

Visualizing the U.S. National Park Service's Inventory and Monitoring Program Data

| | |
|---|----|
| Abstract | 2 |
| 1. Introduction | 2 |
| 2. Previous Work | 3 |
| 3. Design Process | 6 |
| a. Sketches | 6 |
| b. Intended Users and Tasks | 9 |
| c. Interactive Prototypes | 10 |
| d. Usability Tests 1 | 12 |
| e. Expert Interviews | 12 |
| f. Redesign | 14 |
| g. Usability Tests 2 | 15 |
| 4. Final Visualization | 16 |
| a. Surrounding Webpage | 16 |
| b. Dashboard 1: Find Vegetation Qualities by Plot Properties | 17 |
| c. Dashboard 2: Explore Table of Vegetation Plots | 24 |
| 5. Evaluation & Discussion | 26 |
| Assessment | 26 |
| Interaction Techniques | 26 |
| Challenges | 28 |
| 6. Future Work | 29 |
| 7. Acknowledgments | 30 |
| 8. References | 31 |
| 9. Appendices | 33 |
| Appendix A: Personas | 33 |
| Appendix B: Interview Guide Questions | 36 |
| Appendix C: Questionnaire | 37 |
| Appendix D: Data Sets | 39 |

Abstract

The United States National Park Service (NPS) collects and analyzes extensive park resource data through their Inventory and Monitoring Program (I&M). The program, established in 1992 [1], is intended to document and monitor resources in the national parks and to play a key role in managing park ecosystems. The program aims to integrate the natural resource inventories into NPS planning, management, and decision making. The data collected will serve as a baseline for future resource management and may provide early warning signs of adverse events [1].

Because these data are intended to play a large role in education, communication, and collaboration, we saw an opportunity to enhance these efforts by creating an interactive and engaging visualization of the I&M data. While scientists and park staff may be able to communicate through complex statistical analyses of the data, an interactive visualization of the same data may improve communication with policy-makers, funding agencies, and the tax-paying community.

We designed our visualization to allow park rangers, scientists, teachers, and students to explore the National Parks' resources. We chose to include several types of data from the complete I&M datasets with a focus on vegetation, ground composition, geospatial, and topography variables.

In this paper, we present our design and research process as we planned, sketched, created, tested, and redesigned our visualization of the National Park Service's Inventory and Monitoring Program data.

1. Introduction

The United States National Parks are some of our country's most valuable assets. Not only are the parks beautiful places to visit, the shorelines, forests, trails, and education centers in the parks provide a unique experience of hands-on learning and community outreach in all regions of the country. While we rely on the national parks for their beauty and the rich experiences they provide during our visits, the parks are the birthplaces and homes of an abundance of wildlife and complex ecological systems. It is our responsibility to protect them.

In order to protect the parks and maintain their ecosystems, it is important to monitor animal, water, air, soil, and vegetation data to identify trends and know when and how to act upon those trends to mitigate or prevent situations that threaten the parks and their resources. Park Rangers, biologists, chemists, and other scientists are working tirelessly to collect the necessary data. It is a large, multi-faceted effort.

We focused our visualization on soil, vegetation, and the relationship between the two because the state and composition of the soil and vegetation are crucial for healthy ecosystems and resource preservation. Unhealthy soil can be an early indicator of larger problems for vegetation and wildlife, or even flood and fire risk. [The structure and quality of soil](#) affects the movement of water, plant roots, and resistance to erosion [2]. [Healthy soil](#) controls water from rain and snow; healthy soil sustains plants and animals, filters potential pollutants, cycles nutrients such as carbon and nitrogen, and provides physical stability and support [3]. Combining soil data with vegetation data further supports our intended users' ability to extract insights from the collected data via our visualization. We learned during our user research that manipulating and

exploring aspects of soil and vegetation data in a single visualization enables scientists to identify, test, verify, and predict potential environmental threats.

To enrich the soil data, we also decided to include vegetation data in order to allow our intended users to make deeper and more meaningful comparisons. The vegetation data includes plant association, leaf type, leaf phenology, and physiognomic class. Each of these variables is useful in identifying a plant, which helps users of our visualization see correlations between soil, topography, and vegetation.

2. Previous Work

The National Park Service runs a Soil Program to secure information that is needed for managing soil sustainability. The Soil Program ensures that water quality, wetlands, vegetation, and wildlife habitats in the national parks are protected. Collecting and assessing highly detailed information about soil composition is critical for protecting and managing the Parks' natural resources and for predicting the behavior of soil. Soil data is collected from numerous plots in national parks. The number of plots sampled per park varies and data points relevant to soil can easily exceed 22,000 for just one park.

The majority of the data we used for this project came from the National Park Service's [Inventory & Monitoring Program](#) [4]. There have been many small visualizations created using this National Park Data. Several that we found used GIS data (Geographic Information System) to create [basic static visualizations](#) [5].

The Center for Remote Sensing and Mapping Science (CRMS) in the Department of Geography at University of Georgia used the data to [assess the vegetation models and related fire risks](#) [6]. The paper uses vegetation, terrain, and various other historical data points to summarize vegetation patterns and distribution in the Great Smoky Mountains National Park. While we did not go into the detail covered in the paper, we do provide a higher-level, interactive way to view patterns and distributions across many national parks.

Finding interactive visualizations related to our project proved to be difficult. The main source of related visualizations originated from the 1979 [Interagency Monitoring of Protected Visual Environments](#) (IMPROVE) [7]. We stumbled upon IMPROVE by exploring the National Park Service's website. The goals of the visualizations are to document long-term air quality trends. Unfortunately, the visualizations are very rudimentary and provide little opportunities for interaction. For example, the [Spatial and Seasonal Patterns](#) graphic [8] allows users to select between different isopleth map images via a drop-down list, as you can see in Figure 1.

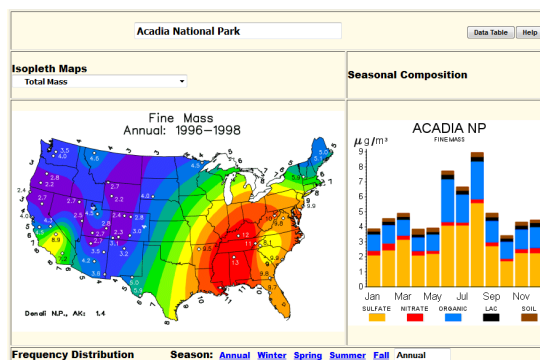


Figure 1

After selecting a map type from the drop-down list, the map image simply changes to another image. Because the map is only a .GIF, there is no intuitive way to explore the data through linking or brushing. The interactivity with this graphic is minimal and not at all visually pleasing. It does, however, provide good data for analysis, and communicates trends clearly. We do believe that improving the media through which the information is communicated can be facilitated by updating its features, graphics, and interactive features. Doing so would encourage more people to explore the data and make discoveries on their own.

IMPROVE also has a [Metadata Browser](#) [9], but it is not designed to work on “current web browsers,” as seen in Figure 2. As users, we found this warning to threaten the perceived validity of the data. Thankfully, we were still able to manipulate the data and see some interesting things via brushing and filtering.

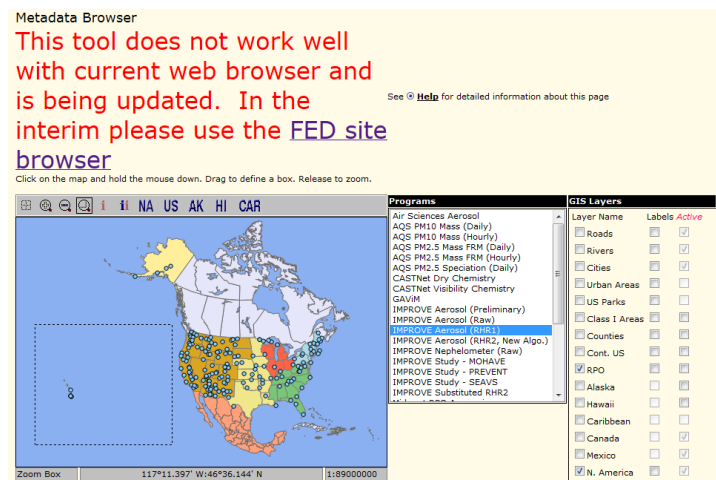


Figure 2

This visualization provides more interactive features than the Spatial and Seasonal Patterns graphics, but we believed we could still improve the experience by providing seamless interactions, expressive animations, and relevant, timely feedback to help guide the user.

While IMPROVE provides us with examples upon which we could improve, the [Shaded Relief](#) [10] and [Relief Shading](#) [11] websites illustrates techniques, design rules, and examples of how to represent different types of topography on maps. These websites also provided invaluable resources, like the incredibly complete and verbose Natural Earth [public domain map dataset](#) [12] aggregated and validated by the [North American Cartographic Information Society](#) [13]. The Natural Earth dataset provides data in three levels of detail: small (1:110m), medium (1:50m), and large (1:10m) scales, seen in Figure 3.

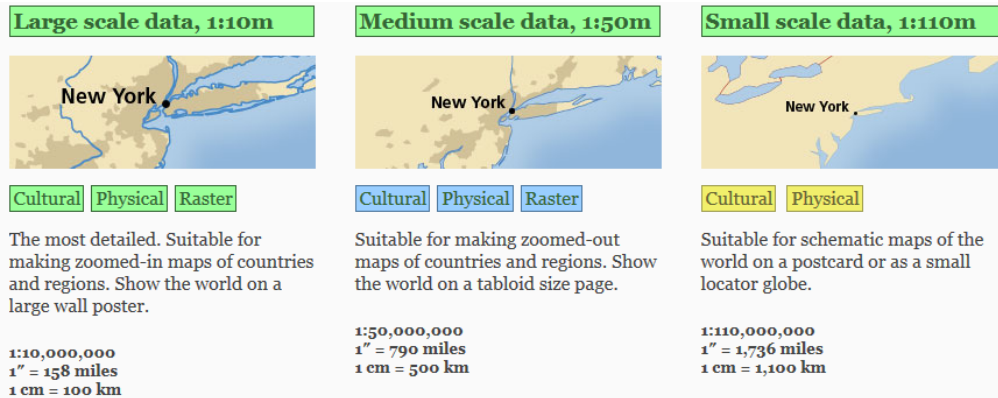


Figure 3

The Natural Earth dataset is available as vector themes in SHP/GeoDB and SQLite. Figuring out how to take advantage of this data proved to be difficult, as we were not planning to use these file formats. Through more experimentation, though, we were able to utilize some aspects of the [vector themes](#) [14] in order to better understand the context of our problem space.

