Bigfoot, UFOs, and Where to Find Them

Kate Lanman* Ziyu Prince Wang† Jacqueline Deieso‡ Brady Hobson§ Sara Kent¶

ABSTRACT

UFO and Bigfoot sightings seem to have faded in the background, with new research proving that these sightings are created by a figment of our imaginations. Yet, there are many that still report these sightings, and are convinced that they have seen something extraterrestrial. There are questions that are posed in this case: why do certain people truly believe they have seen something that is dis-proven? Where do these ideas and belief systems come from? Psychologists are trying to answer these exact questions. Creating a visualization that analyzes this unique group of people fills an important task for better interpreting this data. This will help psychologists better understand the inner workings of individuals who believe in UFO and/or Bigfoot sightings. Since the data is quite unreliable, a clean and organized visualization tool is necessary to help better understand this phenomena. With a focus on location, we are able to dive into the minds of believers, and help recognize if there is a correlation that may reveal an interesting pattern between location and belief systems.

1 INTRODUCTION

Voluntary reports of UFO and Bigfoot sightings can be submitted and found online for many to be seen. Examining these reports may lead to new insights. Our first thoughts were to look at the psyche of the people submitting these findings and what they have in common. To determine this, we will look where the sighting took place, the time of day, or the report given along with their sighting. Our group will analyze these factors in our first home page visualizations and look for correlations in belief systems of these phenomena. We will also ask questions on finding any correlation between physical and environmental attributes on a specific group of sightings. We intend to answer these in our other two smaller visualizations. Looking at these factors like the temperature during the sighting, shape of the seen UFO, or even looking at the phase of the moon during the Bigfoot sighting, we may end up finding interesting results that are statistically significant. Through our 3 main visualizations with our two data sets, we hope to find meaningful results on all of our desired questions.

2 RELATED WORK

Related work that we have looked at prior to making our visualization has served as a point of reference and inspiration for our project. First, the visualization tool UFO Tracker created by Nguyen, Pham, and Dang in *UFO Tracker: Visualizing UFO Sightings* [5] is designed to analyze UFO sightings reported to the National UFO Reporting Center. The tool uses data mining and visualization techniques to provide users with a high-level view of where different types of sightings occur, whether sightings are increasing or decreasing over time, and connections between different events. The usefulness of the tool is evaluated through a case study, and the authors believe that the approach has wider applications in other research domains, such as in analyzing large sets of text data obtained from social media. This tool and paper is useful to us because it is made for a very similar context to our use case, which can inform our work. This paper only focuses specifically on UFOs, but it helps us see a way that sightings have been mapped. One specific feature the tool uses that we think can be useful to us is brushing and linking, as well as being able to zoom in to the map to see more specific features.

The second related work also had an influence on our visualizations. Similar to our ideas, Shuming Bao, Changzhen Wang, and Miao Shui in *Spatial study of religion with spatial religion explorer* [1] created visualization to represent various locations of religious groups within China. In these visualizations, specifically Christianity was targeted, with the purpose of helping the general public better understand the religion in China. This tool uses religious information for "spatial data analysis". Within one specific figure, you are able to press on an area of the map, which would allow the user to access different types of graphs. One specifically focuses on allowing the user to analyze the spatial distribution of religions, and see if there is any correlation between the religions present in a given area. Similar to this project, our goal is to create visualization that will allow a user to see the correlations between sightings, and recognize if there are any dense areas of reported of both UFO and Bigfoot sightings.

3 USE CASE

Suppose a psychologist or sociologist was looking to study people who believe in certain phenomenons. Data on this subject is extensive and it can be difficult to observe patterns from the data's raw form. This visualization tool would reveal trends in phenomenon sightings that cannot be easily seen with just raw data.

Furthermore, in theory, UFO sightings and Bigfoot sightings are relatively disjoint events– one is in the sky and the other on ground. Comparing the geographical locations of where the sightings of these two phenomena take place could be telling. Environments and cultures in certain areas often have a large impact on a person's belief systems. Given visual data on the locations of either type of sighting, the user can further speculate about the location-based belief systems behind the reports. The user can investigate areas that have high amounts of both types of sightings, as well as areas that have little to no activity. By looking at this sort of data for the two phenomena together, the user can form further theories on why certain areas see higher amounts of phenomena sightings.

