INFO 247 · Spring 2014

Final Report

Fire and Drought in California

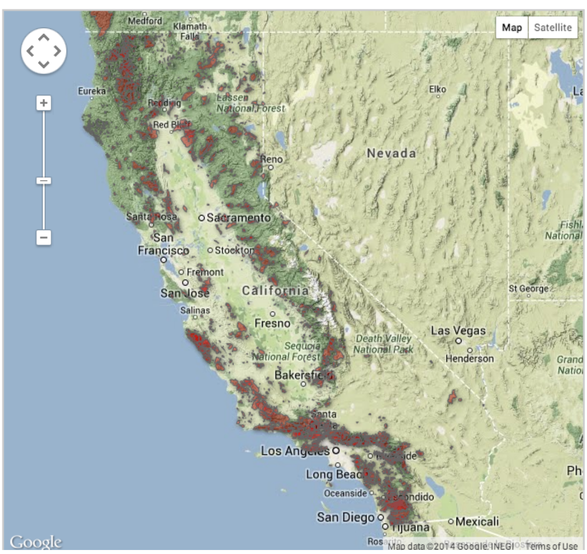
Team Members
Xavier Malina
Peter Swigert
Cameron Reed

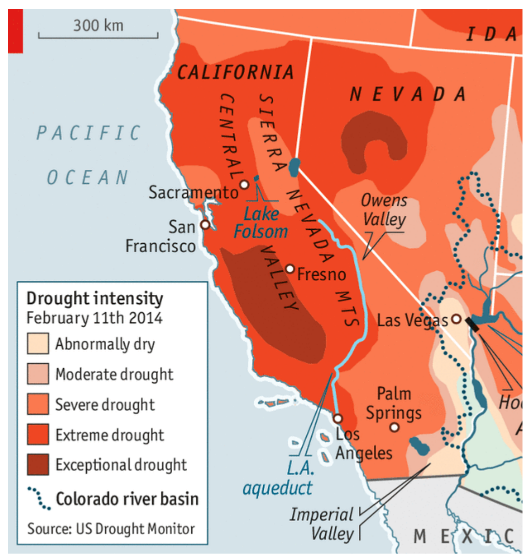
Project Goals

Our primary goal for this project was to design and build a directed narrative on historical drought and wildfire impacts in California. We wanted to provide users with an initial explanation of the core narrative, but then facilitate their exploration of drought and fire data over time through methods like filtering, highlighting and tooltips. To facilitate this exploration, we planned to build a website with interactive visualizations. We wanted to both present an interesting story about California, and enable users to uncover and visually explore the relationship between wildfires and drought across different variables.

Related Work

Drought and wildfire data is often visualized with maps, especially heatmaps and choropleth maps which show the extent and severity of the affected areas. This approach is valuable for answering the question of where things are happening. However, standard heatmaps are not good for showing the connection between drought and wildfire variables. Instead, we would like to visualize the data over time and focus on clarifying the relationship between drought and wildfire.





Left, an interactive map showing areas burned by fire since 1950[1]*. Right, a typical static heatmap showing drought intensity in the Western United States*[2]*.*

[1] Levine, Nathaniel. The Sacramento Bee. Jan 7 2014.
http://www.sacbee.com/2011/07/27/3798480/interactive-map-a-history-of-wildfires.html
[2] “The drying of the West”. The Economist. Feb 22 2014.
http://www.economist.com/news/united-states/21596955-drought-forcing-westerners-consider-wasting-less-water-drying-west

Data

We pulled data from two government sources. Our California wildfire dataset was derived from Dr. Karen Short's national wildfire geodatabase from the U.S. Forest Service, accessible through USDA[3] or the National Park Service[4]. It contains data on all reported wildfires in the United States from 1992 to 2011. The fields we extracted included the geo-coordinates of the point of ignition as well as fire size, ignition date, and cause.

Drought data came from the NOAA National Climatic Data Center[5], which maintains historic climate data for all 344 climate divisions in the contiguous United States. We used the NCDC data search tool[6] to obtain precipitation, temperature, and drought index data for the seven California climate divisions for the period of January 1895 to April 2014.

Tools

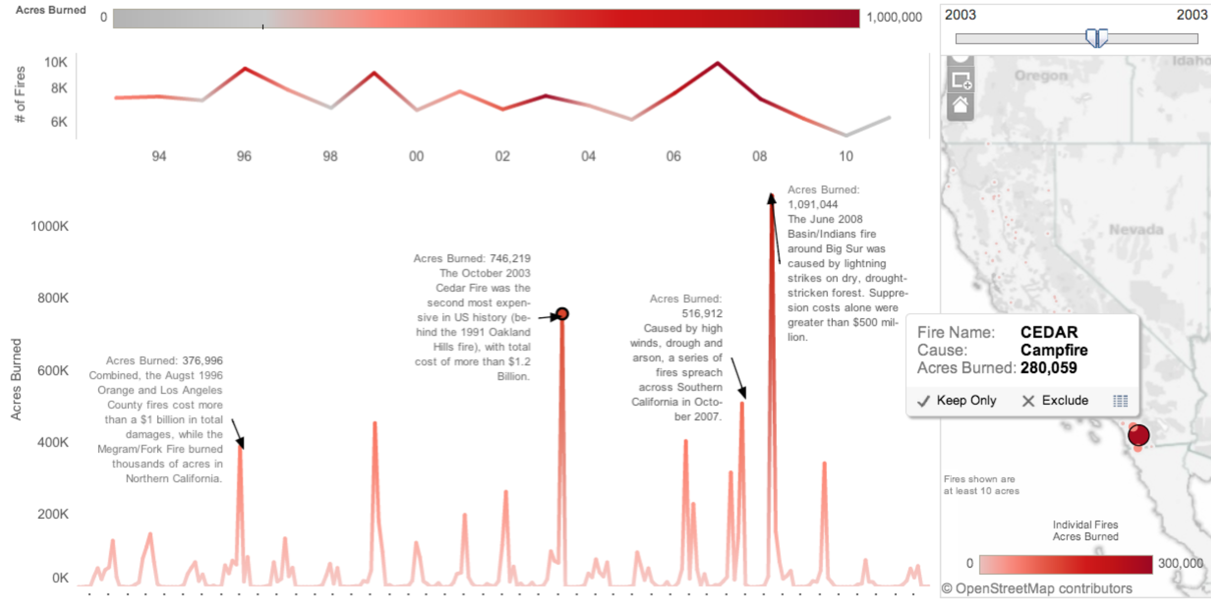
We had initially planned on exploring the data in Tableau and building an interactive visualization in D3. However, after creating some exploratory charts in Tableau we realized that we could achieve the vast majority of the same output that we had hoped for with D3. A D3 component of the narrative was envisaged and implemented to about 60% completion before being put aside in favor of more work being achieved in Tableau. This decision left us much more time to focus on applying information visualization principles to our final product. We also discovered that Tableau dashboards could be embedded in our own webpage, which gave us more control over the context in which the dashboards would be presented. We used the Cool Kitten[7] parallax framework as the foundation for our website and shared our code on GitHub[8].

Visualization

We visualized the wildfire and drought data with six interactive Tableau dashboards arranged on a long scrolling webpage. The main content was broken into three sections: California Wildfire History, California Drought, and Wildfire and Drought. Each section included an introduction to the section topic and explanations of the dashboards. Throughout the document we tried to achieve a balance of explanation and exploration, pointing the user towards interesting patterns and data while allowing them to interact with and navigate the content. Each dashboard is described below along with user feedback specific to that section that we received.

[3] http://www.fs.usda.gov/rds/archive/Product/RDS-2013-0009
[4] https://irma.nps.gov/App/Portal/Home
[5] http://www.ncdc.noaa.gov/cdo-web/search
[6] www.7.ncdc.noaa.gov/CDO/CDODivisionalSelect.jsp#
[7] https://github.com/jalxob/cool-kitten
[8] https://github.com/cameronreed/fire_and_drought

California Wildfire History



This dashboard utilizes raw wildfire data from a USDA Wildfire Data database[9], to show the change in number of wildfire ignitions and number acres burned by wildfires in California since 1992. The upper line graph shows that the total number of fires has been relatively steady over time with no significant trend up or down. The lower line graph shows that there have been increasingly large spikes in the number of acres burned during fire season, showing that fires are burning more acres than before.