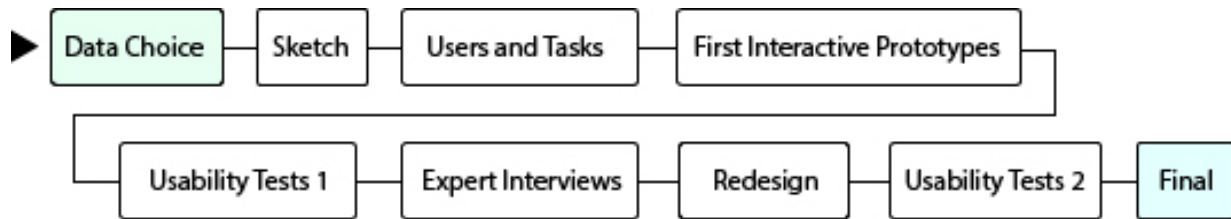
As we looked for further examples of previous work in this area, we found The United States Department of Agriculture's [Geospatial Data Gateway](#) [15]. This gateway is a portal through which the public can order environmental and natural resources data for a monetary fee. Users can select an area of interest, select which data they want to overlay on the map, and customize its format. Unfortunately, this data is delivered via FTP download or CD/DVD by USPS. In order for users to see the data, they are required to "place a data order" and pay a fee. We selected the area of the Olympic National Park in Washington state and a few pieces of data we wanted overlayed on the map. Our fee was \$50 and we would have had to wait a week in order to receive the CD.

After exploring the Geospatial Data Gateway, we found a similar data gateway that focuses on soil. The [Web Soil Survey](#) [16] is similar to the Geospatial Data Gateway in that users can select areas on a map and what data they want to see about that area on a map that they have to buy. However, we found a useful [database of tabular and spatial soil data](#) [17], which we referred to when cleaning up the datasets we used in our visualization.

The most impressive visualization that is related to our project, [Air Atlas](#) [18], is a series of web maps that provide air quality statistics via an interactive medium. These maps provide users the ability to choose several different types of overlays to modify the main display. To choose the overlays, users are provided simple, straight-forward checkboxes to brush and filter the display. We used Air Atlas to help guide our design decisions in order to present soil quality data in a similar fashion.

We were able to access a lot of data from various web services and government websites that were collected by extensive surveys over many years. However, many of the visualizations and graphs are clunky to use, static, or behind a paywall. Our goal was to improve the interactivity and presentation of the data, allowing our target users to consume and act upon it in an engaging, responsive, and accurate way.

3. Design Process

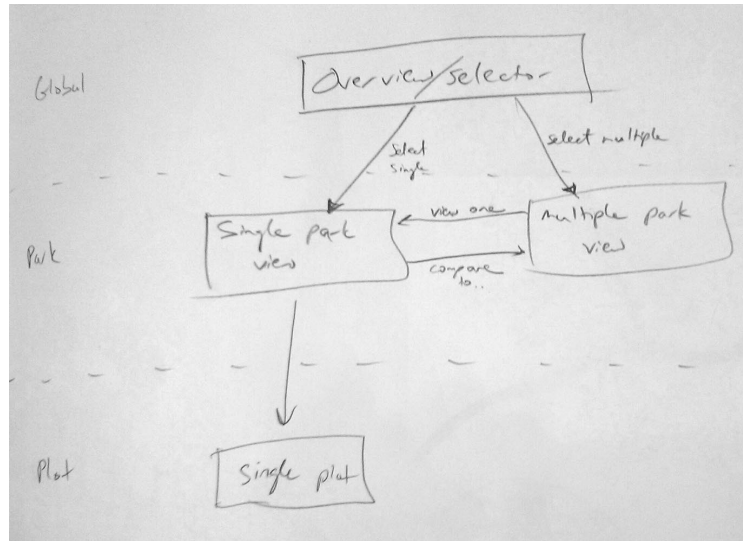


The diagram above depicts the path our design took from data choice through the final product. In this section, we will detail each step from sketching onward.

3a. Sketches

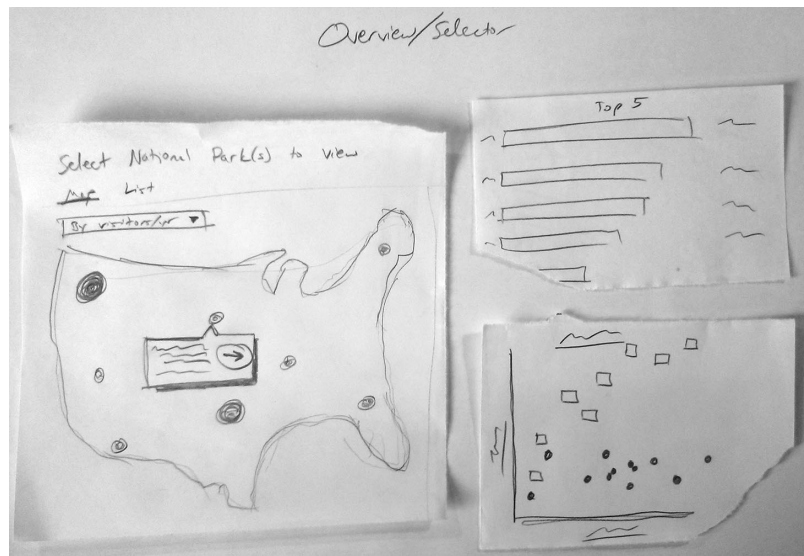
Because our datasets are so robust, our initial design phase consisted of creating 20 to 25 iterations of different interface designs. To record these ideas, and to facilitate discussion and debate within the team, we began with rough sketches. These rough sketches allowed us to test various ways to organize the information. The low-fidelity of the sketches allowed us to try out ideas by quickly erasing, cutting, tearing, and taping the paper. It was a beneficial, iterative, and collaborative exercise that positioned us to design a stronger product. A few examples of our sketches are below.

Overall Hierarchy



Map of information architecture/hierarchy (Summary/overview-level, park-level, and plot-level)

Dashboard View



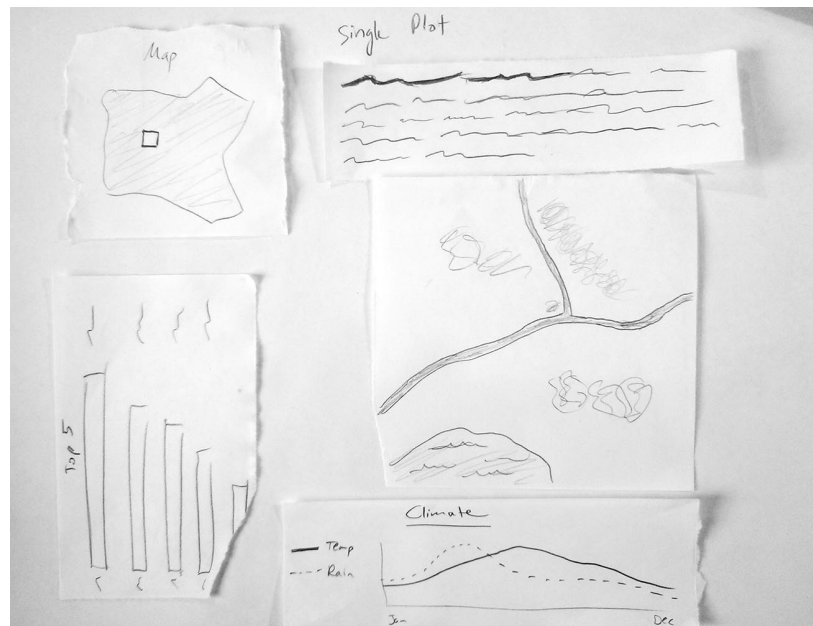
Mock-up of Overview-level view, showing a summary of all the national parks in our data set

Park Level View



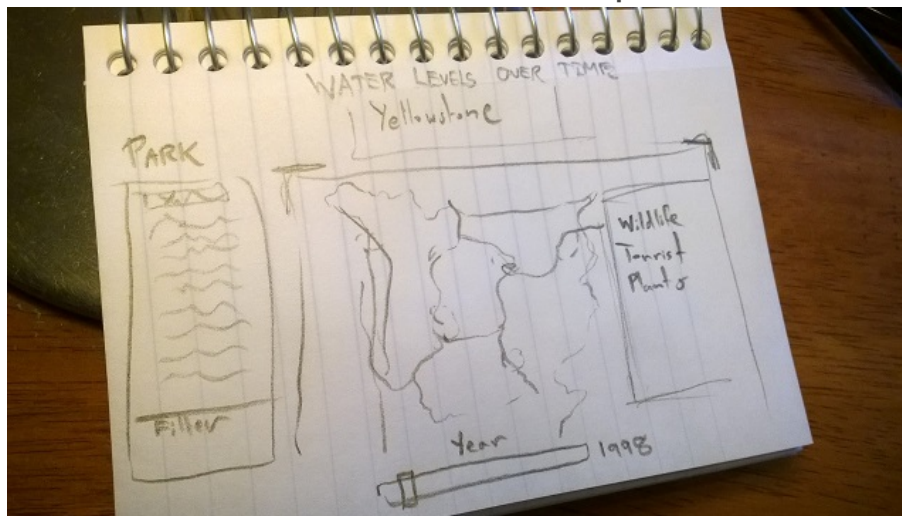
Mock-up of Park-level view, showing summary and details for a single park

Plot Level View



Mock-up of Plot-level view, showing details for a single plot

Detail View Variation with Map



Showing a possible detail view of one park. Map in middle, various filters, navigation on the left. This example shows analysis with a bar that slides between different years.

After sketching, we were able to think more clearly about how users would interact with our product. We struggled, however, with determining how we might improve the poor experiences offered by the visualizations we found when researching related work. As our brainstorming continued, we began interacting with our paper prototypes using gestures such as swiping and noticed that we were making assumptions about our users. We decided to put on our user-centered design hats and pursued proper user research.

3b. Intended Users and Tasks

We created personas to help guide our design process and focus on the needs and goals of our intended audience. We referred to online job postings for Park Rangers at national parks as well as various Wikipedia articles that described the characteristics of park rangers, soil scientists, and educators. We compiled our research into three personas. The following personas are abbreviated; the full personas are available in [Appendix A](#).

Primary Persona

Dr. Sally Smith is a park ranger and soil scientist under the National Park Service. Above all else, the goal of park rangers in United States national parks is to protect and preserve the park's resources and guests, and Sally's contribution to the protection of our national parks is her expertise in soil composition and integrity. She spends most of her time in Yosemite National Park, though her role has her visiting other national parks across the nation when there are concerns about a park's soil. Sally uses her expertise in soil to identify potential correlations with vegetation attributes and soil types to help predict and mitigate any negative environmental events.

Goals

- Get an overview of the soil composition across multiple parks and within different regions of the same park
- Understand relationships between multiple variables and their effects on soil in order to recognize trends and make predictions
- Quickly find the specific data points she needs, such as looking up values for plots nearby a new building project
- Correlate soil and vegetation data, such as comparing specific vegetation types with soil texture types.

Secondary Persona

Ed Westgate is a park ranger who has been working in Arches National Park for the past 5 years. His passion has always been education and getting people excited about the country's natural resources. He took a position in the National Park Service to help spread the message of conservation and protecting the parks' ecology. Because soil and vegetation are, in his opinion, the foundation of an environmental system, Ed wants to explore in which types of soil do certain vegetation attributes occur so he can learn more about their interesting and varied relationships.

