

VisUFO: Organization and Representation of UFO sighting data.

(December 2014)

Manasa Bulusu, Colby Ford, Prajakta Patil, Fnu Aishwarya Ramesh, and Chuan Zhang

The University of North Carolina at Charlotte

Abstract— VisUFO is a visualization system that integrates multiple charting techniques to view geospatial data along with other attributes. The D3.js javascript pack was used to design and create the visualizations. The methods combine geospatial mapping with parallel set charts. The map uses choropleth of population density to unearth interesting relationships between population and UFO sightings. The parallel sets allows users to explore details of the UFO sightings such as duration and shape by region. Users can interact with the system in a simple yet intuitive manner to observe areas of interest based on overview first and details on demand methodology.

Index Terms—UFO sightings, Visual Analytics, Map, Parallel Sets, UFO Visualizations, Choropleths, Population Density

© 2014 Bulusu, Ford, Patil, Ramesh, and Zhang
The University of North Carolina at Charlotte

-
- *M. Bulusu is a graduate student in Data Science and Business Analytics at the University of North Carolina at Charlotte, Charlotte, NC 28212.
E-mail: mbulusu@uncc.edu.*
 - *C. Ford is a graduate student in Data Science and Business Analytics at the University of North Carolina at Charlotte, Charlotte, NC 28212.
E-mail: cford38@uncc.edu.*
 - *P. Patil is a graduate student in Data Science and Business Analytics at the University of North Carolina at Charlotte, Charlotte, NC 28212.
E-mail: ppatil@uncc.edu.*
 - *A. Ramesh is a graduate student in Data Science and Business Analytics at the University of North Carolina at Charlotte, Charlotte, NC 28212.
E-mail: aramesh@uncc.edu.*
 - *C. Zhang is an under graduate student in Computer Science at the University of North Carolina at Charlotte, Charlotte, NC 28212.
E-mail: czhang@uncc.edu.*
-

using injury data. The data is laid on a map using geospatial information and then pie charts are used to represent the breakdown of injuries by type. As the user zooms in and out, the pie charts group together to keep the visualization neat and a summary of the data.

View the visualization at: <http://blocks.org/gisminister/10001728>

Markercluster pie charts

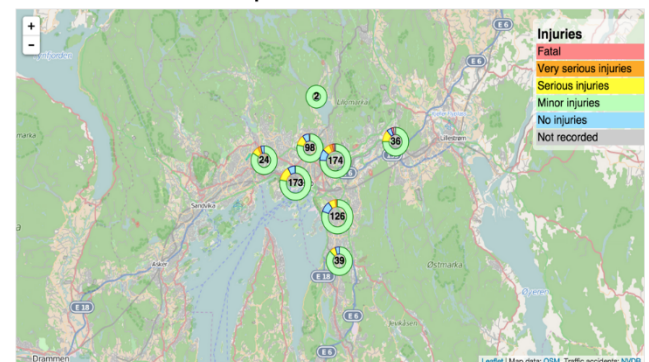


Figure 1: Markerclusters Pie Charts

1 INTRODUCTION

VisUFO is a system designed to represent UFO sighting data in multiple ways. Users may select to view the information geospatially or temporally and view the information by type of sighting, etc. Occurrence data is often hard to represent visually since it occurs over a wide area and includes various attribute. With the UFO sightings, there is information about size, sighting duration, and location. This cannot be shown by map alone. Thus, VisUFO incorporates two types of charts to gain insight into the data, allowing the user to possibly identify patterns or to view the data as a whole.

2 RELATED WORK

2.1 MARKERCLUSTER PIE CHARTS

A work by the user gisminister on blocks.org has created a chart

2.2 MACROMETEORITES

Macrometeorites is a system developed by Roxana Torre. This system plots meteorite fallings and findings as a temporal axis plays through time. This allows for the user to see the multiple sizes of meteorites that have fallen and been found throughout time. Also, the system includes tooltips with more valuable information upon hovering over a data point.

View the visualization at: <http://visualizing.org/visualizations/macrometeorites>

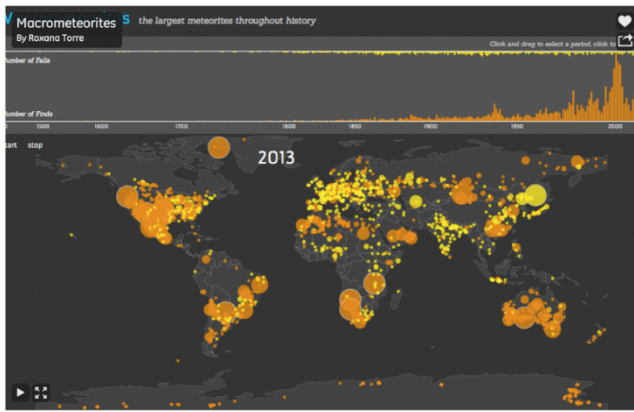


Figure 2: Macrometeorites

2.3 CDC GIS CHRONIC DISEASE EXCHANGE

The American Centers for Disease Control and Prevention hosts a GIS exchange for disease occurrence data. The site contains information about research on how to efficiently represent disease data using GIS information and allows users to upload their data and/or visualizations to the site for use publically.