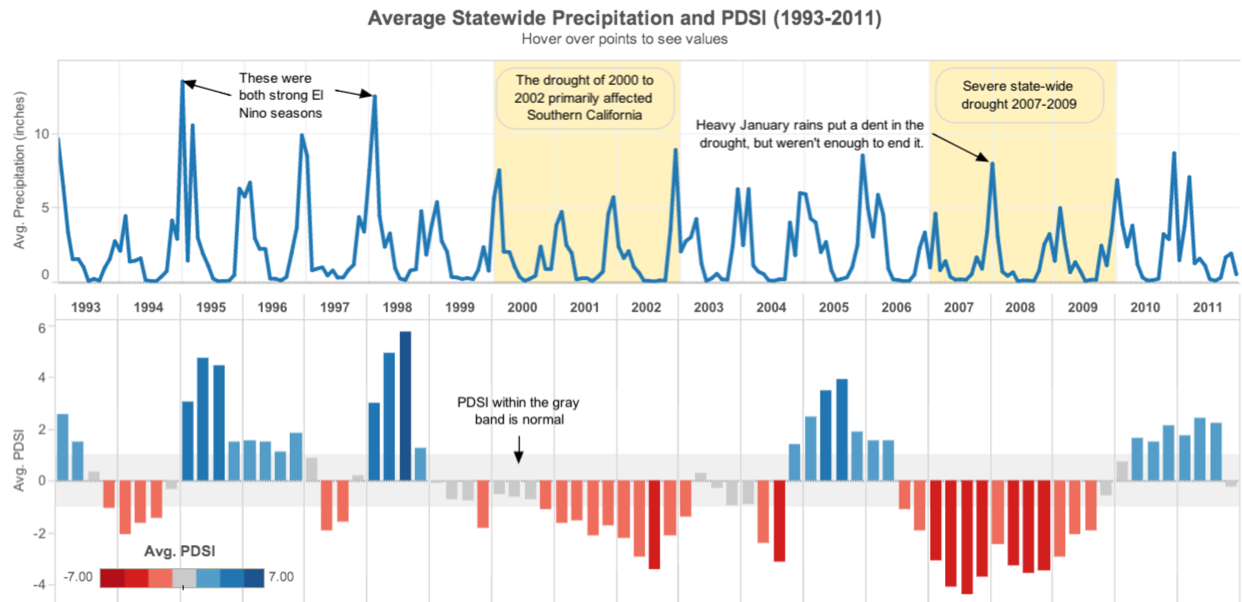
Annotations on four of these spikes are used to provide some description of the fires and the damage they caused. The slider in the top-right allows users to explore where fires appear in California over different time frames. Color and size and both used to encode the number of acres burned on the map. Hovering over any fire provides details on demand, which includes the name of the fire, the cause of the fire, and the number of acres burned. Hovering over any part of the line graph shows the date and the measure (number of fires or acres burned), and selecting a given point filters the map to show those specific fires.

In user testing, users often did not immediately realize that they could interact with the dashboards. We considered including instructions on the interactions but felt that covering all the options would clutter the page and confuse the user, and that it would be more effective to let users learn the interactions as they progress through the page.

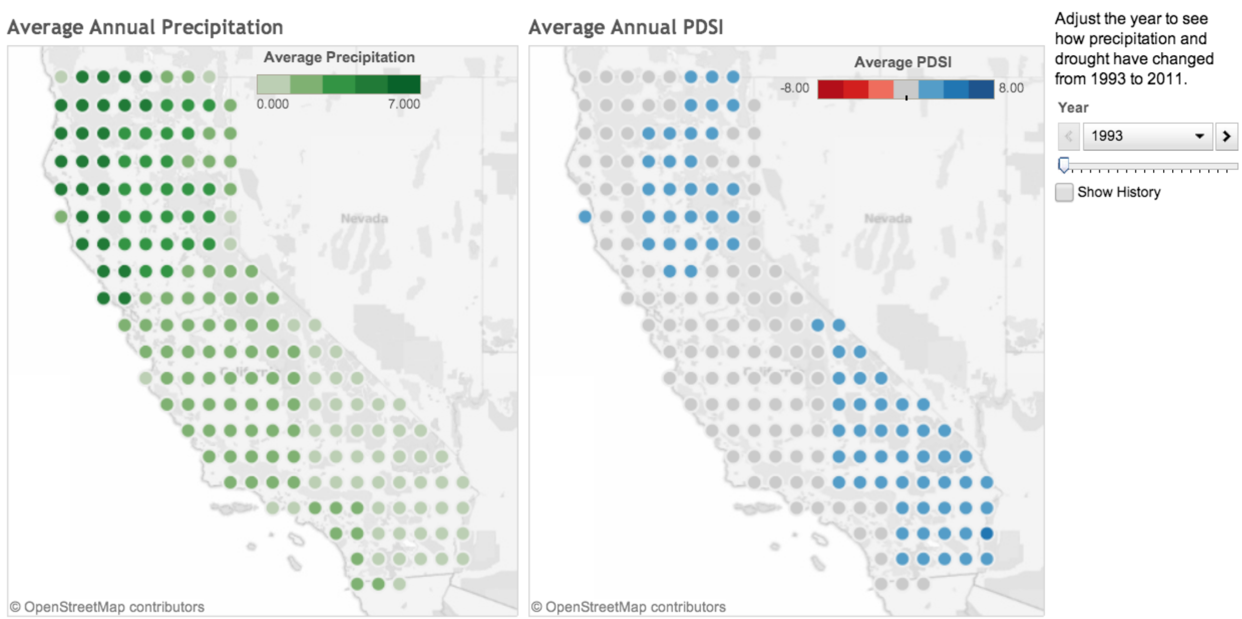
---

[9] Short, Karen C. 2013. Spatial wildfire occurrence data for the United States, 1992-2011 [FPA_FOD_20130422]. Fort Collins, CO: USDA Forest Service, Rocky Mountain Research Station. http://dx.doi.org/10.2737/RDS-2013-0009

California Drought History



The introductory paragraphs and table which precede this dashboard explain the Palmer Drought Severity Index (PDSI), a precipitation-based index for measuring drought. This line and bar chart visualizes statewide changes in precipitation and PDSI over time. The cyclical pattern of the rainy season is contrasted against severe wet and dry spells. Annotations and highlighting are used to point out droughts and heavy rainfall events. Hovering over any particular part of these graphs shows a tooltip with the time period and the exact measure (inches or PDSI) graphed.



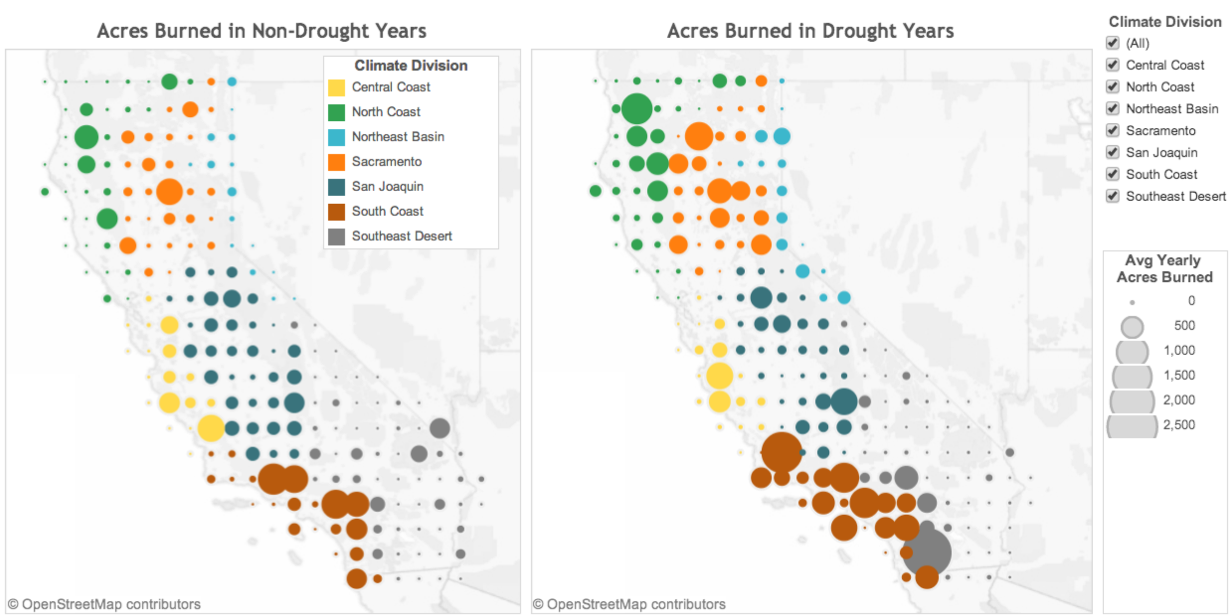
These grid maps show regional variation in precipitation and drought. Users may click through years on the right to see how precipitation and drought have changes during the period of 1993 through 2011. This dashboard is meant to allow users to explore the relationship between these two variables and the regional differences between how precipitation affects PDSI. In general terms, these maps are meant for users to discover that Northern California requires more rain than Southern California in order to be in the normal PDSI range. In addition, they should see that Southern California has been affected by more drought during the time period of 1993-2011.