As a ranger, Ed spends a lot of his time giving educational and interpretive programs to the park's guests, and his programs range from talking about the food chain, showing visitors evidence of the park's changing geology, and leading cultural re-enactments. In his downtime, Ed spends a lot of his time learning about new fields so that he can be ready to answer his enthusiastic guests' wide range of questions.

Goals

- Skim through the data and look for anything interesting or exciting that he can share with his guests
- Understand the soil make-up of his park to be able to respond to visitor questions
- Make comparisons between different parks to be able to compare and contrast Arches with others across the country

Tertiary Persona

Jim Murray is a high school science teacher who has recently switched into teaching Earth Sciences and Ecology courses. His students usually seem disinterested, especially during boring segments like the importance of healthy soil, but he knows that if he can show them something “cool,” they perk up and get excited.

Goals

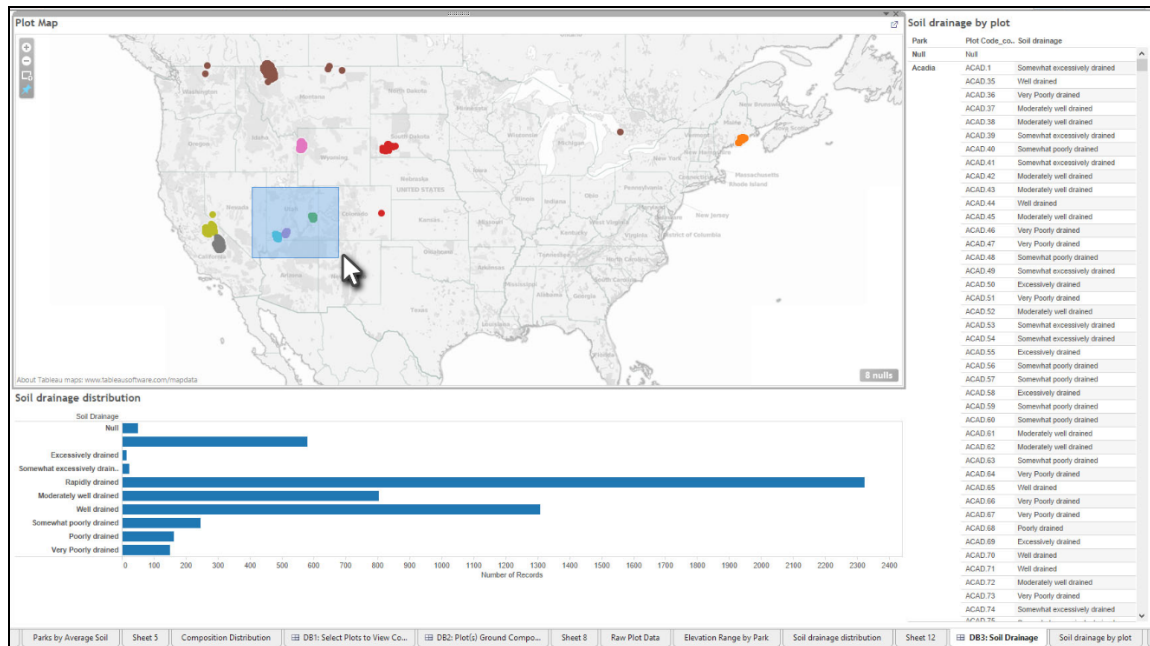
- Teach his high school students about the geologic diversity within their country
- Make the course material more exciting through the use of visuals and exploration

3c. Basic Interactive Prototypes

After creating rough sketches and defining our target users and their goals, we started importing our dataset into Tableau. We did this as a team; all of us used one machine in order to continue the collaborative nature of our design process. At first, we started dragging and dropping parameters onto the X and Y axes to see what sort of view we would get. The immediate and natural encoding that Tableau applied to the data made it come alive. Finally, we were able to explore what interactions we wanted to include in our visualization and how to implement those interactions. For example, one of our most challenging design questions was: how do we design a visualization that allows users to quickly and intuitively navigate from a high-level view to a specific, detailed view?

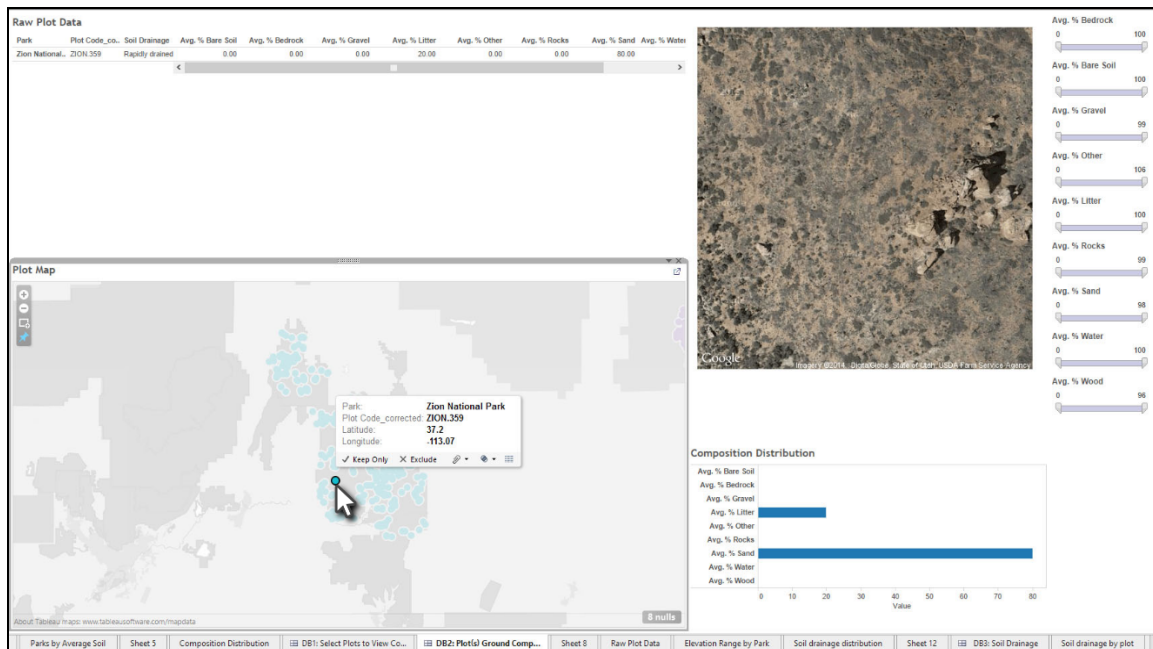
We were able to combine data together to make interesting views quickly. Some of the views were okay, some were good, but it was the great visualizations that had the most significant effect on us that made it into our first prototypes. Tableau helped us narrow down our scope to a consumable amount of information, but we still found ourselves figuring out how all of these visuals would live in a single dashboard.

We divided the dashboard into a grid based on our preliminary sketches. Using these, we were able to implement features such as brushing and filtering while also considering our users. Since we have experts and non-experts as our users, we had to take into account varying perspectives and goals. To do so, we decided that the default view should be a geospatial representation of the data. We thought that a colorful map would draw in users and help keep their attention. We wanted this view to be inviting, yet not so shallow as to threaten the visualization’s validity among our primary user group. To accommodate the needs of both groups, we made the brushing and filtering that allows users to use the visualization as an analytical tool very apparent. With these interaction opportunities in the foreground, users who want to dig deeper into the visualization to answer specific questions or identify patterns know which UI elements allow them to accomplish their goals. With this in mind, we created a prototype of our visualization that enabled awareness, exploration, and analysis, thus satisfying the needs of all three of our user groups.



Selecting some plots to filter the rest of the visualization.

To further support both casual and expert users, we added a feature that allowed users to see a satellite view of a specific plot. We used the [Google Static Maps API](https://developers.google.com/maps/documentation/staticmaps/) by passing the plots' longitude and latitude information from our dataset to the service. When users would click on a plot, a frame next to the map would provide certain data about the plot, and an aerial view of the plot would render in another frame. We thought this was a nice touch, because for casual users it might be difficult to imagine what a temperate coniferous forest with a few *Magnolia grandiflora* trees might look like.

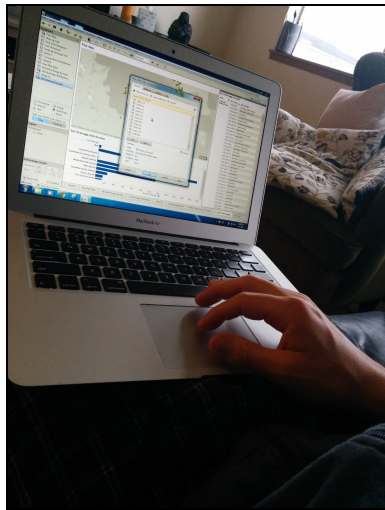


Selecting a specific plot while getting details via tooltip.

3d. Usability Testing Part 1

After we created the initial prototypes, we conducted usability testing to guide further improvement of our visualization. We did this to force ourselves to design with our users and their goals in mind, rather than designing based on opinions and assumptions.

Our initial round of usability studies included three non-expert users with limited knowledge of soil and vegetation. All 3 participants used our basic interactive prototype using the think-aloud protocol. With the goal of testing our visualization's information architecture and appropriateness of our supported interactions, we discovered that completing our tasks was relatively painless, but brushing and highlighting were more difficult than expected. Testing also revealed a need for more labeling throughout the visualization. We also determined from the testing that adding more labels would help our users understand the relationship between the visualization frames and the data being represented.



Quotes

"[The] picture I saw before could have been a cross section of earth so I didn't know that it was necessarily a map view. I don't know if you can customize [the view], but more labels would be nice."

"I found it difficult to select plots around [the map]. I would like to have different tools to select plots rather than just a simple square selector."

3e. Expert Interviews

As part of the feedback we received from our midterm presentation, our team was challenged to interview experts in soil science and geology to better understand our users' goals and use cases. We received generous suggestions and contacts from our peers and one of our professors. All of our interview participants and our friends who connected us with them are listed in the Acknowledgements section of this paper.

We interviewed three experts before we redesigned our final visualization. Each of these interviews lasted between 30 and 60 minutes. Our interview guides included questions about our participants' backgrounds, the importance of the data they use, and the application of that data in the real world. We also prompted the interview participants to think about how a visualization could improve their work or their field. This encouraged them to consider how they might like to interact with the data.

“People like park rangers could use it in their campfire talks: ‘where can we find plants that are most likely to be threatened by drying climates?’ ... It can raise awareness about dynamism and how it relates to environmental variables and how that might be used to better mitigate or plan for changes in the future.”

The following are samples of specific insights from the interviews and related design implications.

Key Insights and Design Implications

Insight 1. Soil data by itself is not very useful, but soil data combined with vegetation data is.

Design implication: Include vegetation in data visualization so that relationships between soil and vegetation can be identified and explored.

