Visualizing Health Indicators

Tamara Barker, James Mallon, Tushar Sethi, and David Walk

Abstract— This paper presents the use of ArcGIS, Tableau, and CiteSpace to create detailed and interactive visualizations to explore the relationships between chronic diseases and behaviors associated with their prevalence. The tools aim to help the user better understand and explore the CHSI dataset and thus ultimately fulfill the original goal of the CHSI report and promote communication on the best health practices. Results were consistent with other independently reported findings and show the benefits of effective visualization techniques.We also included discussions of academic literature on the subject of chronic health issues and an effort to identify a correlation between one chronic health issue – obesity – and criminal activity.

------------------- ✦ -------------------

1 Introduction

As defined by the US Department of Health and Human Services, "The goal of Community Health Status Indicators (CHSI) is to provide an overview of key health indicators for local communities and to encourage dialogue about actions that can be taken to improve a community's health"[1]. The CHSI report contains over 200 measures for each of the 3,141 United States counties. The report contains data not only related to prevalence of chronic diseases but also information related to behavioral factors that contribute to chronic disease-associated deaths, such as tobacco use, diet, physical activity, alcohol and drug use, and sexual behavior.

The CHSI report aims to promote communication amongst healthcare professionals and concerned community members. By making this kind of information publicly available many individuals now have access to this rich data source. However without an efficient means of effectively visualizing the available data the goal of improving communication will be hard to realize. It has been well documented that many people have trouble understanding statistical data [2]. Thus providing rich and interactive visualizations will help the audience better understand and grasp the data. The human brain can perceive complex patterns in data more easily when those data are presented in a graphic (in particular, a map) format, as opposed to tabular displays of numeric values [3].

In this paper we examine two unique and cost-effective ways to visualize the CHSI dataset, thereby helping the user to better comprehend the available data and thus facilitate communication. We also analyze academic literature on the subject of chronic health issues and discuss an effort to identify a correlation between obesity and criminal activity.

2 Dataset

Since it would be difficult to analyze and visualize such a large set of indicators as that of the CHSI dataset in the time given for this project, we decided to focus on a small subset of indicators to find insights and patterns related to healthy lifestyles across the United States. The full CHSI dataset is located at, http://www.data.gov/communities/node/81/data_tools/325# and the data definitions are located at http://www.communityhealth.hhs.gov/COMPANION_DOCUMENT/CHSI-Data_Sources_Definitions_And_Notes.pdf.

We focused our attention on the Risk Factor and Access To Care dataset for our analysis. More detailed information concerning the dataset can be found in the data definitions for the following indicators:

- Column G: No exercise. Information can be found on page 31 in the community health pdf.
- Column H: Few Fruits and Vegetables. Information can be found on 31 in the community health pdf.
- Column I: Obesity. Information can be found on 31 in the community health pdf.
- Column J: High Blood Pressure. Information can be found on 31 in the community health pdf.
- Column K: Smoker. Information can be found on 31 in the community health pdf.
- Column L: Diabetes. Information can be found on 31 in the community health pdf.
- Column M: Uninsured. Information can be found on 31 in the community health pdf.
- Column N: Primary Care Physician Rate. Information can be found on 32 in the community health pdf.
- Column O: Population. This was taken from the Demographics file. Information can be found on 5 in the community health pdf.

The dataset consists of 40,846 cells Of that total, 6004 cells had a value of -1111.1 and 3 of the cells had a value of -2222. The defined data value -1111.1 is referred as an 'nfr' or 'no report' and the defined data value -2222 is referred to as an 'nda' or 'no data available'. After filtering the dataset for these values it was determined that Alaska and DC would not be included in the final analysis since Alaska reported on three indicators and DC did not report on uninsured individuals.

3 Tableau Software

Tableau 6.1 Professional [4] was used to measure the values from the dataset and seek insight and patterns. Tableau is classified as business/data analytics software that allows business and IT to create visualizations from multiple data sources. Tableau "is based on breakthrough technology from Stanford University" that allows a user to import datasets from Text files or Microsoft Excel or Access, then use drag and drop functionality to manipulate a series of data points to create visualizations.

3.1 Methodology

The dataset referenced above was imported into Tableau using the Connect To Data wizard. A New Connection was created then the Import All Data function was enabled. The data points were arranged as either a Dimension or a Measure. Using drag and drop functionality the State Abbreviation dimension was moved into the Columns legend and the indicators moved to the Rows legend. The data was aggregated at the state level to determine the overall health of a state and rank each accordingly. Each measure was adjusted for average since totals would skew towards the most populated and least populated states. To adequately determine a normalized uninsured metric, uninsured was defined as a percentage of each state's population. Once the indicators were visualized multiple features were used to determine the final look of the visualization. To determine the most healthy and least healthy states, the Smoker, Few Fruits and Vegetables, No Exercise, Diabetes, High Blood Pressure, and Obesity indicators were sorted to determine their values. The indicators were defined as follows:

- Smoker: Responded 'Yes' if they smoke.
- Few Fruits and Vegetables: Less than 5 servings a day.
- No Exercise: No physical activities in the last month.
- Diabetes: Responded 'Yes' that a physician informed them of their condition.
- High Blood Pressure: Responded 'Yes' that a physician informed them of their condition.
- Obesity: BMI >30.

3.2 Analysis

The top 3 healthy and non-healthy states in each category were recorded and 5 points were rewarded for the top state, 4 points for the second, and 3 points for the third.

Top 3 Best

	5 Points	4 Points	3 Points
Smoker	**UT**	**CA**	**MA**
Few Fruits and Vegetables	**DC*VT**	**MA**	**RI**
No Exercise	**CO**	**MN**	**HI**
Diabetes	**CO**	**UT**	**MA**
High Blood Pressure	**CO**	**NM**	**UT**
Obesity	**CO**	**MA**	**RI**

Bottom 3 Worst

	5 Points	4 Points	3 Points
Smoker	**KY**	**WV**	**TN**
Few Fruits and Vegetables	**OK**	**LA**	**MS**
No Exercise	**LA**	**KY**	**MS**
Diabetes	**WV**	**SC**	**MS**
High Blood Pressure	**WV**	**MS**	**AL**
Obesity	**MS**	**SC**	**LA**

*DC was not included in the final analysis because they did not report their uninsured individuals.