Testing earlier versions of these dashboards helped us refine our use of color. For example, blue colors were originally used for the precipitation grid map. We found that users confused the precipitation blues with the high PDSI blues, so we used green for precipitation instead. For the line and bar charts, using blue for precipitation line actually reinforced the statewide relationship between heavy rainfall and increased PDSI.

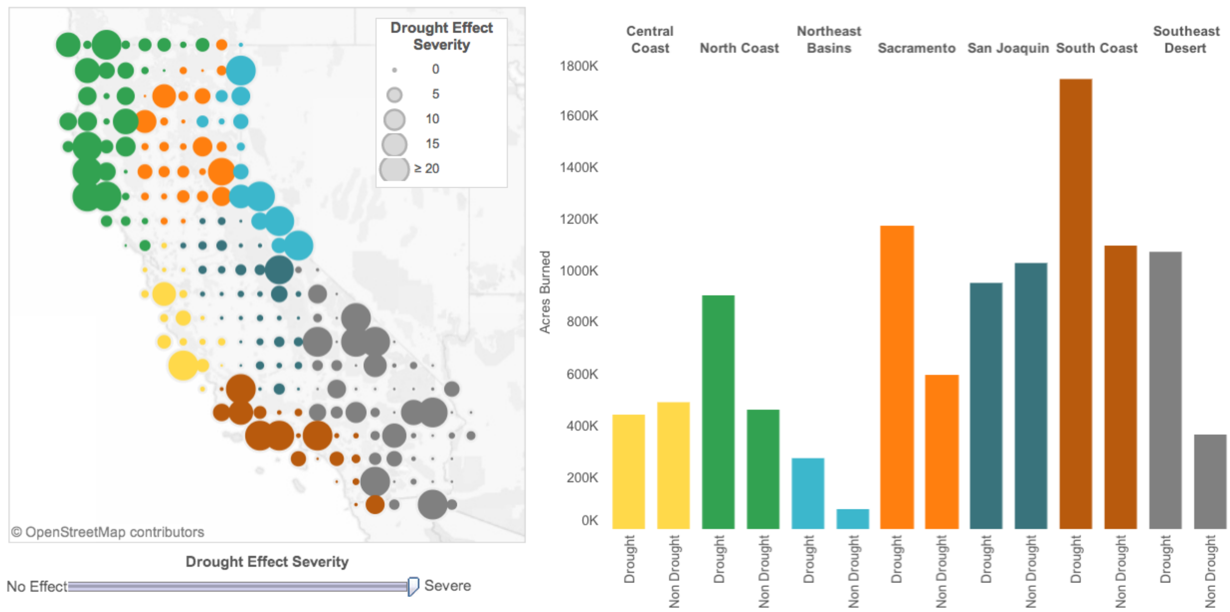
Overall, users expressed appreciation for the ease with which they could visually understand the complex relationship between precipitation and drought. They were able to take away from it what we hoped: namely, a realization that the relationship between precipitation and drought is not 1:1. Users also noted an appreciation for the highlighting of drought years in the first dashboard.

Wildfire and Drought: Patterns and Relationships

The following two dashboards illustrate the main findings of our project—that there is an interesting relationship between drought and fire size that is not equal across the state of California. In both dashboards, color is used to differentiate the climate regions.

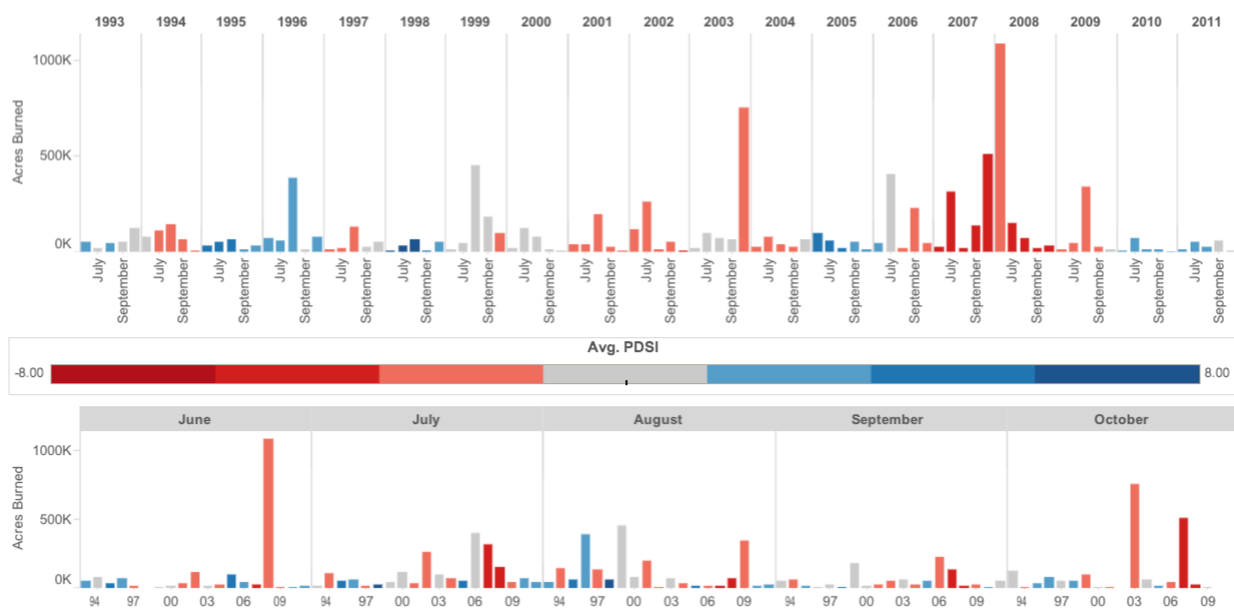


This dashboard provides the user with an opportunity to explore the differences in acres burned between months of drought, and months of non-drought. The size of a circle reflects how many acres were burned near that grid point. The regional color differentiation also makes it easier to identify grid points on each map, which helps the user make comparisons between the maps and notice differences between grid points. This dashboard is intended to pique the users' interest and help them begin to see a relationship between wildfire and drought. Based on user feedback, we switched the left/right order of the charts to enhance the logical flow of the exploration.



This dashboard helps to clarify the connection between wildfire and drought. The grid points are sized according to the ratio of acres burned in drought periods to acres burned in non-drought periods—a measurement of the effect of drought severity on fire size. The larger circles show where fize size is most influenced by drought severity. The bar chart on the right show the actual acres burned in each region in drought and non-drought periods, to help show users additional data, for instance, helping indicate that while the influence of drought on the Central Coast (yellow) is large at two points, that the overall difference for that region is small in absolute terms (seen by comparing both yellow bars on the right).

Comparisons between the map and bar chart show interesting relationships. For example, while both the Northeast Basins region and the South Coast region seem to have fires that are strongly influence by drought, many more acres have been burned in the South Coast region. In addition, the San Joaquin region also had many acres burned, but it drought doesn't seem to have a strong effect on fire there. This may have something to do with much of the San Joaquin Valley being irrigated farmland.



This final dashboard helps to show that the increase in acreage burned by wildfires is driven in large part by a few large fires that occur during fire season (June-October). During drought periods, there have been some exceptionally bad months of wildfire, which generally consist of a few very large fires across the state. This the behind the trending in the fire/drought relationship in the 1993-2011 period.

Users expressed difficulty in realizing that the time scale was not a normal one for this dashboard and that the lower bar chart concerned a specific month over different years. We made some formatting changes to diminish this confusion, although this is perhaps an area where we could further improve the narrative and visualization.

Process

Our processed followed four main steps: initial proposal and research, data identification and analysis, data visualization, and refinement and enhancement. However, particularly with three team members working both independently and collaboratively, these tracks overlapped at multiple points.

Initial Proposal and Research

We began by researching existing visualizations that fell into two categories: exploration and analysis of either wildfire or drought data, and visualizations that combined geographic, temporal and attribute-based analysis. We focused on the former to baseline the project with what current scholars in the field were using and to identify weaknesses, flaws and possible enhancements. We studied the latter to gain inspiration for how to create our own design. Particularly as we knew we would be primarily working with geographic data, we were concerned with how to stay away from traditional choropleth maps. We were particularly interested in adapting the chart and map approach used by Mike Bostock and others for the New York Times drought and crop analysis.

Data Identification and Analysis

In order to visualize wildfires by the number that occurred in the area, exceptional effort was put into finding a visual method to aggregate wildfires together by location without falling back on a traditional choropleth or heatmap.