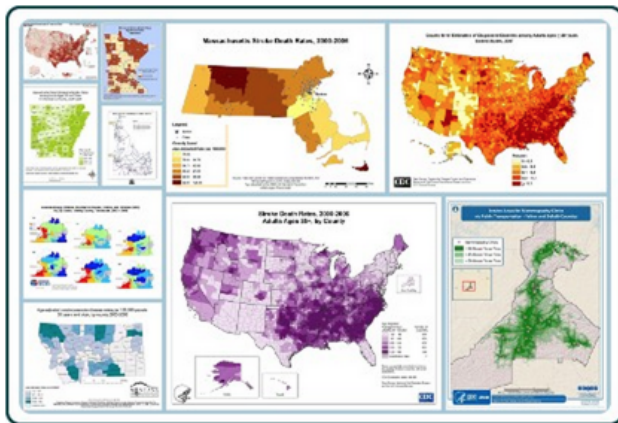


Figure 4: CDC GIS Chronic Disease Exchange

The map gallery of the site can be found here: <http://www.cdc.gov/dhds/mgs/gisx/mapgallery/index.html>

2.4 PROTOVIS

A visualization created by the Stanford Visualization Group called Protovis plots multidimensional data in parallel coordinates. This Allows users to view relationships between multiple attributes at once.

In their example, car data was used to build the visualization that can show the relationship between a car's acceleration, horsepower,

Parallel Coordinates



mileage, etc. in an easy plot.

View this visualization at: <http://mbostock.github.io/protovis/ex/cars.html>

3 SYSTEM DESIGN

3.1 Initial design explorations

Our design discussions have been motivated by important design considerations such as automation, user interactions and use of intuitive design elements. As per Schneider's well-known InfoVis mantra [1], we wanted to present an overview first and details on demand thus providing users with the flexibility of exploring the system at different levels for analysis. That being said, we wanted to partly automate the system to allow users to focus on exploration.

Figure 3: Protovis

The partly automated features of the system includes a Choropleth map of US population density over which glyphs of UFO sightings are positioned to allow users to view geo-location of the data. The parallel sets view provides an effective way to view categorical data of the UFO sightings such as shape and duration of the UFO sightings by region. Here, the regions for the geo locations are computed automatically.

Our initial list of exploratory features included using stacked bar charts on a 3-D map, buttons and filters and coordinated views for the map and parallel sets. However, we found many challenges in implementing a 3D map and hence used 2D which in fact gives a more clutter free overview of the geo locations. We deliberately kept the map and parallel sets independent in this version since we believed it will be more useful to see an overview of the sightings on the map while analyzing details of shapes and duration of UFO sightings by region using the parallel sets.