3.1 Domain Tasks

Our visualization will support two main domain tasks. First, identifying patterns based on the location of Bigfoot and UFO sightings. By identifying patterns in location, the user can see if there are certain areas that report both UFO and Bigfoot, which could reveal information about the types of people or landscape of the region. It would be interesting to see what regions see both types of sightings to inspire future research into why this might be happening. We plan to show this on our first visualization of a map of the United States with all of the points shown on the map.

Second, exploring specific UFO and Bigfoot sightings based on specific attributes of each type of sighting. Many Bigfoot sightings

*e-mail: lanman.k@northeastern.edu
†e-mail: wang.ziyu1@northeastern.edu
‡e-mail:deieso.j@northeastern.edu
§e-mail:hobson.br@northeastern.edu
¶e-mail: kent.sa@northeastern.edu

contain environmental attributes such as temperature, moon phase, and seasons. Whereas the UFO sightings have more physical attributes like shape and duration. Looking more in depth on both types of sightings may find a correlation of this phenomenon. This will be shown on a secondary visualization on a new tab or as a sub-visualization through brushing/linking.

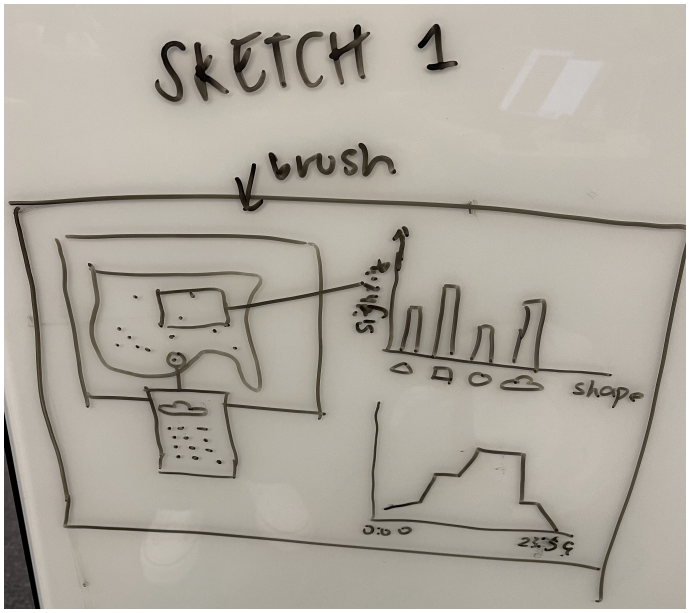
4 Data

Our Bigfoot data came from the Bigfoot Field Researchers Organization (BFRO), and was accessed via data.world.com [6]. Specifically, we accessed the bfro_locations.csv and bfro_reports_geocoded.csv files. Together, these files contain around 5,000 unique reports on Bigfoot sightings as well as times of the events, point location data, and weather data scraped using the Dark Sky API. The original locations data and geocoded data contained 4,250 rows and 6 columns and 5,021 rows and 29 columns, respectively. Each report is identified by a unique number, which we used to merge the two data sets. We then cleaned the data to ignore reports missing data on latitude/longitude or date and dropped columns unnecessary to our analysis such as title and geohash. We then further abstracted the data by filtering for reports made after 2010. Once this was done, we were left with 964 rows and 16 columns of data with each row representing a unique report and the attributes surrounding that event.