One of our team members had previously seen a nearest neighbor visualization using a normal kernel method, and the team began exploring ways to translate the data into a format that would allow this type of visualization.

This particular format for visualizing data is very effective for conveying accurate, easily understood takeaways with minimal effort: by utilizing an even grid, and relying on a single shape that varies in size, two salient preattentive features are played to somewhat cunningly. The regularity lets user pick out patterns quickly - any chance is noticed with ease - and while circles' areas are compared with somewhat more difficulty as far as shapes go, placing them across evenly spaced center points helps counterbalance that, allowing for quick and effortless comparisons. The decision to color regions differently was made so that it would be easier for the user to visually compare very differently sized regions in terms of fire distribution as well as area. Whereas chloropleth's make it difficult to compare the areas of different geographic resions to one another, using the type of grid we decided upon with a grid kernel makes it easy to count how many grid points of the same color there are in each (if yellow only has 6 points in it, and grey has 24 points in it, it helps quickly convey an idea of scale of the two areas).

Our team had experience using MatLab for manipulating geocoordinate data and as a result all data (wildfire and drought) was ported into MatLab. Several custom functions were imported to enable the matching of Fire to Drought data. This involved running a script that looked at each point of wildfire ignition, assessed which climate region it fit into (if point in polygon, then->) and assigning the correct PDSI and precipitation values to that fire record so that fires could be addressed individually relative to their drought condition. This was applied to approximately all 154,000 fire records. Once this had been done, work began on generating a grid for the creation of the normal kernel visualization.

This is how the normal kernel was generated: a grid pattern was created at varying density distributions to determine which would be visually pleasing, and still convey a sense of pattern while remaining truthful to the underlying data. We finally decided on a grid spacing of 0.25 degrees of Latitude/Longitude. Once these grid points were generated, a nearest neighbor algorithm was written and run to assign each fire to its nearest grid point. Finally, the unique ID of each grid point was appended to the fire record for ease, so that it could easily be addressed at a later time. Custom functions used for this included point2polygon, csv2struct, struct2csv and p_poly_dist. This process, from data import to export took approximately 30 hours of coding in MatLab.

Now that the data was in a more manageable format, we began to ask how to draw out the relationship between drought and fire. Tableau served as a very useful tool to direct our attention to areas to explore further. This led us to deciding to utilize the ratio of total acres burned at a given grid point in-drought, and not-in-drought as the data on which to base our final visualization showing the relationship of drought to wildfires.

Data Visualization

We decided to use Tableau for visualizing the data rather than D3 as originally planned. Tableau not only allowed us to explore the data more easily, but we found that it was much faster for building and iterating on the visualizations that best showed the data. We had planned on building a tool that allowed users to explore the data on their own, and we realized that multiple Tableau dashboards would be an effective approach for developing a narrative about the data. Throughout the project we had identified a list of questions that we wanted users to be able to answer, and we built individual charts and linked dashboards one by one that would answer these questions. By placing these dashboards on a long scrolling website, we were able to build a kind of interactive infographic that supported the narrative we wanted users to explore.

Refinement and Enhancement

As we built visualizations that answered questions from our initial list we continuously updated that list with new questions that arose from the data. As such, we built our visualizations and overall narrative based on a combination of our initial research goals and on the patterns in the data itself.

We shared our page with a number of colleagues and classmates throughout its development. Adjustments ranged from removing entire graphics and charts to editing small features of the Tableau interface. Our in-group critiques were particularly effective, allowing us to work on any specific visualization and then bring it to the rest of the team for improvements.

Results

We interviewed users about the interface, features and overall effectiveness of our visualization, and categorized the results into a few key takeaways. We were able to respond to some feedback immediately and alter the visualization. Others we chose to accept as limitations of the approach or the technical design.

Software

There was some confusion about the use of Tableau as a tool, and how the dashboards were being presented. However, we generally found that users liked the consistent styling of the visualizations. There were a few specific challenges; one user discovered the “Keep Only” and “Exclude” buttons that appear on tooltips by default. She was baffled, noting “What’s the point? Why would I ever use those?” We realized that there was no value within our visualization to keep those features, and removed them. It was a good reminder to pay attention to details like default settings, and that stripping out unnecessary features is important (especially when using a tool like Tableau, rather than starting from scratch with your own code).

Purpose

There were diverse reactions to the purpose of the website and the visualizations. We found that during the open house, in which users were expecting us to walk them through our project website, people read very little of the supporting text. They jumped immediately to the visualizations and often wanted specific answers to questions that they came up with from the data. In contrast, people who came to the website just from the link started by reading the introductory text and had an understanding of the website as an exploratory tool. It was effective in getting them to become more interested in the topic, but sometimes left them frustrated that there weren't more clear answers. One user said, "It definitely makes you hungry for more information/explanation." This can be understood as both a failure (you don't want to upset readers/users) and a success (because it meant that users understood and were engaged with the information we present).

Methods

As we had focused primarily on the visualization component, there were many questions on methods from users who had backgrounds in environmental science or other fields related to wildfires and drought. We considered adding more discussion of the data analysis and methods to our website, but decided it would detract from the focus of users on exploring the data. Instead we included an appendix that had our data sources, linked to our Github repo, as well as to the I School where people could find and contact us.

Team Work

	Xavier	Peter	Cameron
Initial Proposal and Research	33%	33%	33%
Data Identification and Analysis	75%	10%	15%
Data Visualization	15%	45%	40%
Refinement and Enhancement	30%	35%	35%

Appendix

Software and Tools Used

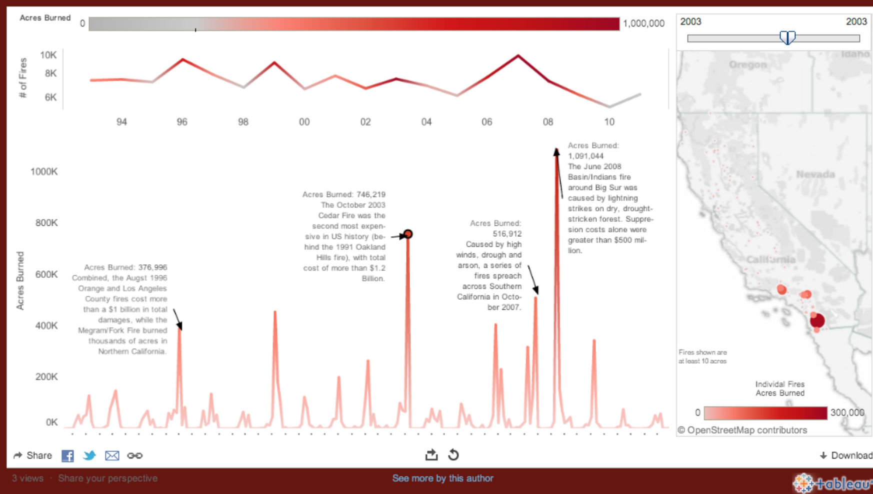
Tableau Version 8.1
MatLab R2013b
Cool Kitten 1.0

Github Repository

https://github.com/cameronreed/fire_and_drought

Screenshots

The full website



California Drought History

Measuring Drought

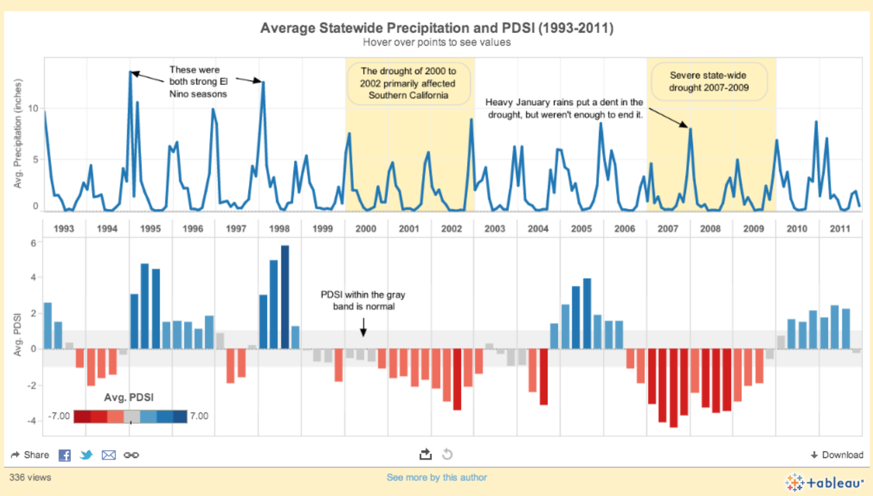
Drought is typically associated with a lack of rain, but it's a little more complicated than that. "Normal" precipitation differs among regions of the United States, so the country is divided into 344 regions for climate monitoring.