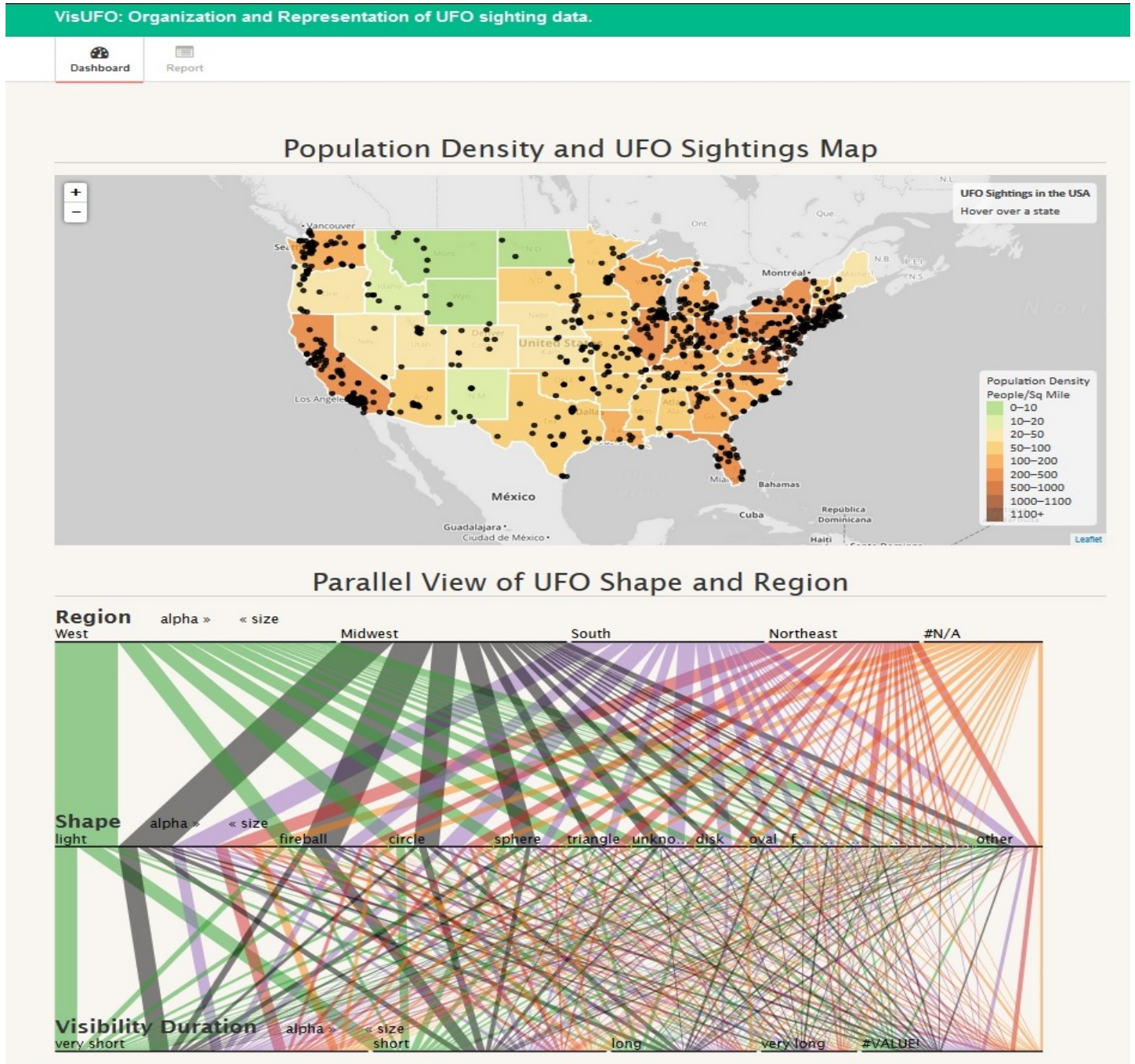


Figure 5: VisUFO System

3.2 VisUFO system

3.2.1 Visual design

For geo locations based multivariate data, we believe the most intuitive and effective visualization is a simple map. However, in the case of UFO sightings, we were interested in exploring additional geo-location features such as population density and/or airport bases. To achieve this, we used a Choropleth map that codes different states with variations in a divergent color scheme to differentiate between highly populated and sparsely populated areas. This is a simple and effective way to unearth some very interesting facts and relationships that go beyond basic location codes. We use additional glyphs on the

Choropleth Map to represent multivariate data points. The objective is to use color, size and shape of glyphs to display multiple attributes of the data points. In the current implementation, we have used circles to represent data points on the Choropleth map. This is a basic implementation and will be extended to accommodate various other attributes such as frequency, duration and shape using size, color and shape in the future.

Parallel sets is an effective visualization for categorical data. We use parallel sets to visualize and allow user exploration for lower level details. While it's powerful for smaller data points, for higher number of data points, it introduces clutter. The parallel sets displays UFO sightings by regions (to reduce clutter), shape and duration. In the data preparation stage, we code each sighting by Region – Midwest, Northeast, West and South. For duration, we converted durations recorded for each sighting to seconds and code them as very short

(<=1 min), short (1min< t< =5min), long (5min< t< =20min), very long (>20 min) for easier visualization. Each of these attributes can be arranged alphabetically or by size and reordered by interacting with the visualization. Tooltips are provided for providing more details to the user.

3.2.2 System Interactions

Glyph based Choropleth Map provides interactions on mouse hover, clicks and wheel. The population density values are displayed in the info box on top right on mouse hover of States. Mouse click and wheel allows users to zoom on the selection state. Fig 6 show the population density based choropleth map where states are coded with different color ranges based on their values. Fig 7 shows a zoom-in interaction of California State.