When the points were totaled the following states were recognized as being the healthiest:

1. CO (20 points)
2. MA (14 points)
3. UT (12 points)

When the points were totaled the following states were recognized as being the least healthy:

1. MS (18 points)
2. WV (14 points)
3. LA   (12 points)

After determining the status of the states, additional indicators were further analyzed to determine any correlations.

3.2.1 Smoking

This metric was used to determine the impact of smoking on all of the other indicators and to understand the role it plays in determining the healthiest states. Before any analysis was done it could be assumed that states with lower smoking rates would rank higher in terms of overall health. The states with the lowest smoking rates were UT (13.47), CA (15.42), and MA (19.22) and those with the highest rates were KY (31.41), WV (27.20), and TN (27.02). UT and MA finished in the top three in the healthiest category and WV finished in the bottom three. It's interesting to note that CA has such low smoking rate for such a large population but they don't rank as high in other indicators. However, if the ranking was expanded, CA would most likely come in the top ten of the healthiest states.

3.2.2 Uninsured Percentage of Population

This metric was created to better quantify the impact on overall health rates as opposed to looking at the total amount of uninsured individuals which would tend to penalize states with larger populations. Before any analysis was done it could be assumed that states with lower uninsured percentages would influence the health status of a state more positively. The states with the largest populations, CA, TX, and NY had percentages of 17.68%, 17.21%, and 13.80% respectively for an average of 16.23%. The states with the lowest populations WY, DC*/VT, and ND had the following percentages: 8.81%, 9.45%, and 14.78% for an average of 11.01%. The state with the lowest percentage was NH at 7.69% and the state with the highest percentage was NM at 22.13%. The average of all states 12.63%. Establishing percentage thresholds on the largest and smallest populated states provides a baseline to compare the uninsured percentages of the most healthy and least healthy states. The healthiest states comprise of CA, MA, and UT and their percentages were, 14.08%, 9.25%, and 13.06% for an average of 12.13% which is slightly under the overall average of 12.63%. MS, WV, and LA are the three least health states and their percentages were 16.36%, 12.82%, and 16.44% for an average of 15.21% which is above the overall average. Although a lower percentage does not guarantee a healthier state, it does appear that the healthiest states have on average a lower percentage of uninsured individuals versus the least healthy states.

3.2.3 Primary Care Physicians

The amount of primary care physicians (PCP) per 100,000 people is anther metric worth exploring. Before any analysis was done it could be assumed that higher rates of PCP's per 100,000 individuals could lead to healthier lifestyles and a healthier state. The state with the most PCP's was MA with 119.3 per 100,000 and the state with the least amount of PCP's was MO with 36.83 per 100,000. The overall state average is 65.33 per 100,000. The healthiest states had the following PCP's per 100,000 individuals: CO, 68.4, MA, 119.3, and UT 49.1. The average for these three states is 78.93 which is above the overall average and individually two of the three top states are above the national average. The least healthy states, LA, WV, and MS had the following PCP's per 100,000, 53, 54.8, and 45.3 respectively for an average of 51.03 which is below the national average of 65.33 with no state exceeding the national average.