“The whole reason for ecologists to collect all of these data is to develop a classification that includes plants and other variables to make a description of a repeatable pattern of vegetation across the land. ... Geology can inform what vegetation will be found in a particular place.”

Insight 2. Scientists want to see the raw data.

Design implication: Provide a direct link to the raw dataset within the visualization.

Insight 3. Elevation, aspect, slope, and soil texture are most important data points. Being able to predict outcomes in plots for vegetation and wildlife depend on these baselines.

Design implication: Try to include this data throughout visualization.

Insight 4. A soil/vegetation visualization would be good for raising awareness, outreach efforts, and enhancing communication between different research and work areas.

Design implication: Used visual communication techniques learned throughout the quarter to best support our users and their goals. Supported both analysis and exploration with our visualization design.

Insight 5. Including a [soil texture triangle](#) can make the visualization be more useful to experts.

Design implication: As users brush and filter through our visualization, they are able to perform more in-depth analyses by using the soil texture triangle. The triangle is a qualitative classification tool used to classify soils based on their texture. It is widely used amongst our target expert users.

3f. Redesign

After interviewing our chosen experts, we made significant changes to our visualization. The image below depicts major changes at this stage:

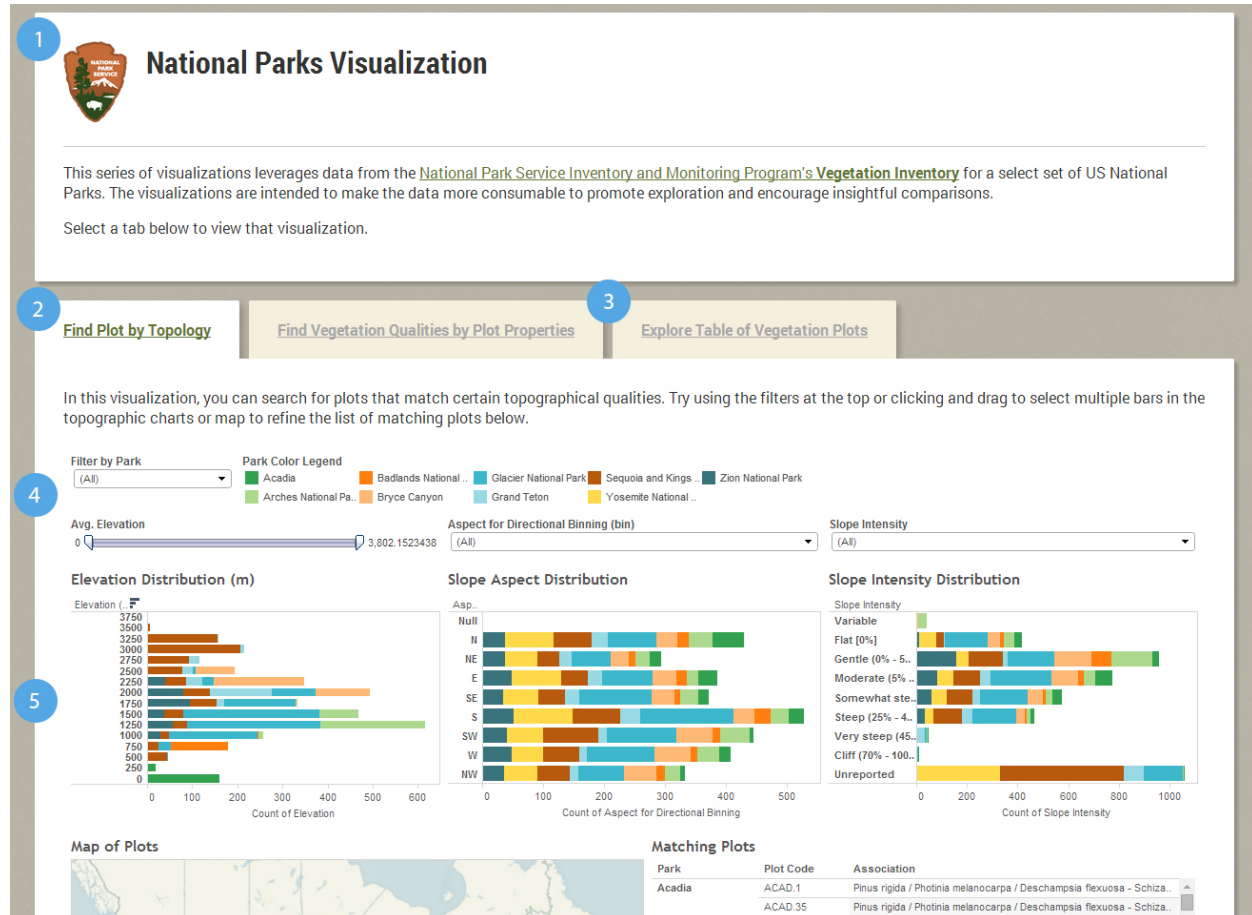


Figure 6

- 1 We tried including all of our information on a single dashboard, but the results were cumbersome. Instead, we decided to build a web frame for the data to house our Tableau dashboards.
- 2 We split up our dashboards with large tabs named after the task the visualization supported.
- 3 As suggested by our experts, we added vegetation visualizations to our design. We also added the suggested soil triangle to the third tab shown above.
- 4 We added more filters to each visualization and grouped them closely together
- 5 Finally, we started to explore efficient ways to use our space. These horizontal bars were very effective (but not stacked, as we learned).

3g. Usability Testing Part 2

With our new design taking form, we conducted usability tests to validate our information architecture. Because we were only able to access experts remotely, two of the “usability tests” were un-moderated. They were, instead, conducted using a questionnaire alongside the interactive prototype. While non-traditional, this method allowed us to collect important feedback. The questionnaire walked participants through a series of tasks and is available in [Appendix C](#). We also conducted one in-person usability test using the think-aloud protocol with a high school science teacher, who represented one of our personas. We asked this in-person study participant to complete the same tasks as our remote participants, but we were also able to observe how he interacted with the interface and to ask more follow-up, open-ended questions.

Users were able to see an overview of our geocentric data on a map of the United States. They had no issues using the map to filter to specific national parks. Users also felt satisfied with the linking of the data across the various filters and views. Our in-person participant, in particular, was able to explore the vegetation plot tab by finding the plots and filtering to see both satellite and terrain pictures easily.

From both the remote and in-person tests, we gained valuable insights. Key insights and quotes are below.

Key Insights and Quotes

Insight 1. Difficult to distinguish column headers from cell data.

Design implication: Make column headers more prominent.

Quote: “Column headings should be larger... font or bold, etcetera...”

Insight 2. Confusion regarding interactivity of some design elements.

Design implication: Use better labeling throughout.

Quote: “I didn't know there was an option to select the bar.”

Insight 3. Industry-specific terms and jargon confusing to general population.

Design implication: Define terms.

Quote: “I don't know what this means... what is aspect?”

Insight 4. Users not immediately aware that they could select and use multiple dashboards in the visualization.

Design implication: Provide instructions above and/or within the visualization.

Quote: “I started in first tab because that was what was there. Once I had a hint I noticed other tabs but I don't think they're not clear enough. I just didn't notice...”

Resulting Design Changes

After this set of tests, we addressed each of these concerns; we refined the headings, set up term-definitions within the frame, and added simple instructions within the dashboard to guide users. We were starting to realize overlap within our three dashboards, and merged two of them. We found that this was the farthest we could condense them. In addition, our users responded negatively to the stacked bar graph, so we changed it to a standard bar graph.

Our visualization consists of two independent dashboards contained within an HTML page. The page begins with an introduction describing the intent of the visualizations, the data source upon which they are based, and a short instruction to select a tab to view the associated visualization. In addition, we added an expandable list of terms and definitions. During our testing, we found that some users were unfamiliar with some of the terms, so we included definitions to aid in understanding, especially since we are partially positioning this visualization as an educational tool. This functionality would be difficult to create in Tableau, but was fairly simple using JavaScript.

Hide list of terms

Aspect

In physical geography, aspect generally refers to the horizontal direction to which a mountain slope faces. For example, a slope on the eastern edge of the Rockies toward the Great Plains is described as having an easterly aspect.

Association

In community ecology and phytosociology an association is a type of ecological community with a predictable species composition, consistent physiognomy (structural appearance) which occurs in a particular habitat type.

Ground composition

The approximate percentage of each ground type that is found on the ground within the plot.

Leaf phenology

A characteristic describing the lifecycle of a plant's leaves based on the seasons.

Leaf type

A characteristic describing the physical nature of the plant's leaves.

Physiognomic class

A level in the classification hierarchy defined by the relative percent canopy cover of the tree, shrub, dwarf shrub, herb, and non-vascular life-form (in the uppermost strata) during the peak of growing season.

Slope intensity

The degree to which a plot is sloped overall. These slopes are grouped into categories based on their incline.

The expanded list of terms.

Below the introduction, a container with two tabs lets the user switch between visualizations. Each tab is labeled with the task that it supports, and each contains instructional text to help the user understand how they can interact with the visualization.

In addition to being able to provide additional instruction, building a surrounding webpage let us take advantage of CSS for defining styles more easily. We used this to apply branding that resembles the National Park Service.

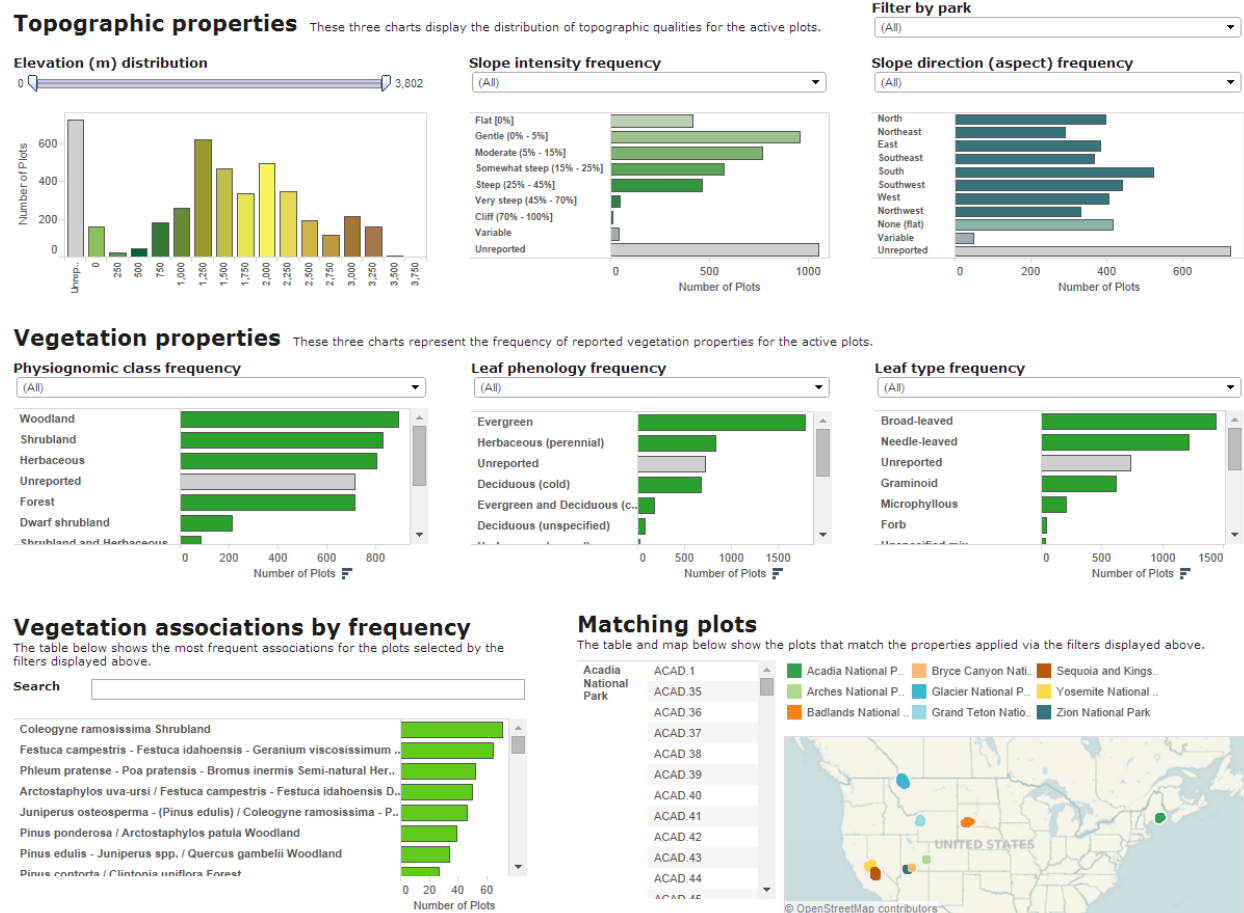
Finally, the web frame allows the accompanying content to be changed without requiring changes to the visualizations themselves. This means that we provide additional instructional text as necessary, for example.