The Palmer Drought Severity Index (PDSI) was developed for comparing drought conditions between different regions. PDSI is based on precipitation, with positive values indicating wet spells and negative values indicating dry spells.

The charts below show annual precipitation and PDSI averaged over the entire state for the period of 1993 through 2011. California's rainy season extends from November to March, during which time the state typically receives 75% of its annual precipitation. Notice in the charts below how years with low rainfall coincide with droughts and negative PDSI.

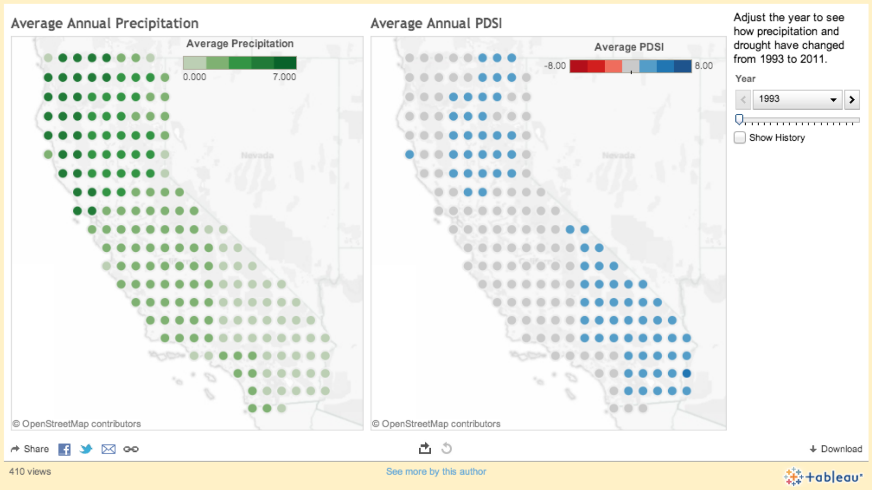
Palmer Drought Severity Index

Condition	PDSI Range
Extreme wet spell	Greater than 4.0
Severe wet spell	3.0 to 3.9
Mild to Moderate wet spell	1.0 to 2.9
Near Normal	-0.9 to 0.9
Mild to Moderate drought	-2.9 to -1.0
Severe drought	-3.9 to -3.0
Extreme drought	Less than -4.0



Regional Patterns

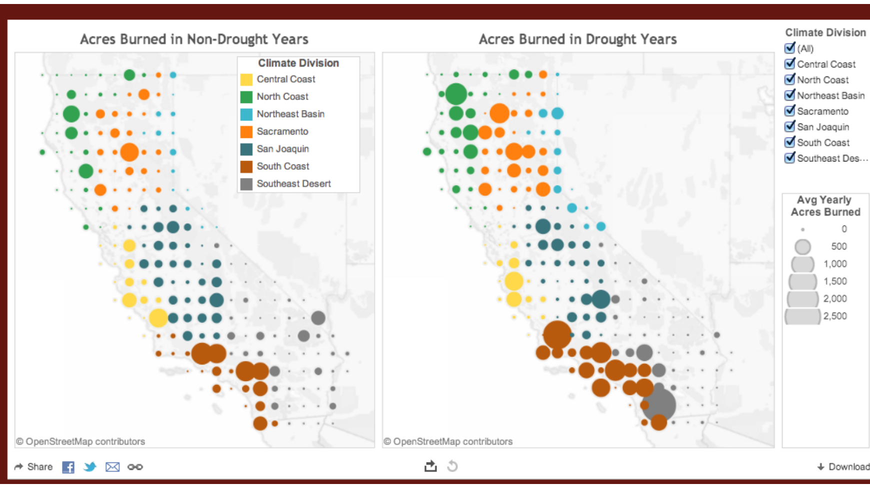
These grid maps show how average precipitation and PDSI change year to year. Use the toolbar on the right to move through the years and see how precipitation levels affect drought differently across the state's climate regions.



Wildfire and Drought:

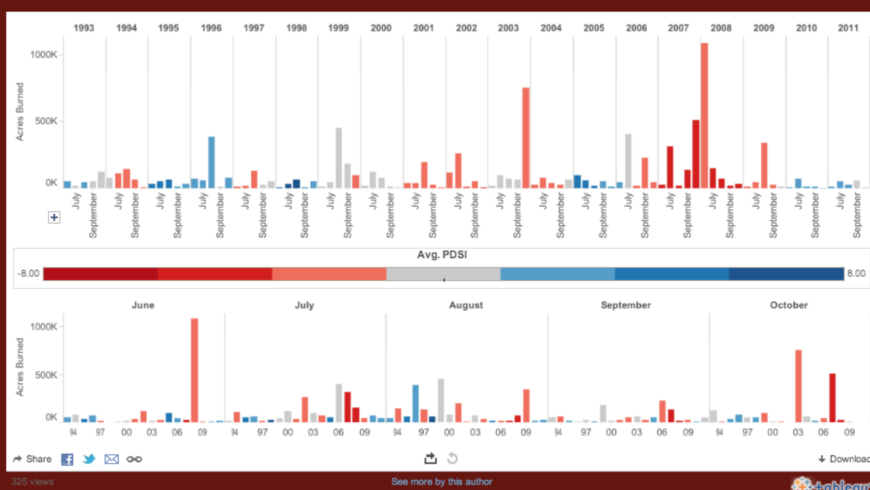
Patterns and Relationships

Applying the same grid pattern to wildfires yielded number of acres burned during drought and non-drought periods for each node. While wildfires are distributed differently across the distinct climate divisions of California, both Northern and Southern California showed significant overall increases in acres burned during drought periods.



The chart below shows ratio of acres burned in drought periods to acres burned in non-drought periods. More acres were burned during drought periods in every region except the Central Coast and San Joaquin.

Investigating just the California wildfire season, June through October, helps clarify the relationship between drought and fire. All major fires and fire seasons since August 1996 have occured during drought years. Additionally, while July and August have more constently high acres burned across the state, the start and end of fire season in June and October have seen the largest fires.



Conclusion

The relationship between drought and wildfire in California is incredibly complex, and relates to diverse inputs ranging across climate change, irrigation patterns, wildfire management strategies and even individual cases of arson. However, our analysis suggests a significant pattern of increasing drought and increasing acres burned by wildfires across much of the state.

Neither drought nor wildfire are problems that will be resolved in the near term. As Californians continue to combat and adapt to their changing environment, there are no easy answers. However, a better understanding of how natural disasters like drought and wildfire interact can help citizens, administrators and politicans understand the hard choices that need to be made in how best to manage these challenges.

Credits

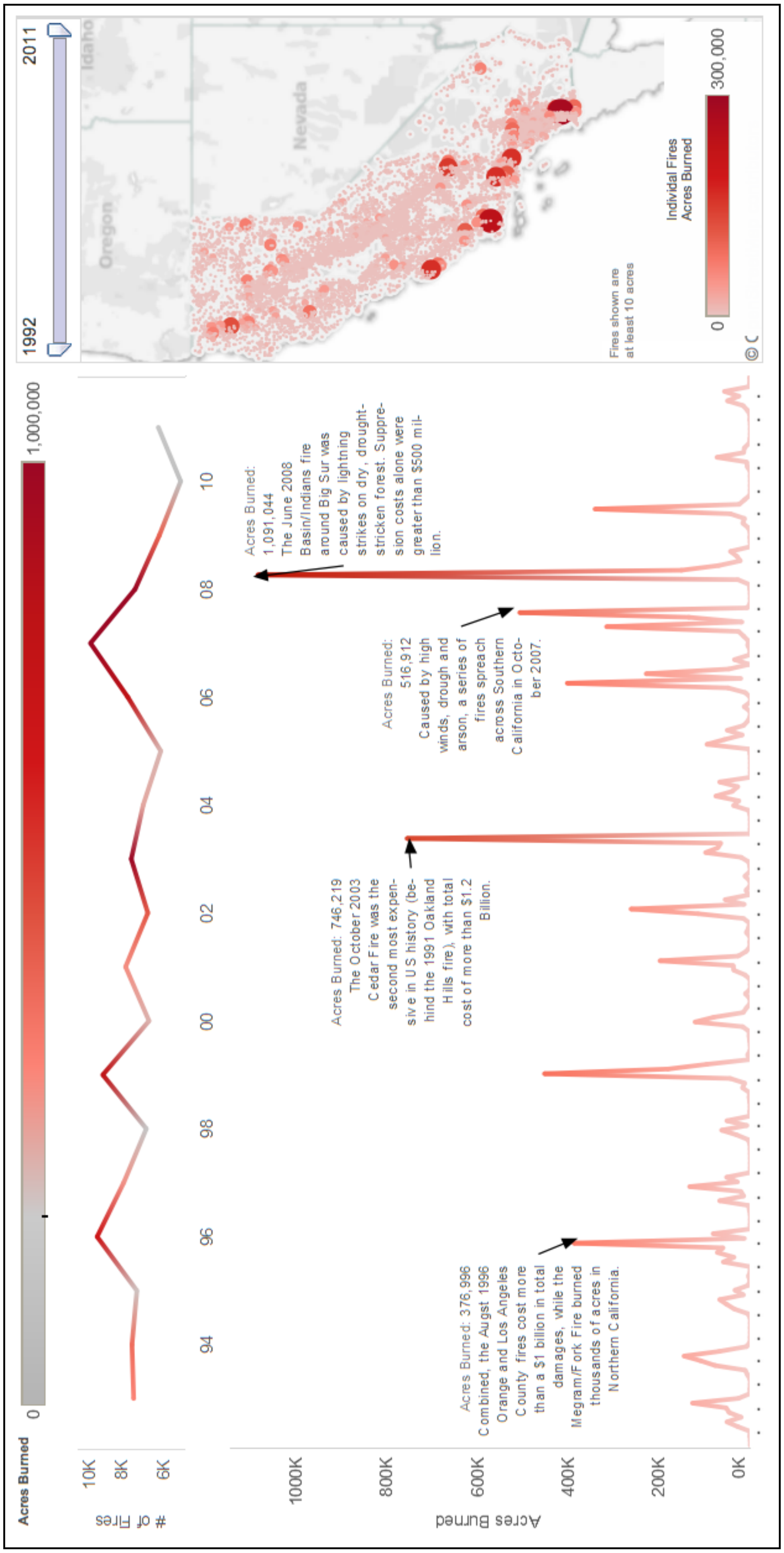
This website was created by Cameron Reed, Pete Swigert, and Xavier Malina as the final project for Information Visualization, a course at the UC Berkeley School of Information.