3.2.4 Multiple Health Indicators

Viewing multiple indicators has the ability to provide another perspective and possibly show a bigger picture. This analysis involved creating a visualization (Figure 1) that shows the total number of obese individuals per state for each of the following indicators, diabetes, few fruits and vegetables, high blood pressure, and no exercise and while trending each with smoking. Obesity levels go from 15 to 30 and states with the lowest levels of obesity are closer to 15 and 20 and those with higher levels are located further on the right hand side of the screen. Dots with a lighter shade of green indicate lower smoking levels and dots that are darker indicate higher smoking levels. Within each indicator (Diabetes, Few Fruits and Vegetables, High Blood Pressure, and No Exercise on the y axis) dots positioned lower in the grid have lower levels of the indicator and those towards the top of the grid have higher levels.
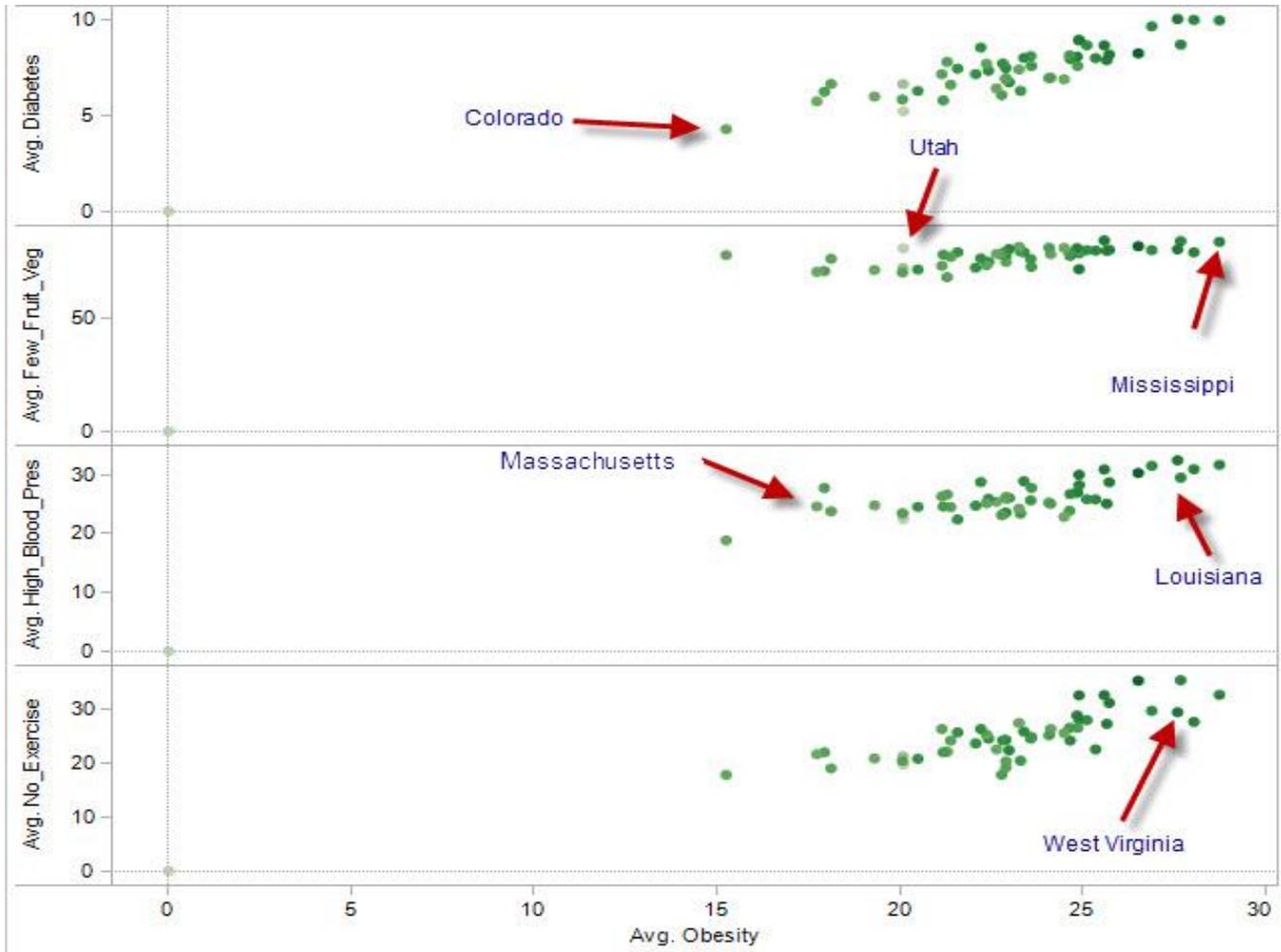
Figure 1 – Visualization of Multiple Health Indicators

As indicated in the visualization, CO has a low level of obesity (closer to 15) on the x axis, a lower smoking rate (lighter shade of green), and a lower level of diabetes since they are towards the bottom of the grid. Conversely, the least healthy states, (MS, and WV, LA) have higher rates of obesity and smoking, and place higher in the grid for each indicator.

3.3   Visualization: Dashboard

The visualization dashboard (Figure 2, next page) created in Tableau is comprised of three areas. The top portion (Indicators) lists the health indicators on the y axis of the dashboard which consists of Obesity, High Blood Pressure, Few Fruits and Vegetables, and the No Exercise, and those indicators are measured by obesity rates. Lighter shaded bars represent states that have lower obesity rates and the height of the bar is influenced by a higher indicator levels. CO which is lightly shaded and has shorter bar indicates a more healthy state in contrast to MS which is more darkly shaded and possesses a higher bar. The legend on the right provides the average rates for obesity. In addition to providing the metrics in numeric form at the top of a bar, a mouse hover will list the state and the average for each indicator.

The bottom left visualization (PCP's per 100,000) provides a snapshot of the number of Primary Care Physicians per 100,000 individuals in each state. MA is identified as the state with the highest number of PCP's and a mouse hover will list each state and their PCP number for quick access to information.

The bottom right visualization (Percent of Uninsured) provides a snapshot of the uninsured as a percentage of population for each state. A highly populated state such as CA has an above average percentage of uninsured but a smaller state such as NM has the highest percentage of all states. Bar's that are lightly shaded represent those that have smaller populations and darker shaded bars indicate more populous states. The legend in the upper right hand corner of the dashboard displays the populations from smallest to largest.
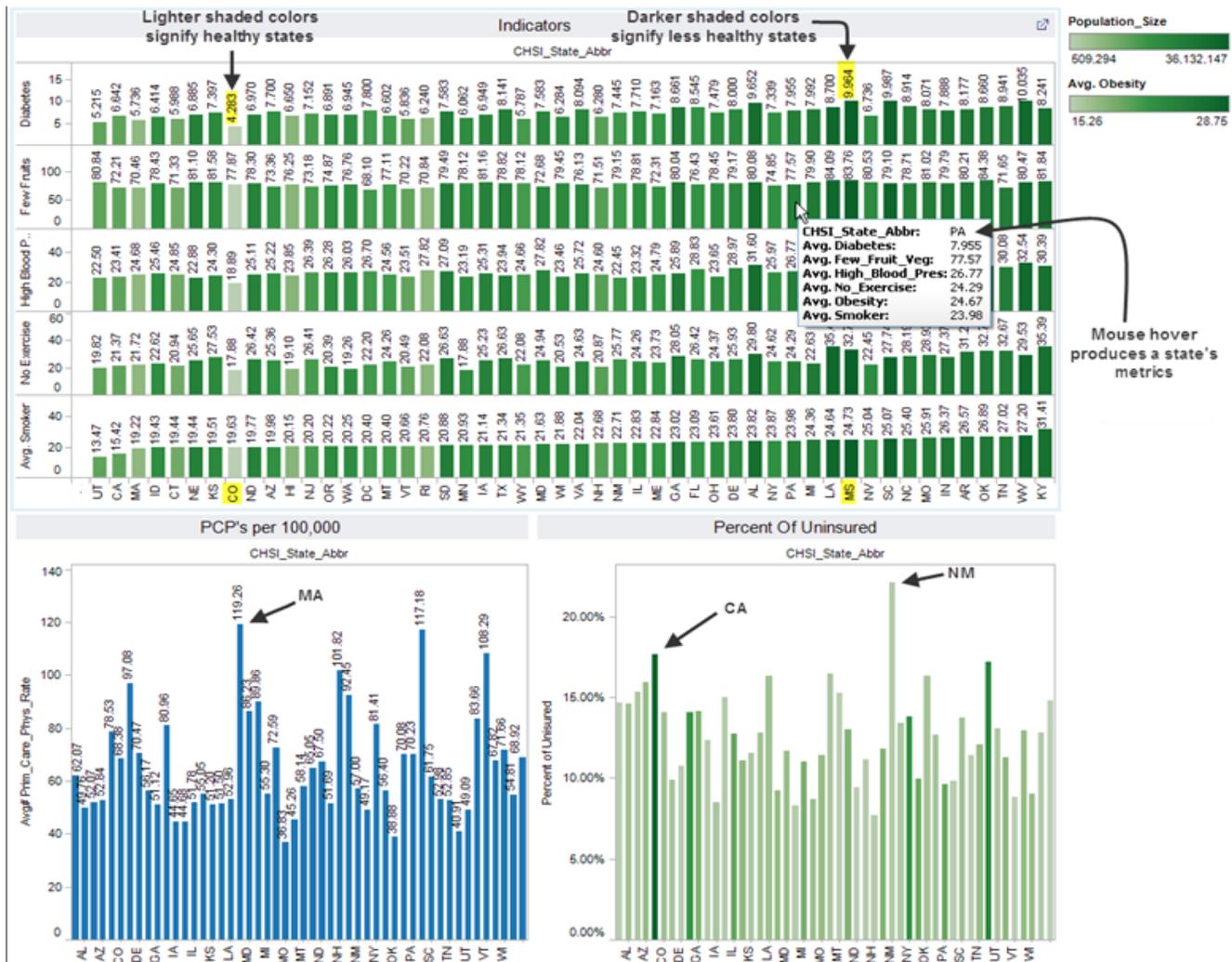
Figure 2 – The dashboard allows the user to group multiple visualizations onto one screen