4b. Dashboard 1: Find Vegetation Qualities by Plot Properties

We designed our first dashboard to display the relationships between topographic and vegetation qualities. We learned from our interview with Dr. Todd Keeler-Wolf that the primary tasks for scientists studying vegetation (whether in the national parks or not) and the reasons for collecting the data were to learn about the conditions under which particular vegetation associations occur.

To address these needs, we created a dashboard that highlights the distribution and frequency of several qualitative values. Each graph uses bars to encode values. This encoding was appropriate because we wanted to allow users to accurately compare the frequency with which each value occurs, and bars provide the highest precision in comparing values. In addition, as discussed by Stephen Few, length and 2-dimensional position are pre-attentive attributes, meaning these encodings are processed before

conscious attention and are therefore easily compared [19]. We used a consistent light gray color to mark the “Unreported” bars so that the user could easily distinguish between unreported and actual values.



Dashboard 1: “Find Vegetation Qualities by Plot Properties”

The individual graphs are interconnected to support extensive filtering of plots. When a user selects particular values within a single graph, the accompanying graphs update to reflect the filtered selection.

The dashboard is organized hierarchically into four primary sections. Though the graphs affect others regardless of their placement on the page, we ordered the sections based on the tasks that we heard from our user research.

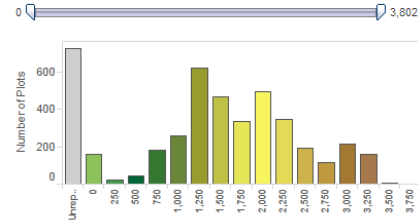
Dashboard 1: Section 1

The **Topographic properties** contains properties related to the physical nature of the land within the plot.

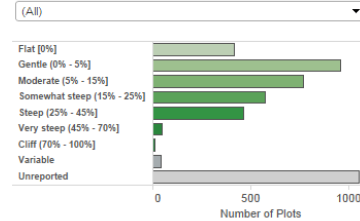
Topographic properties

These three charts display the distribution of topographic qualities for the active plots.

Elevation (m) distribution



Slope intensity frequency

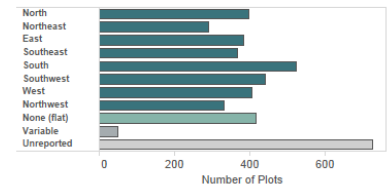


Filter by park

(All)

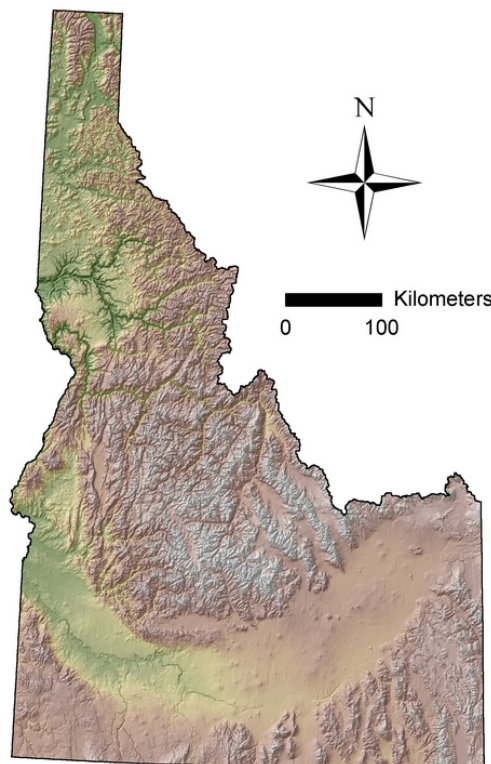
Slope direction (aspect) frequency

(All)



Filtering through topographic properties

The first visualization in this section is a histogram bar chart showing the **distribution of plots by elevation** in meters (including unreported values). Rather than using the same color for each bar, we chose to color the bars according to color scales frequently found on terrain maps. Though the colors do not encode any explicit data from our dataset, they provide a familiar scheme that may make the chart even more recognizable as elevation, as well as giving additional context when the filtered data only comprises a portion of the full elevation range.



Idaho Terrain Map, showing a typical application of color to approximate elevation [\[source\]](#)

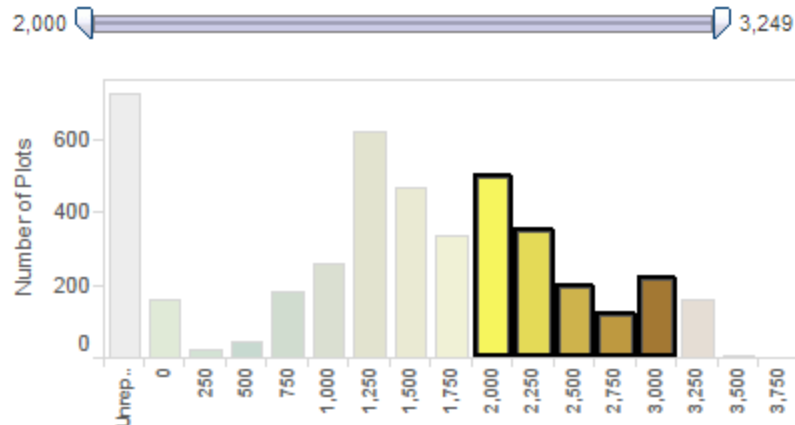
The second visualization in this section displays the **frequency of slope intensities** (including unreported and variable values). It is important to note that the Y axis for this visualization is not linear and therefore not a histogram: for example, the first bar ("Flat") contains slopes of only 0%, the second bar ("Gentle") spans a range of 5%, and another bar ("Steep") spans a range of 20%. This was intentional so that it maps to the categorical values that the surveyors used when collecting the data.

Finally, the third visualization in this section shows the **frequency of aspect**, or slope direction. The values are grouped to representing the nearest cardinal or intercardinal compass points, again including unreported and variable values.

Though the graphs' bars themselves act as filters, each visualization is accompanied by a traditional filter control to aid in discoverability, a need we identified during usability testing.

This section also contains a control for filtering by national park.

Elevation (m) distribution



Example of filtering with bar selection

Leaf phenology frequency

The figure shows a filter menu titled "Leaf phenology frequency". At the top is a dropdown menu currently set to "(All)". Below it is a list of 15 items, each with a checked checkbox: (All), Deciduous (cold and drought mix), Deciduous (cold), Deciduous (drought), Deciduous (unspecified), Evergreen, Evergreen and Deciduous (cold) mix, Evergreen and Deciduous (drought) mix, Herbaceous (annual and perennial mix), Herbaceous (annual), Herbaceous (perennial), Herbaceous and Deciduous (cold) mix, Unreported, and Unspecified mix. The menu is partially obscured by other UI elements at the bottom.

Example of traditional filtering

Dashboard 1: Section 2

The **Vegetation properties** section contains charts related to the general characteristics of the vegetation found in each plot.

The first visualization in this section shows the **frequency of each physiognomic class**, an identifier used to describe the overall mix of vegetation within a plot.

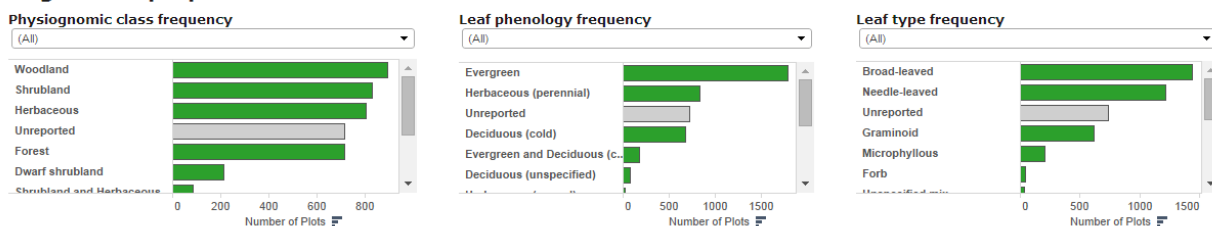
The second visualization show the **frequency of each leaf phenology**, a term used to describe the annual life cycle of plant's leaves (e.g. evergreen), for the dominant stratum (the primary plant used to define the group of plants) in that plot's association.

The third visualization shows the **frequency of leaf types**, again for the plot's dominant stratum.

Note that the height of the visualizations in this section are intentionally limited to highlight only the most frequent qualities for a given set of plots, but the user is free to scroll within each visualization to see those that are less frequent. Again, each visualization is accompanied by a more traditional filter control to aid in discoverability.

Vegetation properties

These three charts represent the frequency of reported vegetation properties for the active plots.



Traditional filtering of vegetation options

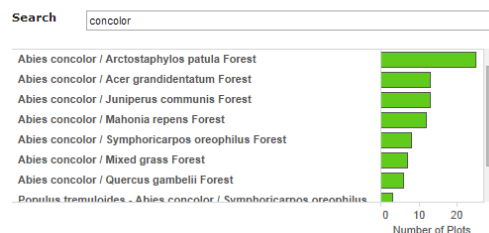
Dashboard 1: Section 3

The **Vegetation associations by frequency** section consists of a single visualization showing the list of associations that occur in the matching plots, sorted by frequency. Like the other visualizations, the functionality of the association list is bidirectional: the list is affected by the selections within other graphs, and selecting one or more associations likewise affects the other graphs.