The visualizations were created with Tableau and are based on the US Forest Service wildfire database, originally accessed through the National Park Service, and historical drought datasets accessed the National Climatic Data Center. The website was designed using Cool Kitten, a framework by Jalxob.

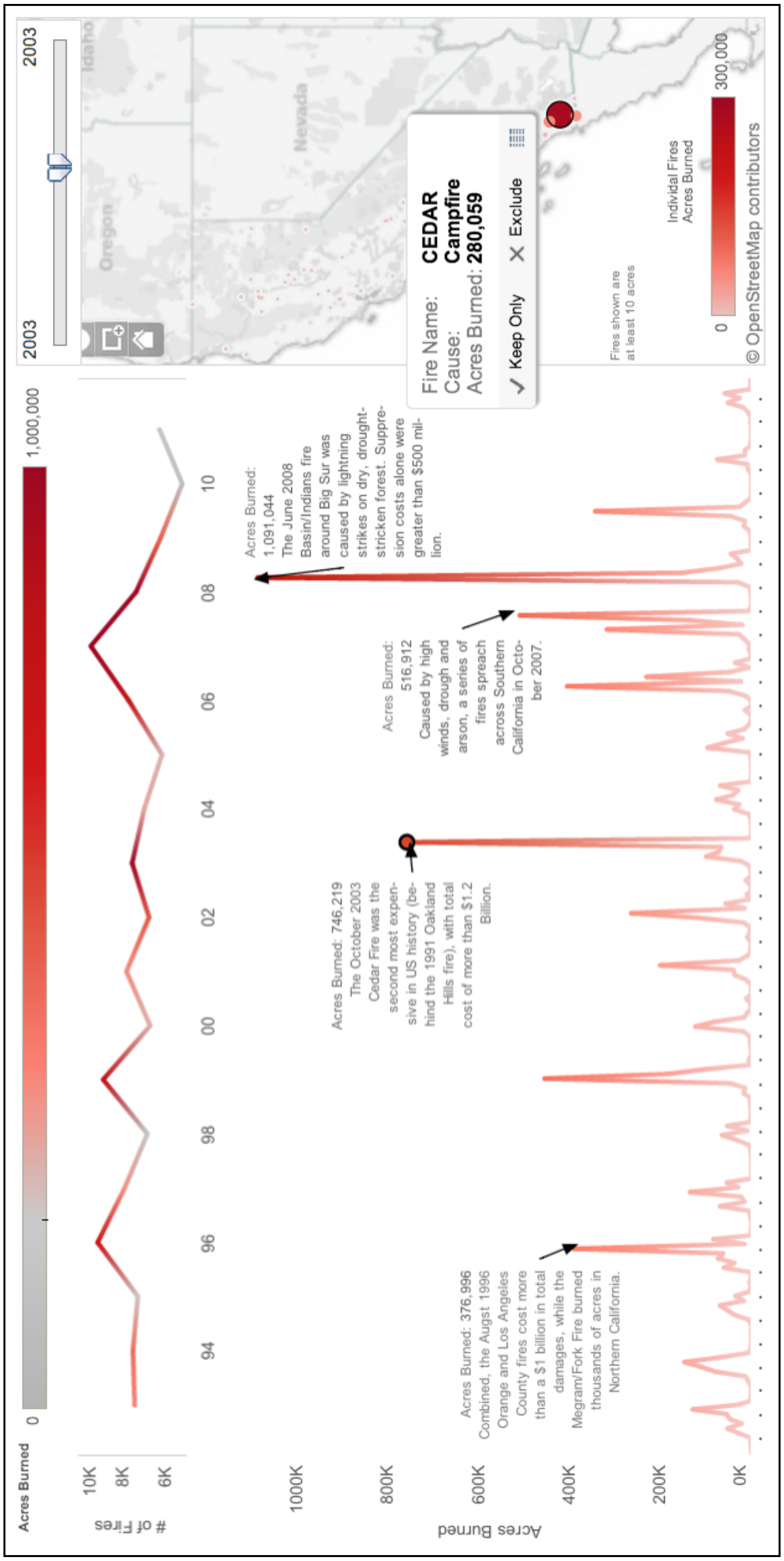
Additional drought information sources: El Niño years, PDSI Explanation, and California Drought.

See project code on GitHub.

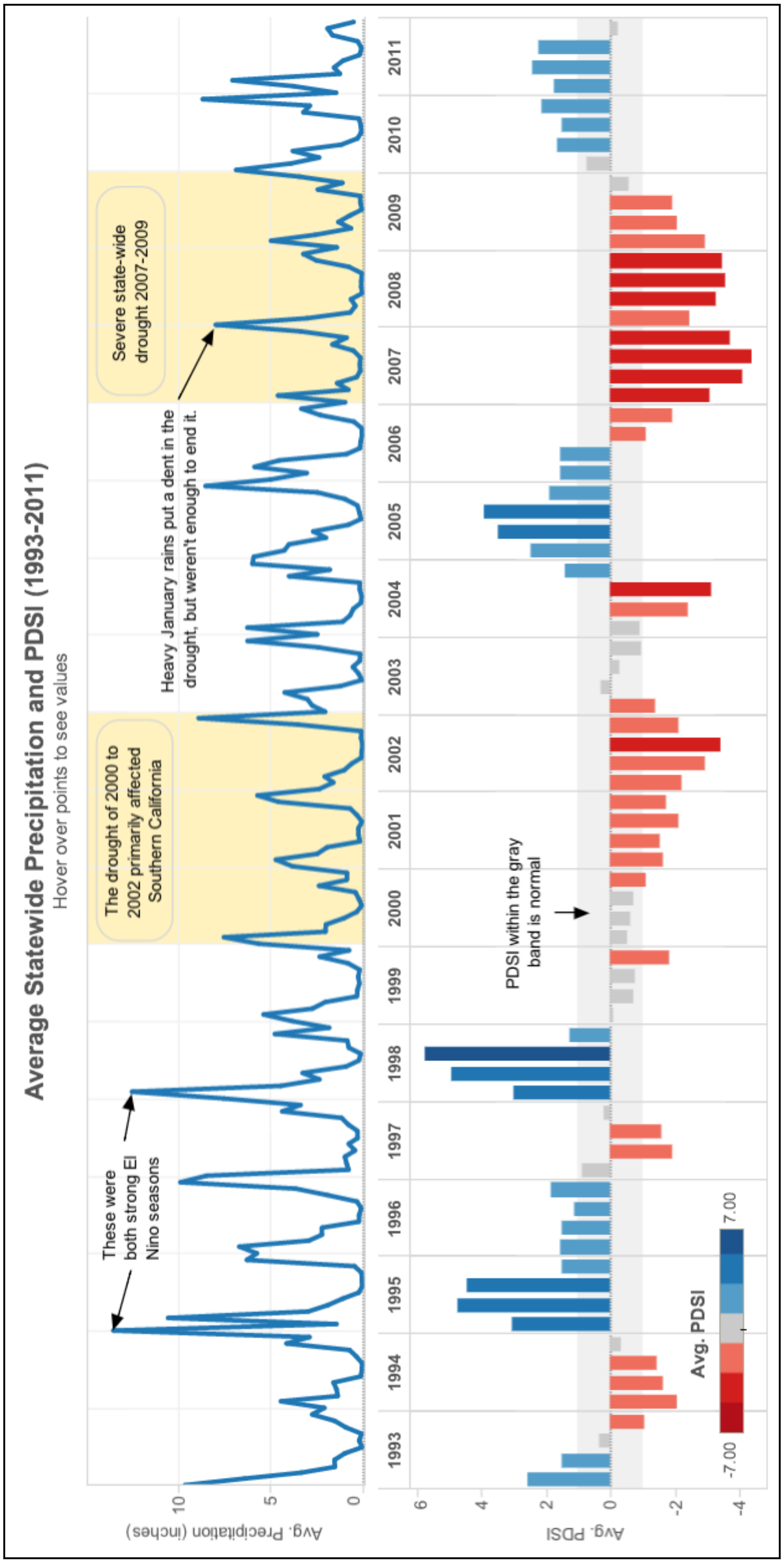
Dashboard 1A: California Wildfire Over Time (1993-2011)