4   Geospatial Visualization

Because the Community Health Status Indicators dataset was organized at the county level, we believed that an interactive web map would be an effective alternative way to visualize the information. The ArcGIS software suite was used to publish the CHSI data as a service and also aided in creating the application [5]. ArcGIS is the creation of Esri, the premiere geospatial software provider, based out of Redlands, California. The ArcGIS software package includes geospatial solutions for the desktop, server, web and mobile for companies, governments and schools across the globe.

Figure 3 (next page) is the starting page of the resulting web application that is available through any web browser. Users can zoom in and out on a region by using the slider in the upper left-hand corner of the map, with their scroll wheel or by double-clicking. A single mouse click will select a county and display the county's health indicators as long as there is not a null value for that indicator (Figure 4, next page). Alternatively, users can search for a specific county using the search pane in the top left corner of the application. With either selection method, the map is zoomed in to the selected county and the indicators are displayed on the right. This allows for the user to easily and quickly select neighboring counties to compare the indicator results of counties.

The user can visualize the data by five different indicators: uninsured, smoker, diabetes, no exercise or obesity. These "layers" in GIS terms can be toggled by way of buttons in the upper left corner of the application. Each indicator has a single hue color scheme that gives the user a clear understanding of what is being displayed. Additionally, there is a legend in the right pane that explains what each color tone represents.

4.1   Methodology

The application is written in JavaScript, HTML and CSS so that it is lightweight, compatible with mobile devices and plug-in free. It takes advantage of ESRI's ArcGIS JavaScript API on the front-end. The CHSI data was imported into Esri's ArcMap desktop software and joined to 2010 Census county polygons [6] by way of a combination of the county and state FIPS codes located in the CHSI dataset. After the join was performed, each county polygon held its own CHSI data. The counties were then symbolized into five equal quantiles for each of the five indicators (uninsured, diabetes, smoker, no exercise and obesity) within ArcMap. Each of these indicators was then saved as separate "layer" in GIS terms so that they could be selected in the final application by the user. With the visual look realized, the data was uploaded as a RESTful service by ESRI's ArcServer software on an Amazon EC2 Windows Server 2008