Parallel sets uses horizontal bars to show the absolute frequency of how often each category occurred. Between the dimension bars are

ribbons that connect categories and split them up. This visualization allows users to sort categories - Region, Duration and Shape based on Alphabetical order or Size in ascending or descending order. When a user selects mouse move/click, a tooltip with additional data becomes visible. The bar gives absolute numbers and percentages for each category as a fraction of the entire dataset. On moving the mouse over a ribbon, it gets highlighted and displays the combination of criteria that the ribbon displays along with absolute numbers and percentages. Categories can be rearranged by clicking and holding the mouse on one and dragging it around, this is seen in Fig 9. This rearranging removes unwanted clutter and gives flexibility to the user to focus on categories and attributes of interest. In Fig 8, we see that the combination West (Region) -> light (shape) -> very short (duration) has maximum absolute frequency of maximum percentage of 3%.

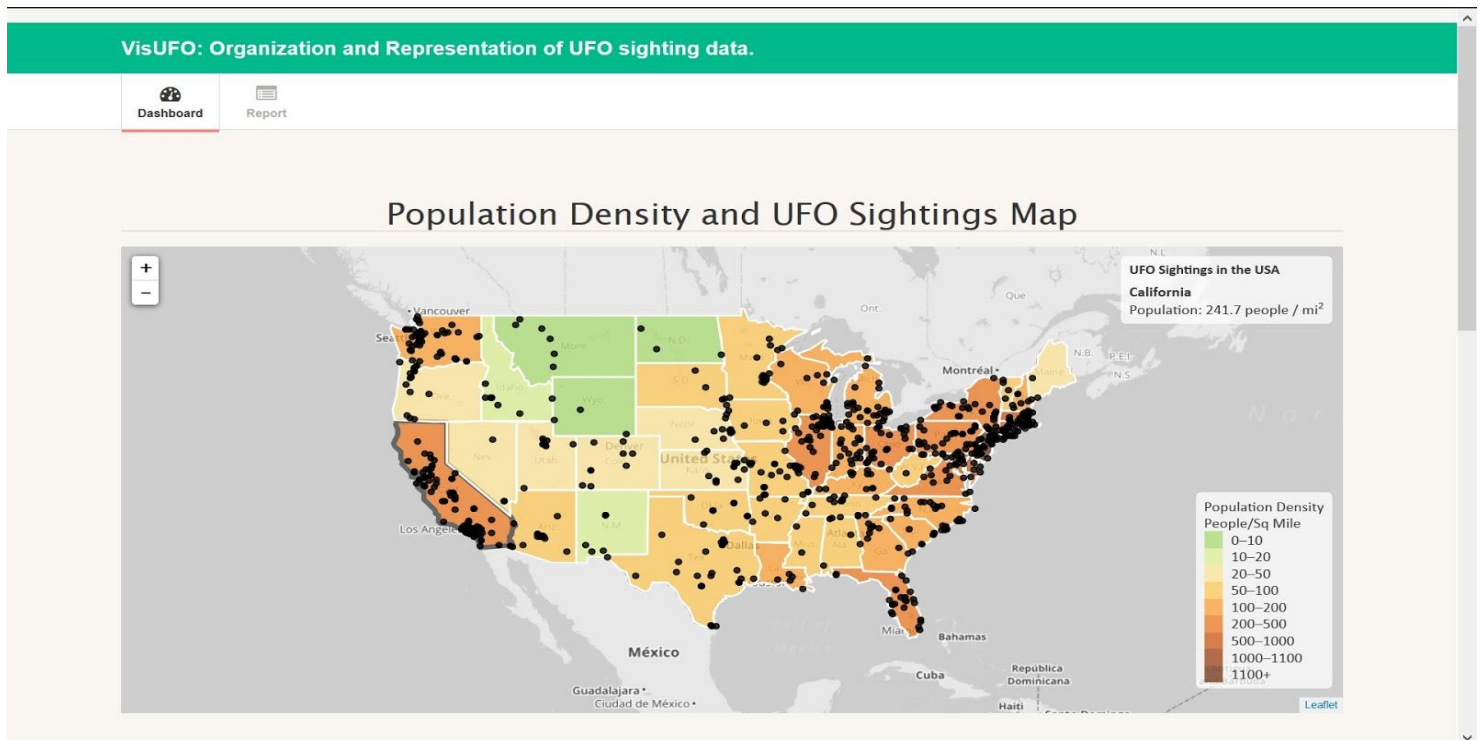


Figure 6: Glyph based Population Density Choropleth Map

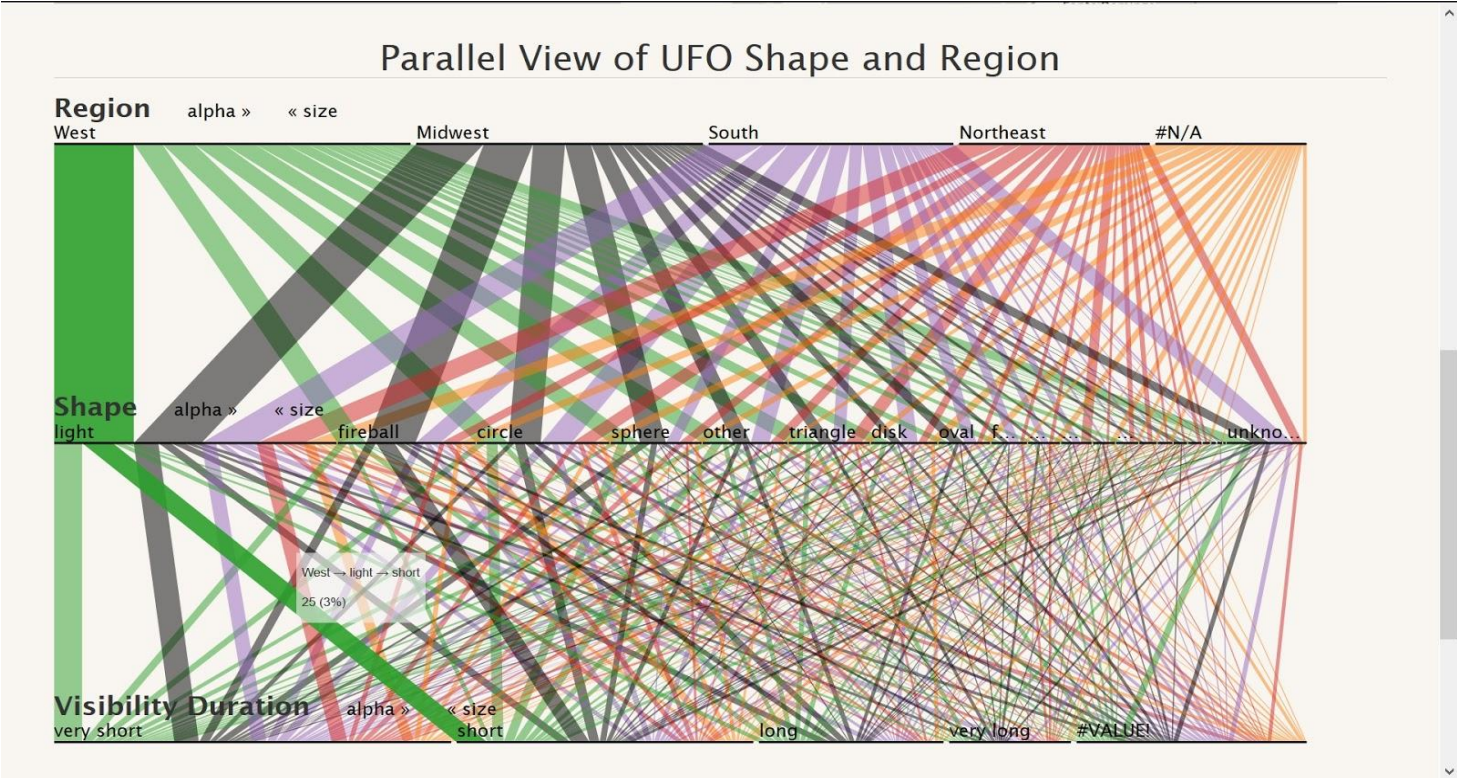
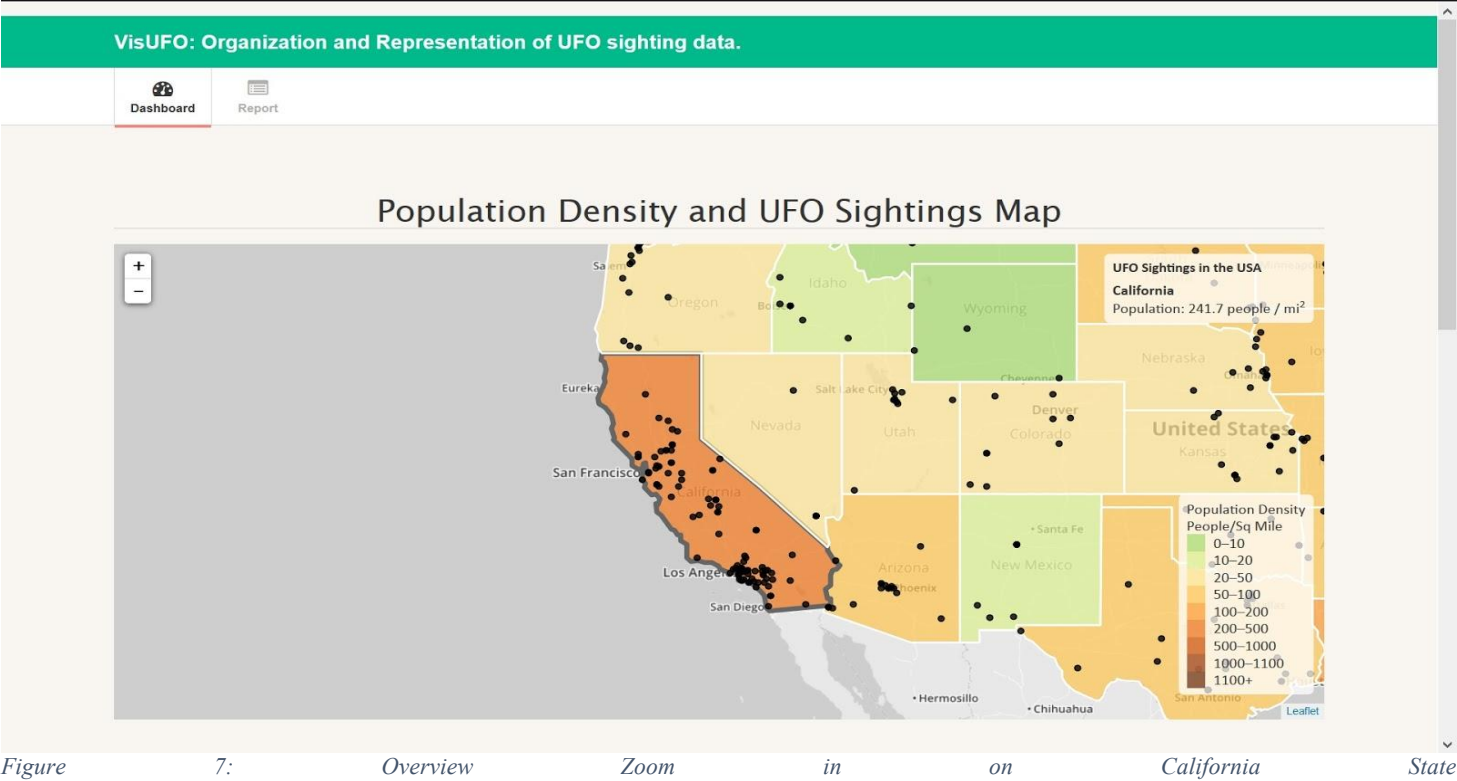