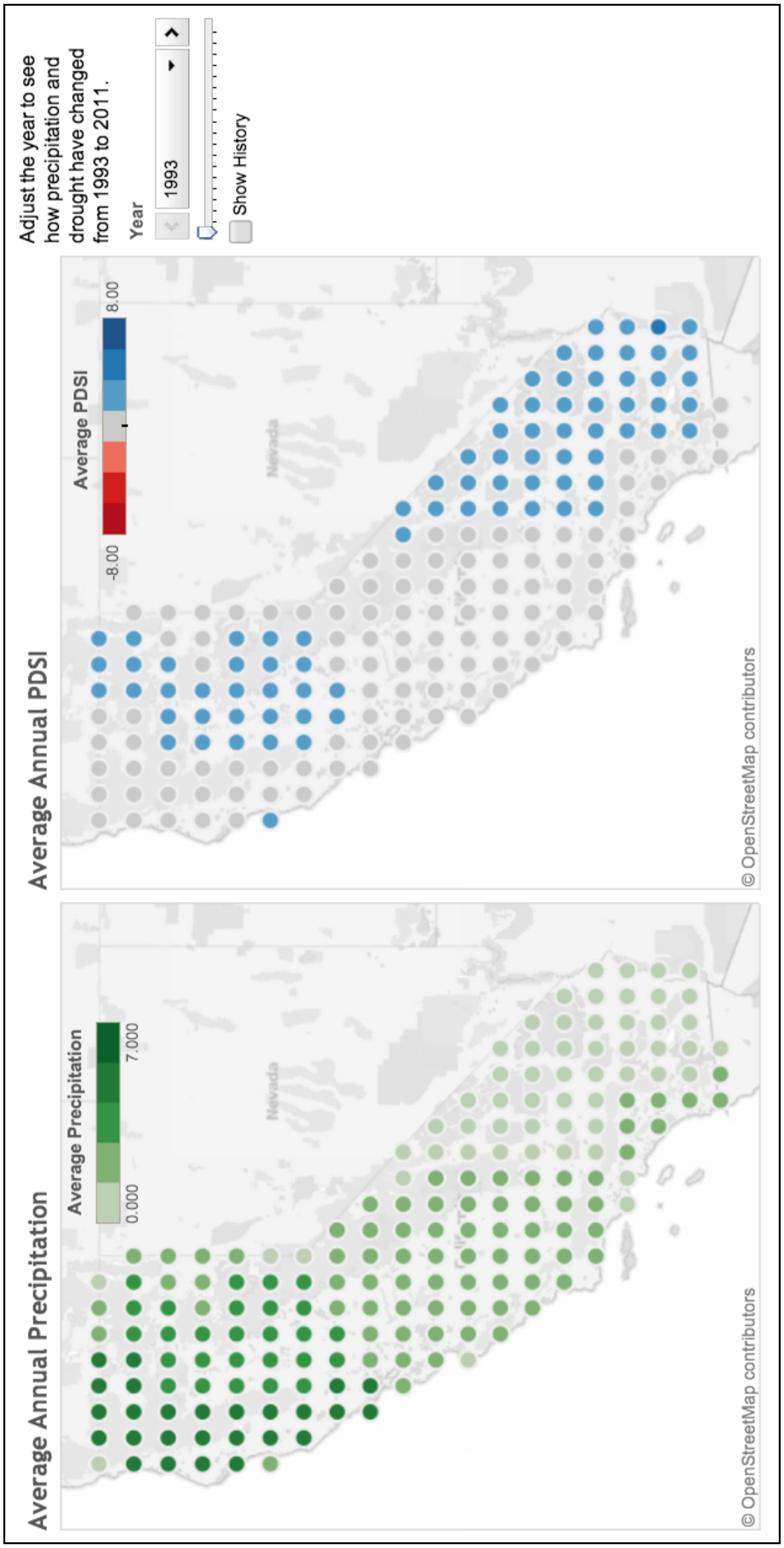
Dashboard 1B: California Wildfire Over Time (1993-2011)
Highlighting a month and hovering over a fire



Dashboard 2: Average Statewide Precipitation and PDSI (1993-2011)