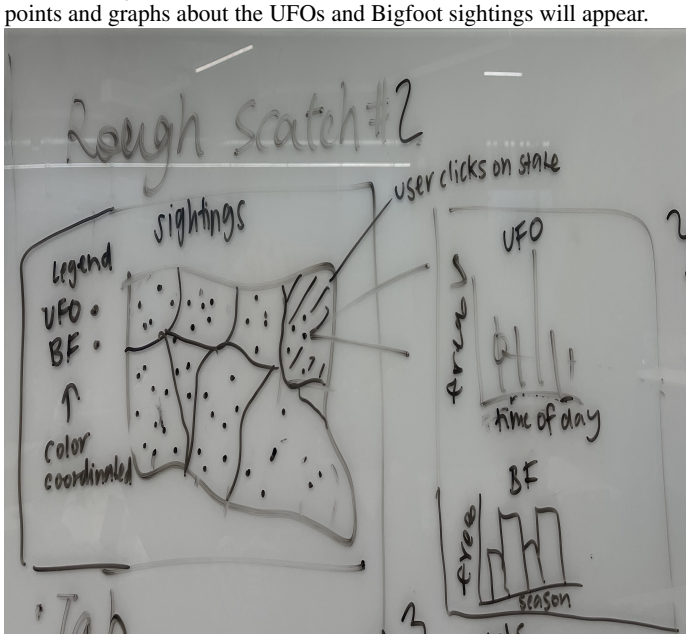
The UFO data originated from the National UFO Reporting Center, then was accessed via data.world.com [7]. It had 14 columns and 141,285 rows before cleaning. Each row in this file is a recorded UFO sighting and contains information such as coordinates, summary, and shape of the object seen. This data set is quite large, and it needed to be filtered to remove empty values. Through cleaning, we removed all the data that had NaN as a value, to make the data more usable. The sighting data goes back to 1970 and stops at 2022, thus, to make it consistent with the Bigfoot sighting data, we trimmed it down to just between 2010 and 2022. Similarly, since the data for the Bigfoot sighting is restricted to the USA area, we also limited the country for the UFO sightings in the US. At the current status, we still have 59170 rows. To match the size of the Bigfoot data, we randomly selected the 1000 rows from the table to visualize. Our method selects rows at random, with each row having an equal probability of being selected. By selecting a subset of the data in this way, we can create a smaller dataset that is still representative of the larger dataset, allowing us to visualize trends and patterns in the data in a more manageable way to the user. For columns, we dropped the report time, report link, summary, and country for a final width of seven columns.

There are biases that come into consideration when accessing this data. Without any concrete evidence of the existence of Bigfoot and UFOs, we are unsure if these are sightings in which people saw something, or if people are just creating a fictional story. However, this data is self-reported as people submit their sightings, so there are limited privacy concerns unless a reporter violates another person's privacy when reporting. The Bigfoot self-reporting form on their website does state that all legitimate reports will be looked at by someone in the company, so this means that there is a second check after someone submits a report. The UFO website also has a self-reporting form, but it also explains to users that certain planets and stars should not be mistaken for UFOs. Same with Bigfoot, there is someone within the company who reads each submission and checks for its validity.