instance [7]. When the user requests the map by visiting the URL or pans or zooms in or out, JavaScript events call static tile images of the map that are served by Esri's ArcServer. Microsoft's Information Internet Services software serves the HTML, JavaScript and CSS elements.
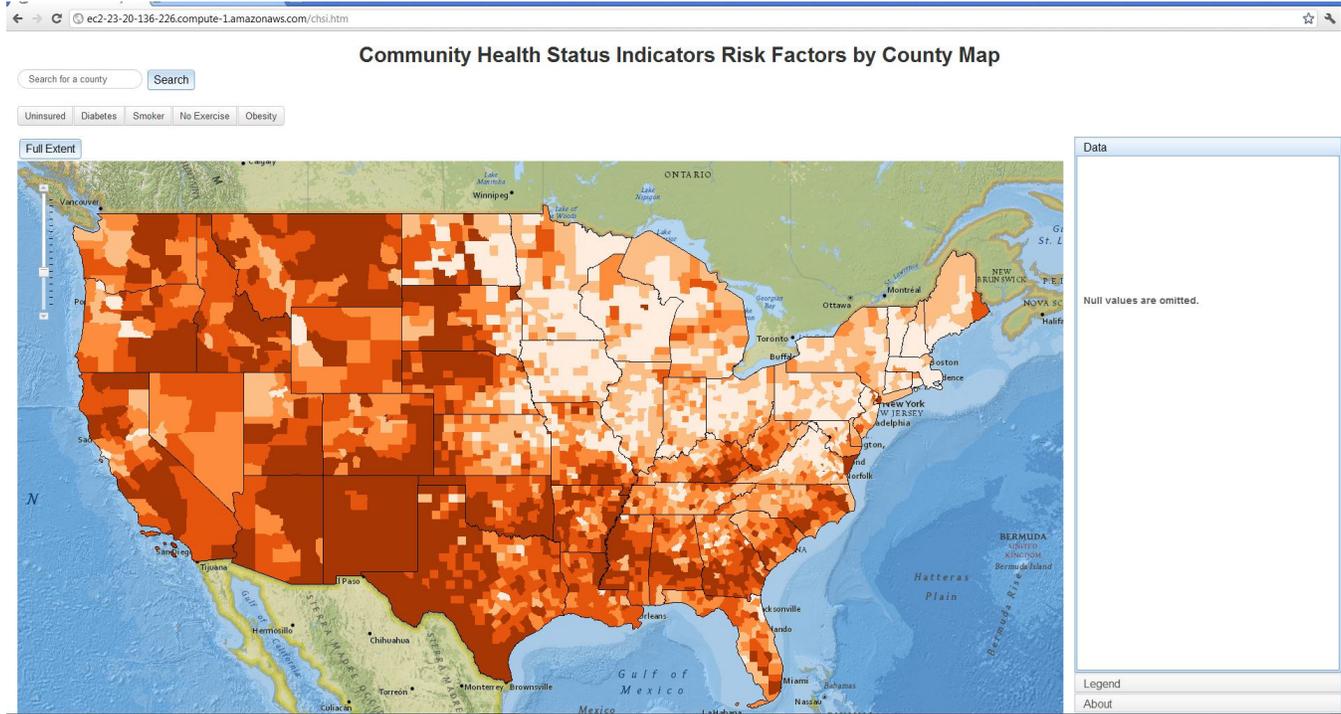

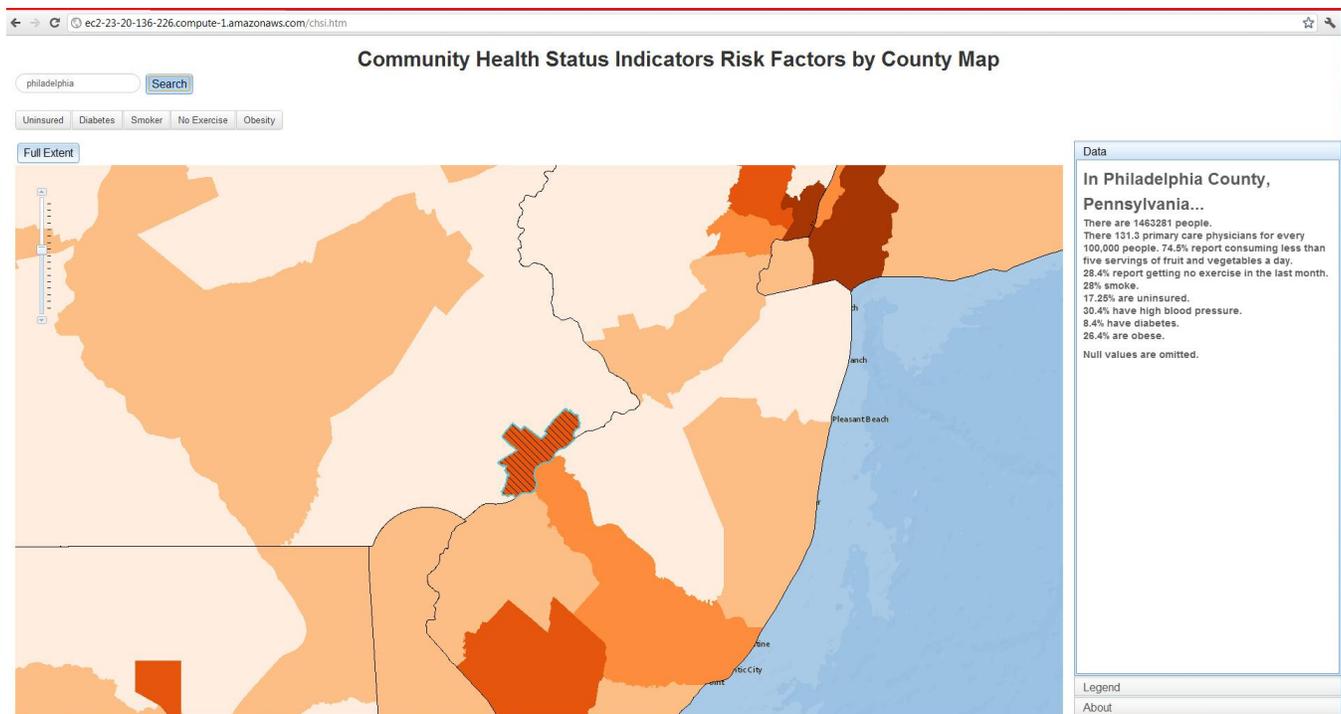**Figure 3 – Community Health Status Indicators Map: Starting Page**

Figure 4 – Community Health Status Indicators Map: Philadelphia Data

4.2 Analysis

The Community Health Status Indicators map is effective in showing spatial patterns in the indicator data. A quick glance at the different layers of indicators provides insight that is not easily gleaned with different visualization methods. For example, the map clearly shows more insured people in the Northeast than the West and South. Overall, the South is significantly more obese than other regions of the country and that indicator mirrors the diabetes rate closely. While smoking patterns aren't as apparent, there is a belt of heavy smoking in Kentucky that goes to the west and south and highly correlates with a lack of exercise. Conversely, places of diverging patterns are also interesting. For example, while there are high numbers of uninsured people in the Southwest of the country, this does not particularly mean that there are also high rates of unhealthy people by way of obesity, no exercise and diabetes rates. This can lead to the conclusion that health in some citizens does not directly correlate to possession of health insurance, which is also apparent in the Tableau visualizations. A video illustrating a walk-through of the map is available at http://youtu.be/3gO5LW01fhY.