Figure 8: Parallel Sets with tooltip

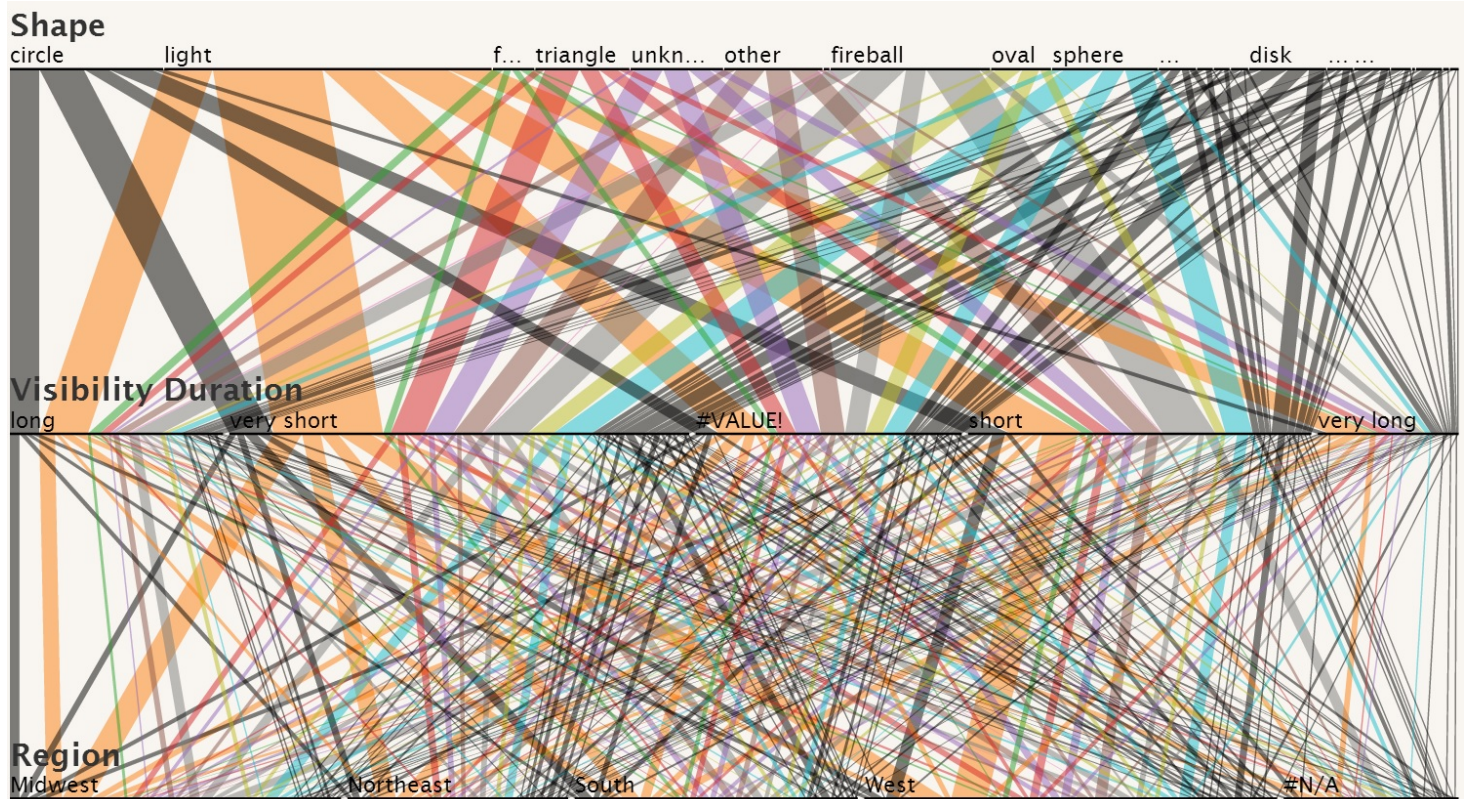


Figure 9: Parallel Sets Rearrange categories to reduce clutter

4. IMPLEMENTATION

The prototype of the system runs on a web browser and is written in JavaScript using the D3 visualization toolkit. We have used Color Brewer for the diverging color scheme for the population density Choropleth map. The visual elements and the dashboard was developed in HTML. The system can be accessed at <http://webpages.uncc.edu/~cford38/research/d3visufo/>.

5. EVALUATION

We conducted a user survey to get a feedback about our system. We asked 5 users to use the system and comment on the usability and understandability of the system. The users were presented with the website <http://webpages.uncc.edu/~cford38/research/d3visufo/>. The users liked the way the system showed correlation between population density and the sightings of UFO. Some of the major recommendations they provided were (1) To present more information (duration, frequency) about each data point (circle) in a tool tip on mouse hover, (2) On being asked if they preferred big circles corresponding to number of sightings per state to multiple small circles, majority responded to retain small circles since it reflects correctly the number of sightings in that area instead of big circles which gives a high level knowledge of the sightings, (3) To present both visualizations – choropleth map and parallel sets side by side instead of one below the other to consume their content easily instead of having to scroll up and down the page, (4) Coordinate both the views since the map alone does not suffice in giving out all the necessary information.

6. FUTURE WORK

Based on the results from the user study, we have considered adding a

few more features to VisUFO to make it more accessible and accommodating. In the second stages of VisUFO, we plan to enhance the visualization with the following features

Stacked 3D Bar Charts – VisUFO currently uses circular markers to represent UFO sightings on the US map. We would like to replace these circular markers with stacked 3D bar charts, each bar chart representing a UFO sighting. The height of each bar chart will represent the total number of sightings for that location. Each bar is divided into colors which depict the shapes of UFOs sighted. This way, first look at the visualization provides lot more information about population density, number of sightings per location and sightings by shape per location.

Airbases/Airports - In order to peak the interests of UFO enthusiasts, we plan to include Airbases' and Major Airports' information onto our map. This facilitates comparison of UFO sightings and the frequency of airplane/aircraft sightings at a particular location.

Coordinated Views - VisUFO uses two visualizations, the map and the parallel sets. Parallel sets provide an interactive way to explore categorical data. User study suggested a need for coordinating the parallel sets with the map to provide a parallel coordinated view, making it a singular visualization. Coordinated views empower information exchange and ease user collaboration.

Temporal Animation - As of now, VisUFO uses only static data for UFO sightings visualization. We consider adding a dynamic temporal representation of UFO sightings based on their sighted dates. We plan to integrate the temporal representation with the US map, creating an animation of UFO sightings from the earliest sighted years to the most recent years, categorized by color representing shape.

Application to other fields of study – This visualization system will work for many fields. From epidemiology to any other occurrence-based data, the style of the geospatial visualization and the parallel sets is relevant to many other areas other than just the superficial UFO sighting data.