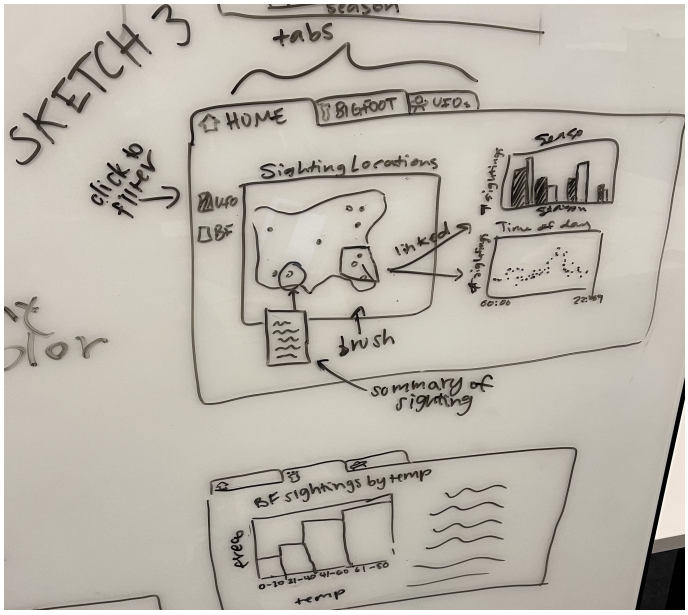
5 Design Process



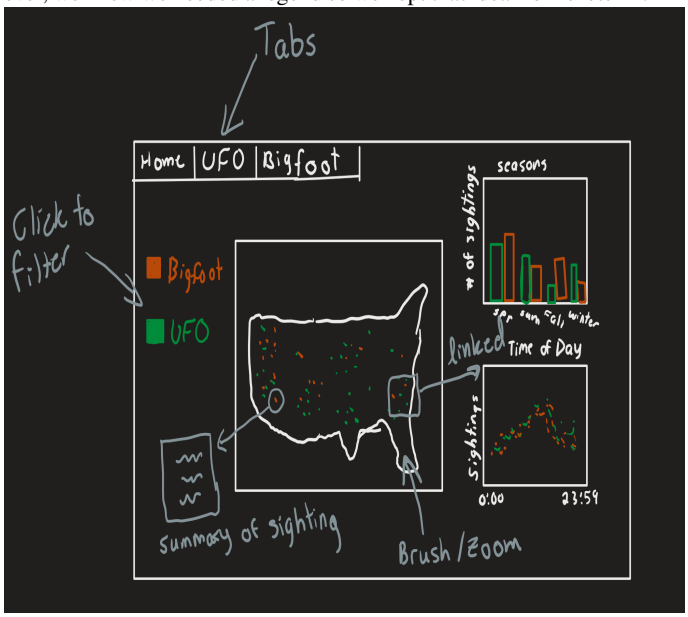
This first rough sketch shows the main ideas we had for our visualization. We want to lay out the data geographically so it is easy for the user to identify where the sighting occurred, and we want the user to be able to click on a specific sighting to read more about it. The brushing in this first sketch shows that a user can brush over points and graphs about the UFOs and Bigfoot sightings will appear.



The second sketch expanded a bit more on sketch one. Instead of using brushing to highlight points, we thought the user could click on a specific state and then those points will show in the visualizations that appear. We also added a legend that would show the difference between the points that are UFO versus those that are Bigfoot sightings.



For sketch three, we decided that the user needs to be able to filter between UFO and Bigfoot. We added tabs for the user to click on to see specific sighting information. We also decided to go back to our ideas for sketch one, because implementing state lines would require us to gather data on longitude and latitude of states. However, we knew we needed a legend so we kept that idea from sketch 2.



The above image shows a detailed look into what our final website's homepage will look like. This final image took ideas from all of our sketches. It will have tabs at the top to switch from the homepage visuals to either the detailed UFO or Bigfoot visuals. On our homepage we will have our main graph of the longitude and latitude markings of the phenomena sightings over a map of the US. The marks on this visualization are points, with channels of both horizontal and vertical positions and the color of whether it is a UFO or Bigfoot sighting [3]. When clicking on the filters you may show Bigfoot sightings, UFO sightings or both. When clicking on a point it will display a summary of the sighting. You can zoom in and out and bush different points on the map. Those points brushed will be linked to the two graphs on the right displaying the seasons these sightings occurred on and the time of day. The season's bar chart will have a mark of lines, with channels of both horizontal and vertical positions and colors to determine UFO or Bigfoot. The time of day scatter plot will have marks of points and channels of both horizontal and vertical position and color to determine a UFO or Bigfoot sighting.

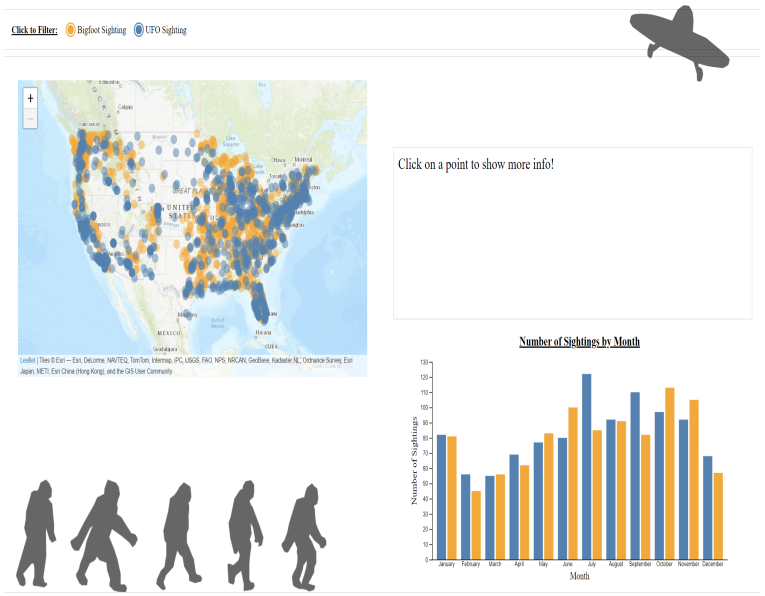
The project remained largely consistent with our initial planning and we were able to accomplish the majority of our established objectives. Specifically, we streamlined the visualization component of our project from a three-page website encompassing location, description, and statistical data with a separate tab for UFO and Bigfoot statistics to a single grouped bar chart that permits filtering through zooming and our legend. This filter enables viewers to manipulate data both on the bar chart and the map, based on different events and the time of their occurrence.