To aid in the location of a particular association (e.g. if a researcher wants to learn where a specific association can be found), we included a simple search bar that matches a string of text located anywhere in the association name.

Vegetation associations by frequency

The table below shows the most frequent associations for the plots selected by the filters displayed above.

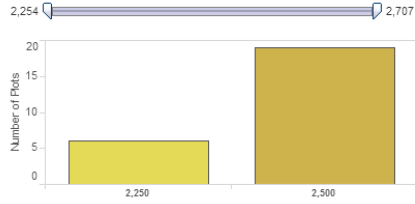


Searching by entering text string

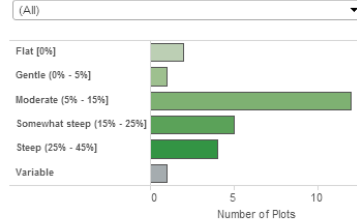
Topographic properties

These three charts display the distribution of topographic qualities for the active plots.

Elevation (m) distribution



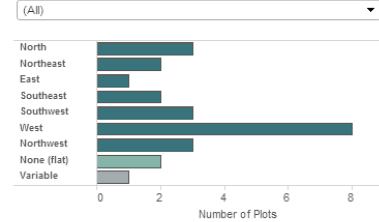
Slope intensity frequency



Filter by park

(All) 

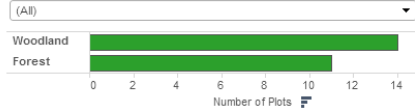
Slope direction (aspect) frequency



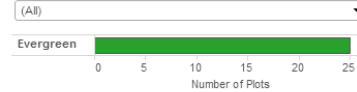
Vegetation properties

These three charts represent the frequency of reported vegetation properties for the active plots.

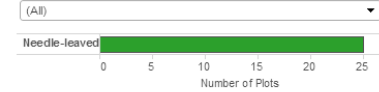
Physiognomic class frequency



Leaf phenology frequency

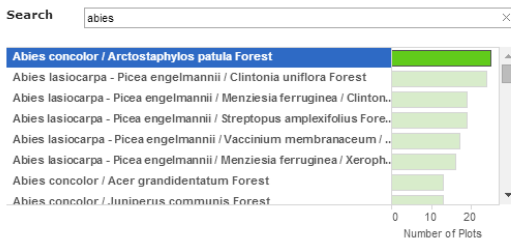


Leaf type frequency



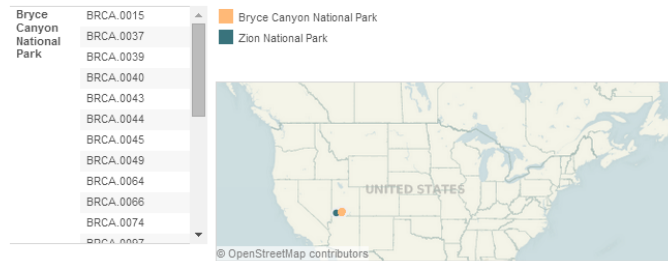
Vegetation associations by frequency

The table below shows the most frequent associations for the plots selected by the filters displayed above.



Matching plots

The table and map below show the plots that match the properties applied via the filters displayed above.



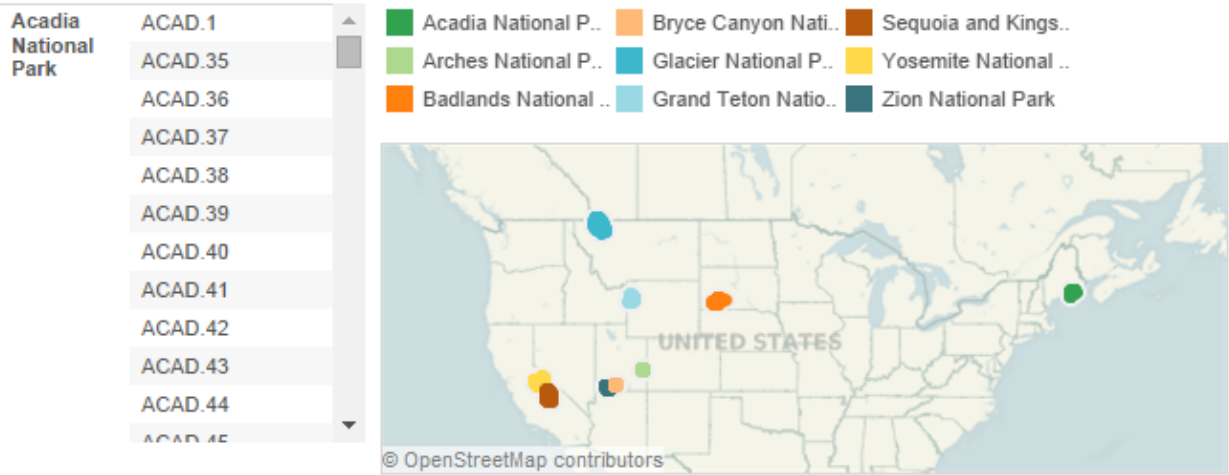
Selecting a single association (*Abies concolor* / *Arctostaphylos patula* Forest) filters the other views. With this selection, an analyst can learn that the selected association occurs more frequently on west-facing slopes than on east-facing slopes. Note also that the Physiognomic class frequency chart shows two terms for the association. While we consolidated terms whose meanings were obviously the same (e.g. "East" and "E"), we avoided making assumptions about others.

Dashboard 1: Section 4

Finally, the **Matching plots** section contains a list and map of plots that match the selections within the aforementioned graphs and filters. Though the intent of this dashboard is not to focus on the discovery or analysis of individual plots, it provides a simple means to identify plots that may be of further interest. For example, a user might select a single association and find that it is only reported as occurring in a few plots. Though the data in the dashboard may not be enough to make conclusions, the user could make note of the plots to conduct further research about that particular association. This feature may also be useful for researchers to locate plots whose data reporting is incomplete (i.e. by selecting "Unreported" in a graph) in order to fill in missing data values.

Matching plots

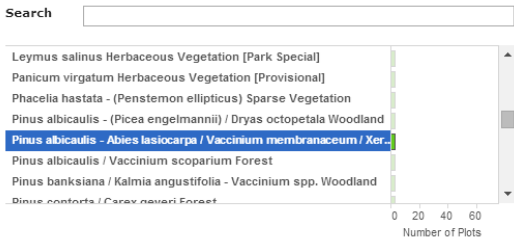
The table and map below show the plots that match the properties applied via the filters displayed above.



The list and map of Matching plots, showing all listed plots (no filters applied).

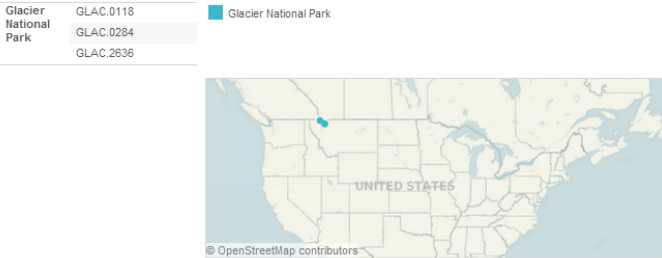
Vegetation associations by frequency

The table below shows the most frequent associations for the plots selected by the filters displayed above.



Matching plots

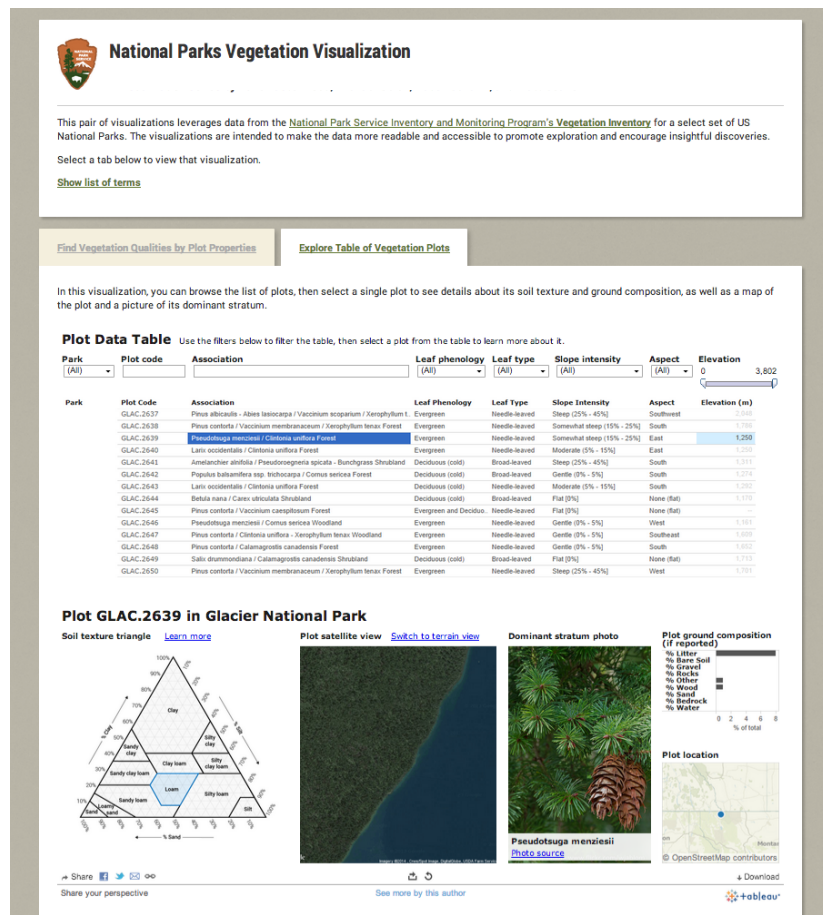
The table and map below show the plots that match the properties applied via the filters displayed above.



Example of filtered list of plots

Finally, in addition to the visualizations themselves, the dashboard combine basic explanatory text and labeling to clarify the intent of each section. Though not intended to serve as a complete guide to using the visualization (more help is contained in the surrounding web page), we learned from our usability testing that simple labeling would help users better understand what they were looking at.

4c. Dashboard 2: Explore Table of Vegetation Plots



Dashboard 2: "Explore Table of Vegetation Plots"

Our second dashboard was designed to serve as an enhanced way to browse the list of plots from our dataset. We learned from our class material, our expert interviews, and our user evaluations that certain visualization techniques were better for displaying some bits of data, either because they were familiar to experts or because they evoke additional insight or emotional response (these particular techniques are discussed below in the second section).

This dashboard is organized into two sections, and is ordered based on interaction methods:

Dashboard 2: Section 1

The **Plot Data Table** section contains selected fields (primarily categorical data) from our original datasets in tabular form. Each plot is organized by park and shows its association, leaf phenology and type, and topographical data. Each column is accompanied by a filter control (optimized for the data type) to allow the user to filter the table down. Clicking upon a single plot activates it and displays more of its qualities in the section below the table.