4.3 Schneiderman's Principles in GIS Design

Schneiderman's seven tasks were particularly relevant in the creation of the map application for this project. In "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," Schneiderman lays out the seven tasks of an effective visualization: overview, zoom, filter, details-on-demand, relate, history, and extract [8].

Overview concerns the ability of the user to receive a view of the entire collection. This is achieved with a "full extent" button that is standard in GIS software. It allows the user an "escape hatch" of sorts that can be used whenever they are lost in the lower scales or want to see the big picture of the entire country on the map.

Zoom in and out on the map are achieved by using the slider in the upper left corner of the map or with the mouse (scroll wheel or double click). This allows for the user to isolate a certain section of the map at different scales in conjunction with the ability to pan in any direction. Like the overview task, this functionality is standard in web maps. In the past some applications provide a "tool" that when clicked allows the user to draw a bounding box with their cursor on the area that they want to view more closely. This mirrored the GIS desktop experience by the way of a "toolbar" of functionality but has been found in more recent times to not be properly intuitive for the web experience and has been abandoned for the most part.

Filter, or the ability to remove elements that are uninteresting, is not fully implemented in this application. In a way, however, the selection of the different "layers" by way of buttons on the top allow the user to control what visualization they see. Also, the ability to zoom in and out allows the user to filter out what regions of the country they find uninteresting for the ones that they do find interesting.

Details-on-demand is the fourth task and is effectively implemented in the application by way of the "DATA" pane on the right. This pane gives the user details on the selected county that cannot be achieved through one single graphic alone. Additionally, these details are given in the form of a sentence so that they are easily read and understood in contrast to numbers alone.

Although there is no specific relate functionality in this application, the spatial proximity of the counties in the United States form a type of relationship. It is difficult to see how else this functionality can be implemented in the application. One way may be to allow the user to drag two counties to a pane outside of the map pane. This isolation would allow for the two counties to be compared in a more focused way. While this is certainly possible to program, it is not trivial, especially for a prototype such as this one.

The history task is occasionally implemented in web maps by way of "bookmark" functionality. This involves a pane in the browser that allows the user to save the extent that they are currently at, often with the ability to add text for context. At any time in the browser session, the user can select the saved bookmark to go back to that extent.

This functionality wasn't provided in this map, as it often isn't elsewhere, because there doesn't seem to be a need for it. The "full extent" ability is standard on web maps and allows the user to start over if they need to. Such interactive map applications are easy to navigate, so zooming and panning to a previous extent is neither time-consuming nor difficult.

Another point of view is that the technology is still maturing. Typically bookmarks do not persist over multiple visits or page refreshes. However, the HTML5 specification allows for persistent storage on the user's computer. This makes it possible to save bookmarks that are available for multiple visits to the map. Even so, web maps are not used so extensively that this is necessary.

The final task is extract, or the ability to take the data outside of the visualization. This can be in the form of a tabular and graphic file in a variety of formats. It could also include the ability to email or send the data over other means. This is possible and sometimes implemented in web applications, but requires server-side logic that costs time and resources.

As web map applications continue to develop and add statistical and spatial functions, in the future people will spend more time using them. At that time the history and extract features will increase in importance because users are interacting with map applications more in general. Technologies like HTML5, jQuery and other JavaScript libraries are pushing the limits of GIS in the browser to give more robust feel and functionality similar to the desktop experience.

5 CiteSpace II

Citespace is a visual representation tool that is available for download at: http://cluster.ischool.drexel.edu/~cchen/citespace/download.html [9]. CiteSpace was developed at Drexel University and is used to analyze and detect patterns found in scientific literature by detecting co-citation clusters. The software can synthesize information from multiple sources including PubMed and the Web of Science. CiteSpace can analyze bibliographic data in time splices to determine relevant coupling and visualizations can be created to identify citation bursts that place interesting scientific topics along a time continuum to understand domain knowledge patterns.