The initial edition underwent utility testing, which provided us with valuable feedback that ultimately culminated in the finalization of our visualization. Among the critiques, test users commented on the overlapping of data on the map, as well as its distribution and overall geographical prominence. These findings suggest that the visual encoding for location proved highly effective. Similarly, the visualization of the month distribution of the sightings was processed and comprehended by viewers with ease. These observations suggest that the non-interactive visualization adequately served its intended purpose.

However, when it comes to the interactive component of the visualization, our viewers intuitively clicked on data points to obtain specific sighting details, as well as the details of the sighting month in the bar chart. Yet, when viewers interacted with the filter, they were initially unsure if it was interactive, and when they realized it was, they intuitively removed the visualization they were asked to keep, leading to some confusion. Additionally, they mentioned that the lack of a response from the filter added to their confusion. Some suggestions were made concerning the layout of the webpage.

After the testing, we decided to implement changes such as modifying the cursor form as a hint when the cursor is above the filter to remind viewers that it is interactive. We also changed the form of the filter after it was clicked to indicate that it was in use. We additionally changed the layout of the page to present the two visualizations side by side to create a stronger connection between them.

6 Final Design



Our final tool includes two main visualizations. The first being the Bigfoot and UFO map of the United States [2]. This map includes

every point of a sighting between the two datasets. The user is able to zoom in and out, brush to move on the map, and click on a point. When a point has been clicked on, a tooltip will show to display the certain sighting's type, date, state, and city/county. The point clicked on, will also display a self-written description in the scroll-box under the graph. To the right of the graph is a grouped bar chart displaying the number of UFO and Bigfoot sightings per month. When hovering over a bar a tooltip will show the type of sighting and the number of sightings. The map and barchart are linked together. When zooming in on the map, the bar chart will adjust to the amount of points being shown in the map's field of view [8]. Lastly, our visualizations include a global filter which includes a button for UFO sightings and another for Bigfoot sightings. When clicking on a button, it will filter that sighting out of both visualizations.

To solve our domain problem using our visualization, you may want to check a certain region/state. Looking at the similarities of sightings in the south like Texas and Florida might give more information about the people there rather than looking at the whole United States as a whole. With the zoom and linking features of our visualization, it is easy to do this. A user could filter so there are only Bigfoot sightings in Florida. This could show a certain region, like central Florida, that has more sightings than places, like southern Florida. This would have the bar chart also automatically update to see the certain months these sightings occurred in to see if the month or season has correlation.

7 Discussion

Our original domain problem was intended to help users identify patterns based on locations of UFO and Bigfoot sightings. These two phenomena are not typically associated together, and we hoped that our visualization reveals patterns allowing the user to gather insights about why certain regions see high or low levels of sightings. We also aimed to allow users to explore specific sightings based on different attributes.

After implementing our final visualization, we addressed the first part of the problem we hoped to solve, by visualizing the sightings on a map. One limitation of our final visualization, however, is we could only show sightings in the contiguous United States due to limitations with our map. This means that we were not able to show the entire United States, Alaska and Hawaii were not able to be implemented. The user is therefore not able to get the full picture of Bigfoot and UFO sightings in the entire United States. The second part of our domain problem was addressed, by allowing the user to click on a specific sighting and read the user description, and information like the date, state, or county.

In the future, if we were to make improvements on our visualization, we would add additional visualizations that demonstrate patterns with the specific type of sighting, adding a moon phase chart for the Bigfoot sightings, for example. Adding additional visualizations like these would not only allow users to deeper explore one type of sighting, but also glean the likelihood of weather patterns creating the illusion of the phenomenon. We also would have liked our visualization to update in real time based on new sightings that are reported. This would be a good change to make so that the user can get the most recent information about new sightings.