Dashboard 2: Section 2

The **Plot <plot code> in <national park name>** section shows additional visualizations and details about the selected plot.

For the first three visualizations, we used a Tableau web page frame to load in external content. We wrote a PHP page that accepted several variable values passed via the plot table. The PHP page used the variable values to generate HTML for showing the visualizations that we chose. Before a plot is selected, a title instructs the user to select a plot. In one of our early prototypes, a plot was pre-selected, and users were unsure what they were seeing. Once the user chooses a plot, the plot's name and park is displayed.

The first component in this section is a **soil texture triangle** that, as we learned from our user interviews, is used extensively in evaluating soil science. The triangle assigns a categorical texture name to a sample of soil based on its composition of sand, silt, and clay. We opted to use a traditional chart due to its familiarity to our target audience. Introducing a new visualization would have required our experts to learn a new encoding scheme. In addition, we included a link to learn more for users who may be unfamiliar with the chart. The chart was implemented by creating a series of images (one for each soil texture) and loading the appropriate image based on the value passed to the PHP page.

The second component is a **close-up map of the plot's location**. When a plot is selected, a satellite view of the plot's location is shown. The user also has the ability to toggle the view to a terrain view in order to see the topography of the area (e.g. slope, aspect, etc.). The views are generated by passing the plot's latitude and longitude values through the Google Static Maps API, with JavaScript functionality providing the toggling between views. In the future, it would be ideal if we could highlight the perimeter of the plot within the satellite view, but this was out of scope for this project.

The third component is a **photo of the dominant stratum** of the selected plot's association. Like the above components, the association value is passed to the web frame via PHP, which trims the full association to the first two words in order to isolate the first species name. The species name is then passed to Flickr API (as the *text* argument in the *flickr.photos.search* method), returning the first search result and displaying it on the page. Though the results are necessarily unpredictable because we are pulling in public pictures dynamically, we have found the method to be reasonably reliable in returning appropriate photos. The photo is accompanied by a species label and a link to the photographer's Flickr profile for attribution.

The next component is a bar chart showing the **percentage of each ground cover type** present in the plot (when available). Rather than a traditional pie chart, we used a bar chart to allowed for better visual comparisons.

Finally, a **small map** is shown to provide locational context to the plot (which is not captured in the zoomed-in satellite view).

5. Evaluation & Discussion

Does our visualization support our users and tasks?

Our design directly supports many of the goals and tasks we discovered through our preliminary and user research. As a result of our user-centered design process, we are confident that our primary park ranger persona will be able to easily compare values across parks and find correlations. Our users commented that the visualization would be a useful tool for education and outreach, which supports the needs of our secondary park ranger persona. While our usability test with the high school science teacher suggested that students might not explore the tool on their own, it would play a key role in a lesson plan or as an assignment to our teacher persona.

Effectiveness of Views

Strengths: Overall, users spoke positively about the concept and the design of our visualization. Users liked the depth of the interaction. One participant responded that filtering was “appealing,” and that it felt “like [he] was doing something.” Although our most recent version has not been tested, we feel that it provides a strong overview on the first tab, and relevant detail on the second. Our views also follow Tufte’s principle that, “Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself” [20]. We took this to heart and focused on clean, efficient labeling.

Weaknesses: While the views effectively communicated a message, we found through usability testing that there were minor problems with the user interface. For example, users found selecting precise elevation values using the slider challenging because even moving the slider control a small distance would result in traveling over many values. Also, it was not always apparent to the users that they could click in the text boxes to type precise values. As a result of these two shortcomings, one usability testing participant responded that it was “impossible to select an exact value.”

Effectiveness of Techniques

We used several of Schneiderman’s infovis techniques in our visualization to support our intended tasks.

Overview

The first tab of our visualization most closely resembles an overview as Shneiderman would describe. It shows most of the data unfiltered and provides a map for context. Users see all of the national parks included in our visualization and are able to select parts of the whole in order to access more detailed views. As a result of this interaction, the context of the visualization is preserved as users navigate to specific views. Our screen space was also limited. As opposed to showing a full overview of a single variable, we chose to depict many variables simultaneously and in moderate detail. [21]

Filter/Zoom

After performing usability tests on the first version of our interactive prototype, it was clear that some of the basic information visualization tasks were lacking. For example, it was difficult for our users to make a connection between the filters and what they were controlling. With this insight we were able to design a much clearer interaction model throughout our final iterations. We did this by

adding help text to the screen and arranging the UI elements in a consistent and intuitive pattern. We also included strategically placed traditional filters that appeared closer to their associated views to make it more obvious what would change if our users interacted with the sliders. Subsequent usability testing validated the effectiveness of our design changes, making filtering more straightforward in all future versions of our visualization.

Because we were utilizing the built-in zoom interactions of Tableau, the interaction model used to zoom in and out of our geospatial view was very simple. Users had some issues with not knowing that double-clicking resulted in zooming into the map, but they learned quickly and adjusted their behavior as a result. Users were also comfortable with dragging semi-translucent squares over bars and map elements in order to filter and zoom through information.

Details On Demand

Because our data are very dense, it was crucial for us to support providing users details on demand rather than cluttering views with too much information. To accomplish this, we included detailed information in tooltips. When a user wants to know more about a specific data point or location, they were able to hover or click on the UI element and a tooltip appears with more detailed information.

Relate (Brushing/Linking)

Before conducting usability testing, our brushing and linking methods did not support our user tasks well. For example, a filter that controlled one of our physiognomic bar charts was poorly labeled and situated far away from it. As a result, it was not clear to our users that the slider controlled the information presented in the chart. After our usability tests revealed this insight, however, we rearranged our interface and information architecture to provide greater synergy between our information visualization's input and output interactions.

During our second round of usability tests, which occurred after we made changes to our interface to better support brushing and linking, one participant commented that it was very "appealing" when he made a change in one view of the visualization and it affected the other frames. This demonstrated to us that the changes we implemented were effective.

Insight from usability testing guided additional design decisions. We included more labeling and help text that explicitly stated what the user could interact with to produce their desired views. For example, we included the following user help text: "Use the filters below to filter the table, then select a plot from the table to learn more about it."

Infovis Tasks Not Supported

Our visualization minimally supports history and extraction. We stuck with Tableau's default behavior, which allows users to 'reset' a page or save a view as an image or PDF file. None of our users mentioned these types of features. While they might be expected in some sophisticated tools, history and extraction would not change the core user experience of our visualization.

Challenges

While planning, designing, and completing the visualization, we faced various challenges:

Scheduling Participants

Because our primary and secondary target users work in a highly specialized field, it was difficult to find people who matched the description of our target users. Additionally, many of these specialists work in remote locations, so we were unable to meet with them in person. However, we received generous introductions from our colleagues and professors and conducted phone interviews with subject matter experts.

To supplement our limited access to expert users, we also conducted usability tests with non-experts to discover issues with the UI controls.

Cleaning the Data

A large component of our project involved cleaning the datasets. As we started exploring the data, inconsistencies arose: headings differed or were misaligned, terminology was inconsistent across datasets, and some quantitative data were clearly inaccurate. For example, when we imported the longitudinal and latitudinal data into Tableau we quickly noticed that some plots' coordinates were recorded incorrectly. After further exploration and data cleaning, we were able to safely correct inaccuracies, or to exclude the data where appropriate.

We decided not to exclude plots that had some inaccurate data. We did so because including the plot data that was reported was still very important to our users.

Balancing User Needs

Initially, we wanted to cater to experts and the general population. As we started implementing expert feedback and domain specific terms, we realized that there was a distinct conflict between the needs of our audiences: the visualization needed to be useful for experts but evocative and friendly enough for non-experts.

Even though balancing user needs was a challenge we faced throughout the design, testing, and implementation of our visualization, we were able to come to a compromise that allowed users of varying expertise to benefit from the visualization.

6. Future Work

Although our visualization is fully functional, there is room for improvement and future research:

Include more national parks: The I&M program continues to collect data and manages projects that span many years. A large number of vegetation datasets are slated for completion later in 2014. The majority of these will be appropriate to add to our visualization. While there are minor components that were built with this limited set of national parks in mind (e.g. the color legend of the 9 national parks), the visualizations are otherwise extensible and scalable to any number of national parks by adding them to the dataset. In the future, it would be valuable to incorporate more parks so that more US regions and associations are represented.

Add historical weather data: One of the experts we interviewed noted that historical weather data would be helpful in combination with our soil and vegetation data. Integration of this type of time-dependent data poses numerous challenges but may be interesting to explore further.

Combine dashboards: Our initial goal was to include all of the variables onto a single, dynamic dashboard. As described above, we started with multiple dashboards, then consolidated based on results from usability testing, but finalized on a two dashboard user interface. Further research should be conducted to ensure that users can effectively and efficiently explore the data with this UI.

Manage Tableau's limitations: While Tableau was incredibly useful within its boundaries, we did brainstorm some ideas that were not feasible to implement within Tableau. It may be interesting to test out other toolsets such as D3 with these data for more customization.

Conduct additional in-person usability testing with field experts: Findings from our usability tests were valuable and quickly led to design changes. We had to rely, however, on some remote feedback. Because we were not able to watch experts interacting with our tool and interact with them while they were doing so, our findings are relatively limited and further usability testing with domain expert users should be pursued to gain a better understanding of the true target users' user experience.

Participatory research and design: Many of the I&M datasets require domain specific knowledge to fully comprehend and use. Collaborating with data collectors and domain experts throughout a user-centered design process may result in an improved visualization with an improved user experience for target users.

Create a visualization framework for future inventory programs: One thing we noticed early on is that much of the I&M data is presented but sometimes only loosely connected. An abstract goal of ours was to show the advantage of connecting the data into a larger framework. Future work should include creating a general visualization frame for I&M data that allows for new or updated data to be easily imported. This type of dynamic and collaborative tool could provide structure and cohesion to the many datasets.