5.1 Methodology

In addition to analyzing the CHSI data, we were curious to learn about the academic research that has been done on the topic of chronic health issues in the last 12 years. We utilized CiteSpace for this analysis and downloaded citations from the Web of Science, covering the years 2000 – 2012 with the words "chronic health" in the title. The Web of Science returned 1,803 records, of which we downloaded the first 1,500. We analyzed the records in two passes, first clustering by document, then clustering by author.
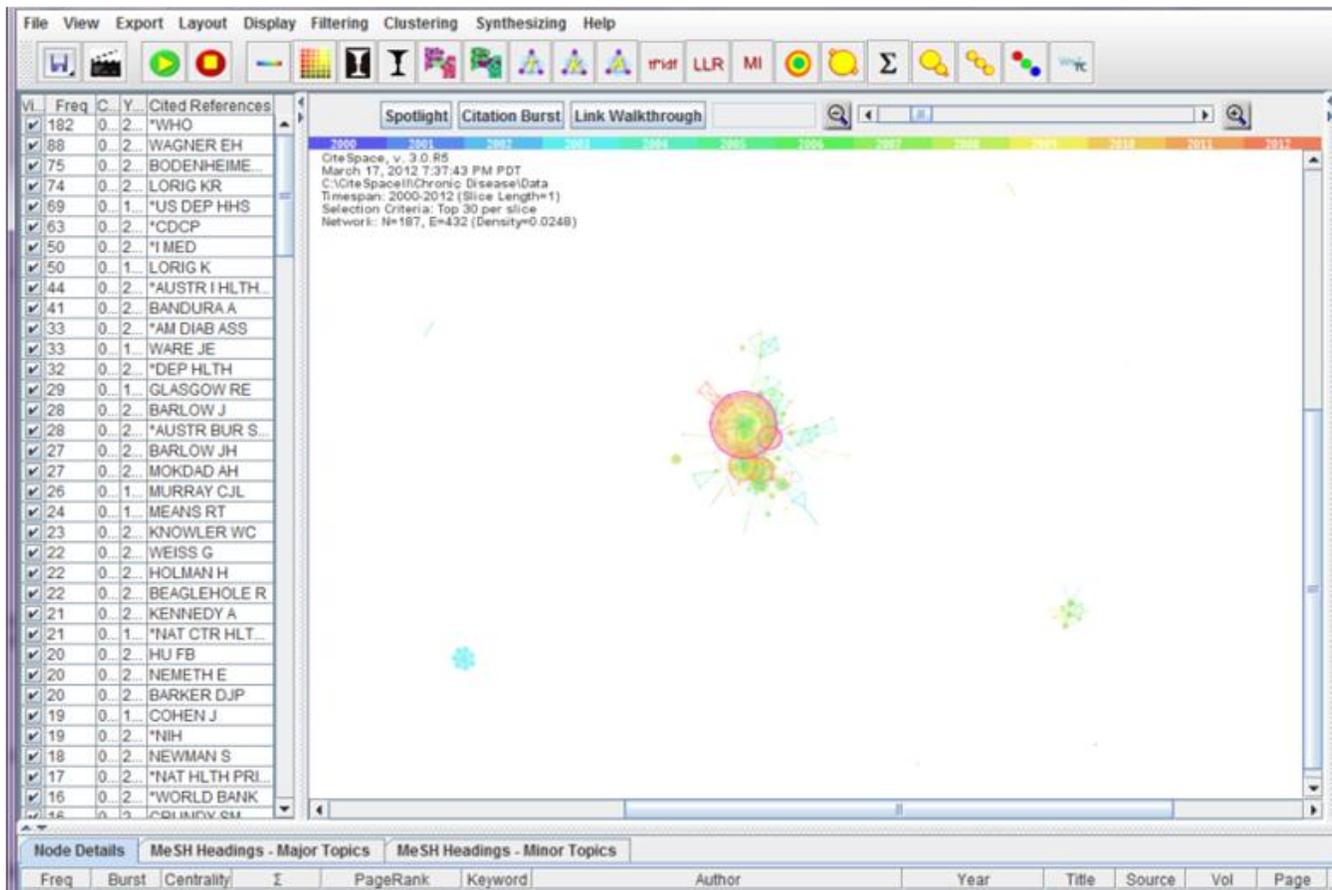
5.2 Analysis

Clustering by document, CiteSpace imported 1,472 of the 1,500 records and divided them into 33 co-citation clusters. The cluster visualization is shown in Figure 5. Interestingly, the clusters are all in the early to middle years of the time-frame we analyzed, suggesting that some significant research was conducted and reported during those years.

Figure 5: CiteSpace II Chronic Health Citation Clusters

The top ten clusters, their sizes, and index terms are as follows:

Cluster	Size	Index Terms
7	28	Obstacles
11	23	Scale, chronic disease
23	15	Estrogen plus progestin, aging
6	13	Left-ventricular hypertrophy, anemia
12	13	Lessons, education-programs
20	13	Middle-income countries, burden
31	12	Iron, inflammation
15	11	Longer-term outcomes, chronic illness, improve
28	8	Delivery of health care, disease management
14	7	Organizations, chronic illness, peer-support

From analyzing the top index terms indicated above, it is clear that the management of chronic health issues, patient education, and obstacles to improving health have been of great concern to academic researchers in the past 12 years. Zooming in on the top five citations, we found that four of the top five citations are in the top 10 clusters. Clearly, Lorig KR performed some research that others found quite useful, as the author had two of the top five citations.

Citation	Cluster	Co-Citations
Lorig KR (1999)	11	44
*WHO (2005)	20	36
Lorig KR (2001)	11	33
Wagner EH (2001)	16	33
Bodenheimer T (2002)	15	31

A visualization of the clustering of the same records, but by author, is shown in Figure 6.

Figure 6: CiteSpace II Chronic Health Author Clusters

Through this anlaysis, we found that, perhaps not surprisingly, the author with the greatest citation count is *WHO (2002), with a citation count of 182. Other top authors include *US DEP HHS (1986), ranked 5th with a citation count of 69; *CDCP (2004), ranked 6th with a citation count of 63; *I MED (2002), ranked 7th with a citation count of 50; and *AUSTR I HLTH WELF (2002), ranked 9th with a citation count of 44. That five of the top 10 most-cited authors are a global health organization (WHO), various national organizations (US Dept of HHS, Center for Disease Control and Prevention, and the Australian Institute of Health and Welfare), and a non-profit organization (the Institute of Medicine) suggests both that a significant amount of research is conducted by such organizations and that – generally rightly so – such organizations are considered experts in the study of chronic health issues. The use of visualization tools by these organizations to analyze and illustrate their findings could potentially help the information visualization community expand its reach to other organizations that find the results of these organizations interesting and useful.

6   Obesity and Crime

Another area we explored was prompted by a paper by Gregory N. Price titled "Obesity and crime: Is there a relationship?" [10]. The paper states that, "our results are consistent with obesity increasing the incentives an individual has for participating in criminal activities" and that the researchers' results suggest that reductions in the incidence of obesity will improve public health and public safety by reducing crime. While exploring the associations of obesity and community we wanted to examine the possible association between obesity and crime using the CHSI data. We combined the data from the CHSI dataset with data from the FBI database on the four most crime-ridden states: California, Georgia, New Mexico and Texas.

The combined dataset contains the original CHSI data with the incidences of murder, forcible rape, robbery, aggravated assault, etc.

We then used the ManyEyes application to visualize the dataset with TreeMap visualizations. The group created treemaps for the four states of CA, GA, NM and TX to highlight their obesity, cases of theft, and cases of aggravated assault by county. On visual comparison there does not seem to be much correlation between the counties with high obesity and the counties with high cases of theft and assault. One of the problems with the visual examination is the unfactored data source which did not take into consideration the Cox proportional hazard. However, this does not mean that a linkage does not exist. Proper statistical methods need to be applied to the data source in order to verify whether or not a link exists. Further, more time-intensive, research is definitely warranted. The visualizations created during this analysis are available in the attached appendix.

