disasters/grid/year>.
- Federal Geographic Data Committee. 2001. *United States National Grid*. https://www.fgdc.gov/standards/projects/FGDC-standards-projects/usng/fgdc_std_011_2001_usng.pdf.
- Flanagan, Barry E., Edward W. Gregory, Elaine J. Hallisey, Janet L. Heitgerd, and Brian Lewis. 2011. "A Social Vulnerability Index for Disaster Management." *Journal of Homeland Security and Emergency Management* 8 (1): Article 3. <https://doi.org/10.2202/1547-7355.1792>.
- Flanagan, Barry E., Elaine J. Hallisey, Erica Adams, and Amy Lavery. 2018. "Measuring Community Vulnerability to Natural and Anthropogenic Hazards: The Centers for Disease Control and Prevention's Social Vulnerability Index." *Journal of Environmental Health* 80 (10): 34–36.
- Fuchs, Johannes, Petra Isenberg, Anastasia Bezerianos, and Daniel Keim. 2017. "A Systematic Review of Experimental Studies on Data Glyphs." *IEEE Transactions on Visualization and Computer Graphics*, 23 (7): 1863–1879. <https://doi.org/10.1109/tvcg.2016.2549018>.
- Haines, Andy, Sari R. Kovats, Diarmid H. Campbell-Lendrum, and Carlos F. Corvalán. 2006. "Climate Change and Human Health: Impacts, Vulnerability and Public Health." *Public Health* 120 (7): 585–596. <https://doi.org/10.1016/j.puhe.2006.01.002>.
- Jia, Peng, Youliang Qiu, and Andrea E. Gaughan. 2014. "A Fine-scale Spatial Population Distribution on the High-resolution Gridded Population Surface and Application in Alachua County, Florida." *Applied Geography* 50: 99–107. <https://doi.org/10.1016/j.apgeog.2014.02.009>.
- Langford, Mitchel. 2013. "An Evaluation of Small Area Population Estimation Techniques using Open Access Ancillary Data." *Geographical Analysis* 45: 324–344. <https://doi.org/10.1111/gean.12012>.
- Lemon, Oliver Joseph, and Ian Pratt. 1997. "Spatial Logic and the Complexity of Diagrammatic Reasoning." *Machine Graphics and Vision* 6: 11–21.
- Martin, David, Chris Lloyd, and Ian Shuttleworth. 2011. "Evaluation of Gridded Population Models Using 2011 Northern Ireland Census Data." *Environment and Planning A* 43 (8): 1965–1980. <https://doi.org/10.1068/a43485>.
- McNabb, Liam, and Robert S. Laramee. 2019. "Multivariate Maps—A Glyph-Placement Algorithm to Support Multivariate Geospatial Visualization." *Information* 10 (10): 302. <https://doi.org/10.3390/info10100302>.
- Mennis, Jeremy. 2003. "Generating Surface Models of Population using Dasymetric Mapping." *The Professional Geographer* 55 (1): 31–42. <http://doi.org/10.1111/033-0124.10042>.
- Morrow, Betty H. 1997. "Stretching the Bonds: The Families of Andrew." In *Hurricane Andrew: Ethnicity, Gender, and the Sociology of Disasters*, edited by Walter G. Peacock, Betty H. Morrow, and Hugh Gladwin, 141–170. London: Routledge.
- . 2002. "Identifying and Mapping Community Vulnerability." *Disasters* 23 (1): 1–18. <https://doi.org/10.1111/1467-7717.00102>.

- Peacock, Walter Gillis, Howard Kunreuther, William H. Hooke, Susan L. Cutter, Stephanie E. Chang, and Phillip R. Berke. 2008. "Toward a Resiliency and Vulnerability Observatory Network: RAVON." *HRRC reports: 08-02R*. Texas A&M University. <https://www.nehrp.gov/pdf/RAVON.pdf>.
- Shneiderman, Ben. 1996. "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization." In *Proceedings 1996 IEEE Symposium on Visual Languages*, 336–343. Los Alamitos, CA: IEEE Computer Society Press. <https://doi.org/10.1109/vl.1996.545307>.
- Strode, Georgianna, Victor Mesev, Julianna Maantay. 2018. "Improving Dasymetric Population Estimates for Land Parcels: Data Pre-Processing Steps." *Southeastern Geographer* 58 (3): 300–316. <https://doi.org/10.1353/sgo.2018.0030>.
- Strode, Georgianna, Victor Mesev, Benjamin Thornton, Marjorie Jerez, Thomas Tricarico, and Tyler McAlear. 2019. "Geovisualization of Florida Land Use and Land Cover using Bivariate Mapping, Statistical Legends, and Visual Analytics." *The Florida Geographer* 51 (1). <https://journals.flvc.org/flgeog/article/view/118173>.
- Tarling, Hannah Andrea. 2017. "Comparative Analysis of Social Vulnerability Indices: CDC's SVI and SoVi®." Master's Thesis, Lund University.
- Tierney, Kathleen. 2006. "Social Inequality: Humans and Disasters." In *On Risk and Disaster: Lessons from Hurricane Katrina*, edited by Ronald J. Daniels, Donald F. Keitl, and Howard Kunreuther, 109–128. Philadelphia: University of Pennsylvania Press. <https://doi.org/10.9783/9780812205473.109>.
- US Fire Administration. 2015. *National Fire Incident Reporting System. FEMA*. https://www.usfa.fema.gov/downloads/pdf/nfirs/NFIRS_Complete_Reference_Guide_2015.pdf.
- US National Grid Information Center. n.d. *Data*. Accessed March 13, 2020. <https://usngcenter.org/portfolio-item/usng-gis-data>.
- Wang, Chongming, and Brent Yarnal. 2012. "The Vulnerability of the Elderly to Hurricane Hazards in Sarasota, Florida." *Natural Hazards* 63 (2): 349–373. <https://doi.org/10.1007/s11069-012-0151-3>.
- Ward, Matthew O. 2002. "A Taxonomy of Glyph Placement Strategies for Multidimensional Data Visualization." *Information Visualization* 1 (3-4): 194–210. <https://doi.org/10.1057/palgrave.ivs.9500025>.
- . 2008. "Multivariate Data Glyphs: Principles and Practice." In *Handbook of Data Visualization*, edited by Chun-houh Chen, Wolfgang Härdle, and Antony Unwin, 179–198. Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-540-33037-0_8.
- Willmott, Cort J., Scott M. Robeson, and Kenji Matsuura. 2013. "Geographic Box Plots." *Physical Geography* 28 (4): 331–344. <http://doi.org/10.2747/0272-3646.28.4.331>.
- Wolkin, Amy, Jennifer Rees Patterson, Shelly Harris, Elena Soler, Sherry Burren, Michael McGeehin, and Sandra Greene. 2015. "Reducing Public Health Risk During Disasters: Identifying Social Vulnerabilities." *Journal of Homeland Security and Emergency Management* 12 (4): 809–822. <https://doi.org/10.1515/jhsem-2014-0104>.
- Wright, James D., Peter H. Rossi, Sonia R. Wright, and Eleanor Weber-Burdin. 1979. *After the Clean-up: Long Range Effects of Natural Disasters*. Beverly Hills, CA: Sage.
- Zhou, Mengjie, Jing Tian, Fuquan Xiong, and Rui Wang. 2017. "Point Grid Map: A New Type of Thematic Map for Statistical Data Associated with Geographic Points." *Cartography and Geographic Information Science* 44 (5): 374–389. <https://doi.org/10.1080/15230406.2016.1160797>.