8. References

- [1] National Park Service. 2009. Strategic plan for natural resource inventories: FY 2008 - FY 2012. Natural Resource Report NPS/NRPC/NRR—2009/094. National Park Service, Fort Collins, Colorado.
- [2] Natural Resources Conservation Service. "Soil Health." United States Department of Agriculture. USDA, Washington, DC. n.d. Web. 05 June 2014.
<<http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/health/>>.
- [3] Natural Resources Conservation Service. "Soil Health Management." United States Department of Agriculture. USDA, Washington, DC. n.d. Web. 05 June 2014.
<<http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/health/mgmt/>>.
- [4] "Inventory & Monitoring in Parks." *NPS Inventory and Monitoring Program*. National Park Service, Fort Collins, Colorado., 22 January 2014. Web. 05 June 2014.
<<http://science.nature.nps.gov/im/about.cfm>>.
- [5] United States. National Park Service. "Example GIS Projects in Yellowstone Park." *National Parks Service*. U.S. Department of the Interior, 27 May 2014. Web. 05 June 2014.
<<http://www.nps.gov/yell/naturescience/example-gis-projects.htm>>.
- [6] Madden, Maguerite. "Vegetation Modeling, Analysis and Visualization in U.S. National Parks." *Center for Remote Sensing and Mapping Science (CRMS)*. Dept. of Geography, University of Georgia, Athens, Georgia, n.d. Web. 01 June 2014.
<<http://www.isprs.org/proceedings/XXXV/congress/comm4/papers/540.pdf>>.
- [7] "Interagency Monitoring of Protected Visual Environments." *IMPROVE*. Colorado State University., n.d. Web. 18 April 2014. <<http://vista.cira.colostate.edu/improve/Default.htm>>.
- [8] "Spatial and Seasonal Patterns." *IMPROVE*. Colorado State University., n.d. Web. 18 April 2014.
<http://vista.cira.colostate.edu/improve/Data/Graphic_Viewer/seasonal.htm>.
- [9] "Metadata Browser." *IMPROVE*. Colorado State University., n.d. Web. 18 April 2014.
<<http://vista.cira.colostate.edu/improve/Web/MetadataBrowser/MetadataBrowser.aspx>>.
- [10] Patterson, Tom. "Ideas and Techniques about Relief Presentation on Maps." Shaded Relief, 21 May 2014. Web. 05 June 2014. <www.shadedrelief.com>.
- [11] "Welcome." Relief Shading, 28 August 2008. Web. 04 June 2014. <www.reliefshading.com>.
- [12] "Natural Earth." N.p., 2014. Web. 04 June 2014. <<http://www.naturalearthdata.com/>>
- [13] "Home." North American Cartographic Information Society. N.d. Web. 05 June 2014.
<<http://www.nacis.org/>>
- [14] "1:50m Physical Vectors | Natural Earth." *Natural Earth*. Natural Earth, n.d. Web. 11 May 2014.
<<http://www.naturalearthdata.com/downloads/50m-physical-vectors/>>.
- [15] "Welcome to GDG." *Geospatial Data Gateway*. Natural Resources Conservation Service, n.d. Web. 3 May 2014. <<http://datagateway.nrcs.usda.gov/GDGHome.aspx>>.

- [16] "Web Soil Survey - Home." *Web Soil Survey - Home*. Natural Resources Conservation Service, 12 June 2013. Web. 05 June 2014. <<http://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm>>.
- [17] "Web Soil Survey." *Web Soil Survey*. United States Department of Agriculture, n.d. Web. 1 June 2014. <<http%3A%2F%2Fwebsoilsurvey.sc.egov.usda.gov%2FApp%2FWebSoilSurvey.aspx>>.
- [18] "AirAtlas." *NPS: Explore Nature » Air » Air Atlas*. National Park Service, 3 July 2012. Web. Apr. 2014. <<http://www.nature.nps.gov/air/maps/AirAtlas/index.cfm>>.
- [19] Few, Stephen. "Thinking With Our Eyes." *Now You See It: Simple Visualization Techniques For Quantitative Analysis*. Oakland, CA: Analytics, 2009. 40-42. Print.
- [20] Tufte, Edward R. "Graphical Integrity." *The Visual Display of Quantitative Information*. Cheshire, Conn. (Box 430, Cheshire 06410): Graphics, 1983. 77. Web. 7 Apr. 2014.
- [21] The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, Shneiderman, Proc. IEEE Conference on Visual Languages, Boulder 1996.

9. Appendices

Appendix A: Personas



Age: 44

Occupation: Park Ranger, soil scientist

Location: Yosemite National Park, CA -- but visits other parks to consult with other park rangers

Primary persona:

Dr. Sally Smith

National Park Service park ranger, specializing in soil science

Sally is a park ranger and soil scientist under the National Park Service. Above all else, the goal of park rangers in United States national parks is to protect and preserve the park's resources and guests, and Sally's contribution to the protection of our national parks is her expertise in soil composition and integrity. She spends most of her time in Yosemite National Park, though her role has her visiting other national parks across the nation when there are concerns about a park's soil.

While Sally began as a guide several years ago, she began focusing on soil after participating in one of the first vegetation inventory projects in Yosemite National Park in the early 2000s. She became fascinated with the detailed data that she had been helping collect, including soil composition, elevation, slope, water drainage, vegetation distribution, and more all throughout the park.

These days, Sally's main tasks are ensuring that the parks' soil goes through its natural processes, accounting for changes in soil composition based on changing environments, and advising park and project stakeholders about any hazards related to the soil.

The Vegetation Inventory project, which will eventually span all national parks, has provided an overwhelming amount of valuable data about the parks' soil and vegetation, and Sally uses the data often in order to understand differences in regions of the parks she helps manage. However, in its raw, tabular form, the vast amount of data is difficult to make much sense of; she spends a lot of time looking up specific data points. In particular, she doesn't have a good way of getting an overview of multiple plots of land or making comparisons between plots. Although she knows the nature of the data well enough to pick out the data points she needs, she has a tough time helping other rangers who are less familiar with the data.

Goals

Get an overview of the soil composition across multiple parks and within different regions of the same park

Understand relationships between multiple variables and their effects on soil in order to recognize trends and make predictions

Quickly find the specific data points she needs, such as looking up values for plots nearby a new building project

Obstacles

Lots of data to sift through manually

Impossible to look at the data at an aggregated level

Data is presented entirely tabularly so it's difficult to recognize extreme values and trends



Secondary persona:

Ed Westgate

National Park Service park ranger,
focusing on education

Ed is a park ranger who has been working in Arches National Park for the past 5 years. His passion has always been educating and getting people excited about the country's natural resources. He took a position in the National Park Service to help spread the message of conservation and protecting the parks' ecology.

As a ranger, Ed spends a lot of his time giving educational and interpretive programs to the park's guests, and his programs range from talking about the food chain, showing visitors evidence of the park's changing geology, and leading cultural re-enactments. In his downtime, Ed spends a lot of his time learning about new fields so that he can be ready to answer his enthusiastic guests' wide range of questions.

Ed heard about the Vegetation Inventory project and was inspired by the amount of time and effort that went into documenting the parks' vegetation and soil. While he didn't participate in the collection of the data, he's eager to learn about what it has to offer. However, when he looked at the data, he was quickly overwhelmed and unable to discover anything that might be interesting to his visitors. He was hoping to learn about how geography, climate, geology, and plant and animal life affect the soil and future vegetation, but the plot-level data was too granular to be able to understand.

Age: 53

Occupation: Park Ranger, educational guide

Location: Arches National Park, Utah

Goals

Skim through the data and look for anything interesting or exciting that he can share with his guests

Understand the soil make-up of his park to be able to respond to visitor questions

Make comparisons between different parks to be able to compare and contrast Arches with others across the country

Obstacles

Doesn't understand how the data was collected, so it's difficult to understand what it all means

No overview means that he can't get a quick, high-level understanding, which is essential for the diverse topics that his educational talks must span

Data on individual plots isn't very useful to him since he talks about the park in general



Tertiary persona:

Jim Murray

High school teacher, focusing on
Earth Sciences and Ecology

Jim is a high school teacher who has recently switched into teaching Earth Sciences and Ecology courses. His students usually seem disinterested, especially during boring segments like the importance of healthy soil, but he knows that if he can show them something "cool," they perk up and get excited.

Jim found the Vegetation Inventory data sets for the US national parks, and he thinks the data would fit in perfectly with the chapters they're reading in the next month. However, he can't figure out what the data shows, and he can't imagine that his students would be interested in the data in the current format. But if he could make the data interactive and visual, he thinks he could get his students excited about learning more about their national parks.

Goals

Teach his high school students about the geologic diversity within their country

Make the course material more exciting through the use of visuals and exploration

Obstacles

Students often need some kind of visuals to understand or get excited about course material

Unfamiliarity with underlying data makes it difficult to use the data as a learning source

Age: 39

Occupation: High school science teacher

Location: Rancho Bernardo High School, California



Appendix B: Interview guide questions

Interview Questions

Note: each interviewee did not answer an identical set of questions. For example, our expert from the National Park Service was asked direct questions about the datasets whereas our soil expert from Seattle Tilth was asked more questions about education and outreach.

- Tell us a little bit about your background:
 - What is your role at [org]?
 - What do you consider to be [org's] most important effort?
- [Dr. Keeler-Wolf only] In practice, what do you expect will happen with the data from these datasets? (How will they be utilized?)
- What are the most important justifications for collecting soil composition data?
- What kind of visualization of these data would be the most useful?
- Is it most useful to look at one park's data at a time? Or multiple? (probe for depth)
- Are there any particular correlations that you look for between any of the data dimensions in the datasets?
- How would an interactive visualization of soil composition data from the I&M Project be best used? (Prompt: Education? Outreach? Exploration?)
- Are there any kinds of teaching tools that would be useful to you in improving your education and outreach efforts?
- Would a data visualization of soil composition in national parks be interesting to your community as a way of learning about soil composition?

Appendix C: Questionnaire with tasks for expert users

1. What is the "association" and "soil texture" for plot GT-03A036 in Grand Teton National Park?

How easy or difficult was it to find the answer to this question?

1 2 3 4 5

Very Easy

Very Difficult

Why did you rate the difficulty this way? In what ways was it easy or difficult?

Any additional comments about this task?

2. Find the satellite and terrain pictures of the plot you found in the previous question (GT-03A036). How easy or difficult was it to find?

1 2 3 4 5

Very Easy

Very Difficult

Why did you rate the difficulty this way? In what ways was it easy or difficult?

Any additional comments about this task?

3. In which parks can you find plots with north-facing (including NW, NE, and N) slopes with an intensity between 0%-25% at an elevation between 370-1,200 meters?

(Check all that apply)

- Acadia
- Arches
- Badlands
- Bryce Canyon
- Glacier
- Grand Teton
- Sequoia and Kings Canyon
- Yosemite

- Zion
- None
- I was not able to find the answer

How easy or difficult was it to find the answer to this question?

1 2 3 4 5

Very Easy

Very Difficult

Why did you rate the difficulty this way? In what ways was it easy or difficult?

Any additional comments about this task?

4. What is the most common "association" above 2500m of elevation?

5. What is the most common "physiognomic class" in this elevation range?

How easy or difficult was it to find the answers to the above two questions?

1 2 3 4 5

Very Easy

Very Difficult

How interested are you in continuing to explore this visualization?

1 2 3 4 5

Not at all interested

Very interested

Why did you rate the difficulty this way? In what ways was it easy or difficult?

Any additional comments about this task?

6. Do you have any additional comments about our information visualization?

Appendix D: Datasets

Vegetation Datasets on the National Park Service Website:

<http://science.nature.nps.gov/im/inventory/veg/products.cfm>

Original data sources:

[Acadia](#)

[Arches](#)

[Badlands](#)

[Bryce Canyon](#)

[Glacier](#)

[Grand Teton](#)

[Sequoia](#)

[Yosemite](#)

[Zion](#)

Final, compiled, and polished dataset from which our dataset is running:

<http://www.mattsoave.com/projects/nationalparksvis/content/Compiled%20data.xlsx>