8 Conclusion

When starting this project, we identified a couple main goals for our visualization. The first was to display the two types of sightings based on their geographical location. We implemented a map that shows the sightings and colors them based on the type. Our second goal was to display specific information about each sighting. We implemented an on click event, so when a point is clicked, it shows the full report summary for each sighting. Another goal we had was to find correlations between the sightings. We made a linked bar chart that shows amounts of sightings per month, based on map location.

The data for this visualization was from real people who reported sightings to the Bigfoot Field Researchers Organization and the National UFO Reporting Center. We cleaned this data and aggregated them into central CSV files, and used random sampling that was representative of the original data set in order to display a number of points that is easily digestible for the user [4].

The final implementation of this visualization allows the user to explore patterns in the data based on location and timing. The visualization also has the capability for the user to explore specific sightings, and read some of the reports to answer questions on why people believe in these phenomena. Through the visualization, we can study the language used to discuss these phenomenon and hopefully uncover the motivation and psychology behind these reports.

For the project we all contributed in different areas. Kate mostly worked on the leaflet map and linking the points between the two visualizations. Jackie also worked on the leaflet map and legend/filter for the visualizations. Ziyu focused on the bar chart and creating the dataset for the UFO data. Sarah created the two tooltips for the leaflet map and the bar chart. Brady worked on the on click function of creating the description into the scrollbox. We all worked together during class, on our own times, and with scheduled group meetings.

Acknowledgments

The authors wish to thank Professor Ab Mosca for all their advice and support on this project, as well as the rest of the DS4200 class who assisted in testing the visualization.

References

[1] S. Bao, C. Wang, and M. Shui. Spacial study of religion with spatial religion explorer. *2014 22nd International Conference on Geoinformatics*, 2014. doi: 10.1109/GEOINFORMATICS.2014.6950848.

[2] M. Bostock. U.s. map, 2020.

[3] Dixon and Moe. Html color codes, 2015.

[4] J. M. Importing data from multiple csv files in d3, 2014.

[5] V. T. Nguyen, V. Pham, and T. Dang. Ufo tracker: Visualizing ufo sightings. *IEEE International Conference on Big Data*, 2018. doi: 10.1109/BigData.2018.8622418

[6] T. Renner. Bigfoot sightings. dataworld.com originally from BFRO, 2017. Last updated in 2019.

[7] T. Renner. Ufo sightings dataset. dataworld.com originally from NUFORC, 2017. Last updated in 2019.

[8] T. S. Get a list of markers/layers within current map bounds in leaflet, 2014.

A Data Abstraction

This appendix includes the data abstraction for the two individual final data sets on Bigfoot and UFO sightings. Both data sets are formatted as table with items and attributes. For both the Bigfoot and the UFO data, rows represent one unique sighting of the respective phenomenon. The attribute types for each data set are laid out below.

Table 1: Bigfoot Data Abstraction

Attribute	Type	Sub-types (if applicable)	Description
observed	Categorical		Description of event from observer
location_details	Categorical		Location description from observer
county	Categorical		County event took place in
state	Categorical		State event took place in
season	Categorical		Season event took place in
latitude	Ordered	Quantitative, diverging	Latitude point of event
longitude	Ordered	Quantitative, diverging	Longitude point of event
date	Ordered	Ordinal, sequential	Date of event
number	Categorical		Unique identifier of report
classification	Categorical		Classification of reliability of report
temperature_high	Ordered	Quantitative, diverging	High temperature from day of event
temperature_mid	Ordered	Quantitative, diverging	Mid temperature from day of event
temperature_low	Ordered	Quantitative, diverging	Low temperature from day of event
moon_phase	Ordered	Ordinal, cyclic	Moon phase from day of event
summary	Categorical		Summary of weather from day of event

Table 2: UFO Data Abstraction

Attribute	Type	Sub-types (if applicable)	Description
city	Categorical		City of the sighting
state	Categorical		State of the sighting
date_time	Ordered	Ordinal, sequential	Time of the sighting
shape	Categorical		The shape of the UFO
duration	Ordered	Ordinal, sequential	Compact all the info of the sighting
text	Categorical		Witness description
city_latitude	Ordered	Quantitative, diverging	Latitude of the sighting
city_longitude	Ordered	Quantitative, diverging	Longitude of the sighting