7 CONCLUSION

Although the VisUFO system is meant to visualize UFO sighting data, our system has implications in many fields such as in epidemiology with disease prevalence, political science with voting statistics, and many other areas.

ACKNOWLEDGMENT

The authors wish to thank Dr. Jing Yang of the University of North Carolina at Charlotte for the opportunity to do this research project.

REFERENCES

- [1] B. Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of IEEE VL '96*, pages 336-343, 1996.
- [2] Koua, E.L.; Kraak, M.-J., "A usability framework for the design and evaluation of an exploratory geovisualization environment," *Information Visualisation*, 2004. IV 2004. Proceedings. Eighth International Conference on, vol., no., pp.153,158, 14-16 July 2004
- [3] Ping, Y., Xinming, T., Shengxiao, W. "Dynamic cartographic representation of Spatio-Temporal data". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII(XXXVII-B2):7-12, 2008
- [4] Zhiyuan Zhang, Xiaonan Tong, Kevin T. McDonnell, Alla Zelenyuk, Dan Imre, and Klaus Mueller. "An Interactive Visual Analytics Framework for Multi-Field Data in a Geo-Spatial Context." *Tsinghua Science and Technology*, 18(2):111-124, April 2013
- [5] Matthew O. Ward. 2002. "A taxonomy of glyph placement strategies for multidimensional data visualization". *Information Visualization* 1, 3/4 (December 2002), 194-210.
- [6] M. Nöllenburg, "Geographic Visualization"; in *Proc. Human-Centered Visualization Environments*, 2006, pp.257-294.
- [7] An Oracle White Paper (April 2010)
- [8] R. Borgo, J. Kehler, D. H. S. Chung, E. Maguire, R. S. Laramée, H. Hauser, M. Ward, M. Chen *Eurographics State of the Art Reports*, 2013
- [9] Patrick Job Ogao, ISBN 90-6164-206-X ITC, Dissertation number 89
- [10] NIJ Award 2009-SQ-B9-K009, Dr. William Ribarsky(PI), Dr. Kalpathi Subramanian(Co-PI) Department of Computer Science, University of North Carolina at Charlotte Federal funds provided by the U.S. Department of Justice, October 1, 2009 - August 31, 2012.
- [11] Stefan Seipel and Leonor Carvalhaz Department of Industrial Development, IT and Land Management, University of Gävle, Sweden Centre of Image Analysis, and Uppsala University, Sweden, University of Turku, Finland
- [12] Thomas Kapler and William Wright Oculus Info Inc. (www.oculusinfo.com)
- [13] Louise S. Hadden, Abt Associates Inc., Cambridge, MA Mike Zdeb, Albany School of Public Health, Rensselaer, NY Robert Allison, SAS Institute Inc., Cary, NC http://support.sas.com/rnd/datavisualization/maponline/html/geo_code.html)
- [14] Kartograph (www.kartograph.org) Google Developers – Visualization Geo Map

(<https://developers.google.com/chart/interactive/docs/gallery/geomap>)

Manasa Bulusu A Bachelor of Science in Biotechnology, an MBA graduate and a graduate student in Data Science, Manasa is a research enthusiast. She has research interests in fields of biological sciences, marketing research and analytics. Manasa is currently a Data Science graduate student at the University of North Carolina at Charlotte, prior to which she was associated with a variety of industries, Pharmaceutical to Retail Banking, in the field of Marketing. Alongside academic activities, Manasa engages in photography and playing the guitar as an amateur.

Colby Ford B.A. in Applied Mathematics, M.S. in Data Science and Business Analytics anticipated May 2015. Colby has research interests in the biological sciences and the overlay into data science. He is currently a graduate student at the University of North Carolina at Charlotte and works for Mariner in Charlotte, NC as a Statistical Data Scientist specializing in statistical modeling and visualization design and is a frequent speaker on the subject of Data Science at companies like Microsoft and others. He is a member of the American Statistical Association and the Society for Industrial and Applied Mathematics.

Prajakta Patil Bachelor of Engineering in Computer Science, Graduate student in Data Science and Business Analytics. She has research interests in Data Analytics. Prior to her current graduate program she was associated with an industry in Datawarehousing and Business Intelligence.

Fnu Aishwarya Ramesh Bachelor of Engineering in Instrumentation Technology, Masters in Strategic Management from Cass Business School, London, Aishwarya has research interests in data sciences, business strategy and analytics. Aishwarya is currently a Data Science graduate student at the University of North Carolina at Charlotte, prior to which she has worked in Broadcast & Consumer Electronics industries in technical as well as strategic roles. She is a lifetime member of the esteemed Beta Gamma Sigma Society.

Chuan Zhang is an under graduate student in Computer Science at the University of North Carolina at Charlotte