7   Discussion

Our study of the CHSI data only reviewed health indicators for one year but having access to CHSI data from previous years would provide the ability to discover and track trends on an annual basis and measure indicators and their influence on health rates in time splices. Expanding the dataset to track other health indicators could provide additional insights and their impact on healthy attitudes. A further analysis could look at states within regions since it appears that states in the Northwest have healthier lifestyles than those in the Southeast. A closer look at cultural indicators may provide a deeper understanding on the relative health of one state versus another and it may reveal why CA, which has the second lowest smoking rate, does not have similar healthy indicators when measured. Although for such a large population CA represents an overall healthy lifestyle. Further analysis is needed to determine why states such as KY, WV,

and TN which border one another have such high smoking rates and what makes CO so healthy. As we strive for healthier attitudes and lifestyles future datasets will provide a multitude of indicators to quantify. Further analysis may reveal insights that affect individual choices, public education and policy, and criminal activity..

8   Conclusion

Through ArcGIS and Tableau we have shown ways to visualize the CHSI dataset. Our results were not only visually interesting and information rich but also in line with the latest health trends as reported by Gallup. According to their report, chronic conditions are more prevalent in the most obese metro areas [11]. Our findings using the CHSI dataset corroborate the Gallup findings. Another poll done by Gallup identified West Virginia, Kentucky and Mississippi as the three unhealthiest states [12]. Our own analysis, which was based upon a different data source, suggested that Mississippi, West Virginia, and Louisiana were the three unhealthiest states.

While researching this project, we found similar attempts have been made to visualize health indicators using various data sources. One example is the County Health Profiles project developed by GE Visualizations. This project tracks many of the same indicators as CHSI, however their collection methodologies differ. Interestingly, both our own project and the GE Visualizations project showed that southern states tend to have higher rates of obesity and that north-eastern states tend to have the highest levels of insured individuals. These findings validate our findings via different visualization techniques and data sources. One of the key takeaways from this project has been that it is possible to produce effective and accurate visualizations in a cost-effective manner.

The CHSI web map at this point is only a prototype that demonstrates functionality that should be included in an interactive spatial visualization of the health data. Additional work needs to be done on the overall styling of the page. Back-end optimizations can be implemented, such as caching the data so that the map loads faster. User testing would be beneficial in coming to conclusions on what is effective in terms of user experience. In order to accurately gauge how well this model works, we hope to gather information directly from actual CHSI GIS users and analysts. It will be very useful to learn how users and analysts interact with the GIS software and get their feedback on how to improve the visuals. This information will provide important insight into how effective we are in conveying the indicators to the intended audience. Nevertheless, the application can be useful for visualizing any data at the county level and as open source will be available to anyone that would like to use and improve upon it. By having a better understanding of how people view, use and respond to these kinds of maps we are in a better position to communicate public health data for policy and action. The source code is available at https://github.com/davewalk/CHSI.

While not as interactive as the CHSI map, the Tableau software is an excellent tool for comparing a variety of health factors and their related outcomes. Although we only used the educational-licensed version of the software, we noted that creating dashboards with multiple data sources was very easy due to Tableau's efficient user interface. Tableau also offered more variety in terms of data visualization techniques.

As communities focus their efforts on educating the public on healthy lifestyles, referenced research material will provide valuable insight on the topics and trends to be used for discussion. Easy to read public service literature can be created so all members of the community can benefit from the findings. These resources should be readily available in paper format and electronic versions so they can be accessed in public spaces, community libraries, and community websites. Healthy lifestyles workshops and forums can be created for those that seek additional information pertaining to a

chronic disease in which they have an interest. Software such as CiteSpace II can assist in determining trends and patterns in the literature that can be used to trace the insights found in visualizations generated in Tableau and ArcGIS from CHSI statistics. The combination of divergent visualization tools can become quite useful for analysis as communities strive for more healthy attitudes and lifestyles.

References

[1]   U.S. Department of Health and Human Services. Community Health Status Indicators.
      http://www.communityhealth.hhs.gov/homepage.aspx?j=1.
[2]   Hewlett-Packard Development Company, LP. "The Power of Visual Communication." (2004).
      http://www.hp.com/large/ipg/assets/bus-solutions/power-of-visual-communication.pdf.
[3]   MacEachren, AM. "How maps work: representation, visualization, and design." (1995). New York (NY). The Guilford Press.
[4]   Tableau: http://www.tableausoftware.com
[5]   ArcGIS: http://help.arcgis.com/en/webapi/javascript/arcgis/
[6]   U.S. Census Bureau 2010 Data:
      http://www.census.gov/geo/www/tiger/tgrshp2010/tgrshp2010.html
[7]   Amazon EC2: http://aws.amazon.com/ec2/
[8]   Shneiderman, Ben. "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations." (1996). 18 Jan. 2012.
[9]   Chen, C. (2006) CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. Journal of the American Society for Information Science and Technology, 57(3), 359-377.
[10] Price, Gregory N. "Obesity and crime: Is there a relationship?". Economics Letters, Volume 103, Issue 3, June 2009, Pages 149-152, ISSN 0165-1765, 10.1016/j.econlet.2009.03.009.
      http://www.sciencedirect.com/science/article/pii/S0165176509000834.
[11] Witters, Dan. "More Than 15% Obese in Nearly All U.S. Metro Areas." (2012).
      http://www.gallup.com/poll/153143/Obese-Nearly-Metro-Areas.aspx.
[12] Stockdale, Charles B., Michael B. Sauter, and Ashley C. Allen.. "America's Most Miserable States." (2012).
      http://247wallst.com/2012/03/02/americas-most-miserable-states.