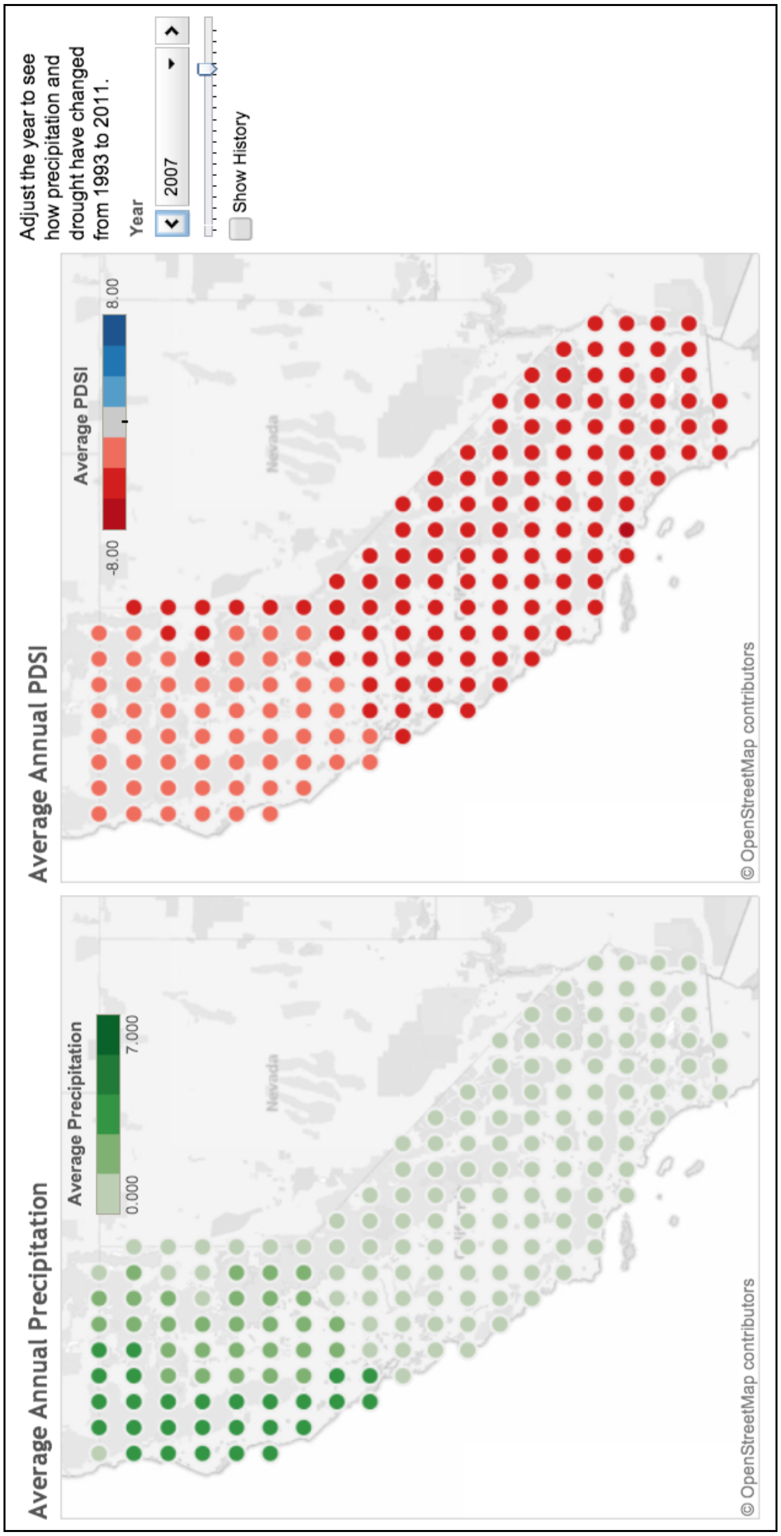


Dashboard 3A: Average Annual Precipitation and PDSI

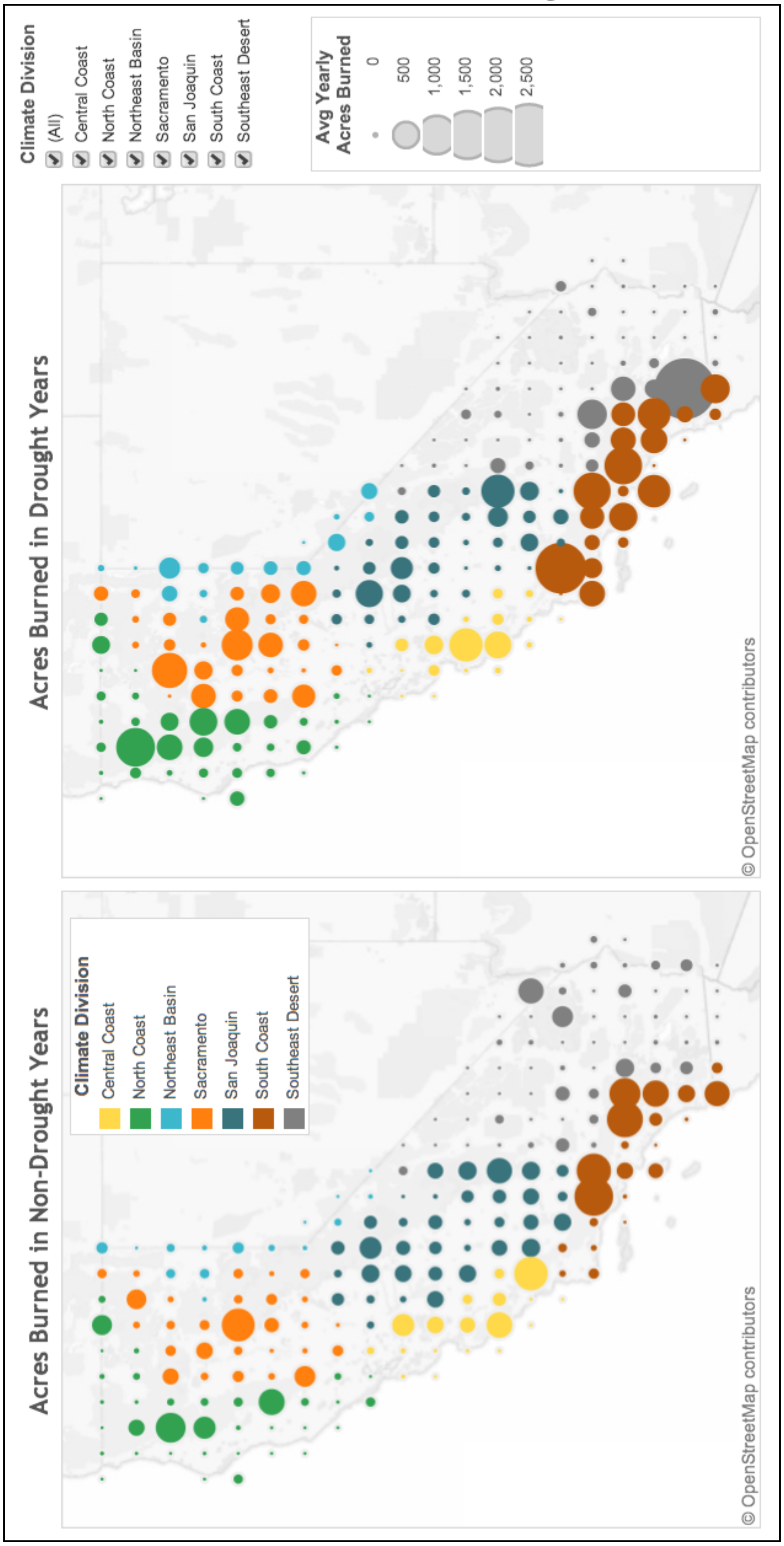


Dashboard 3A: Average Annual Precipitation and PDSI

Clicked through to another year

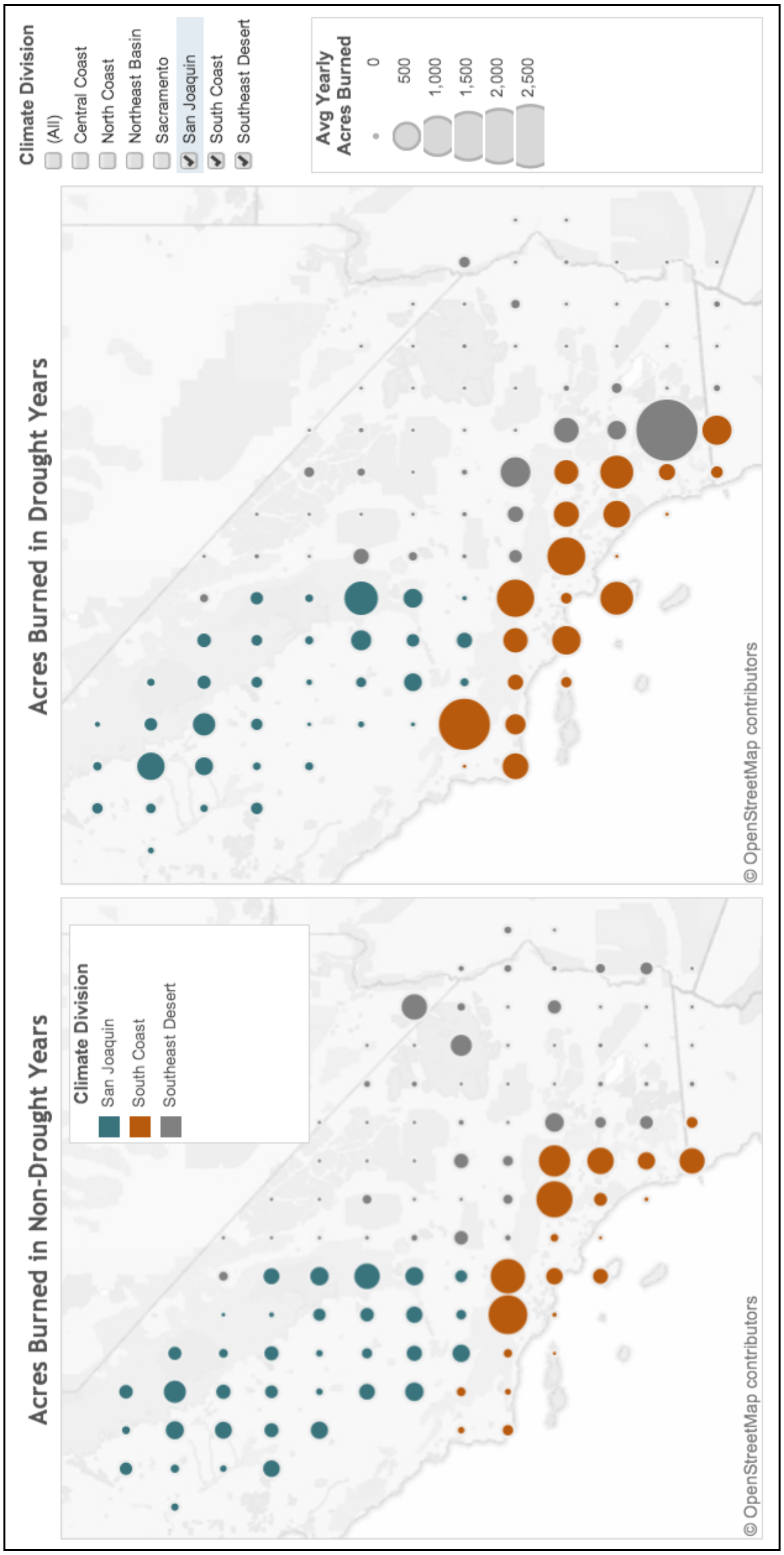


Dashboard 4A: Acres Burned in Drought Years and Non-Drought Years

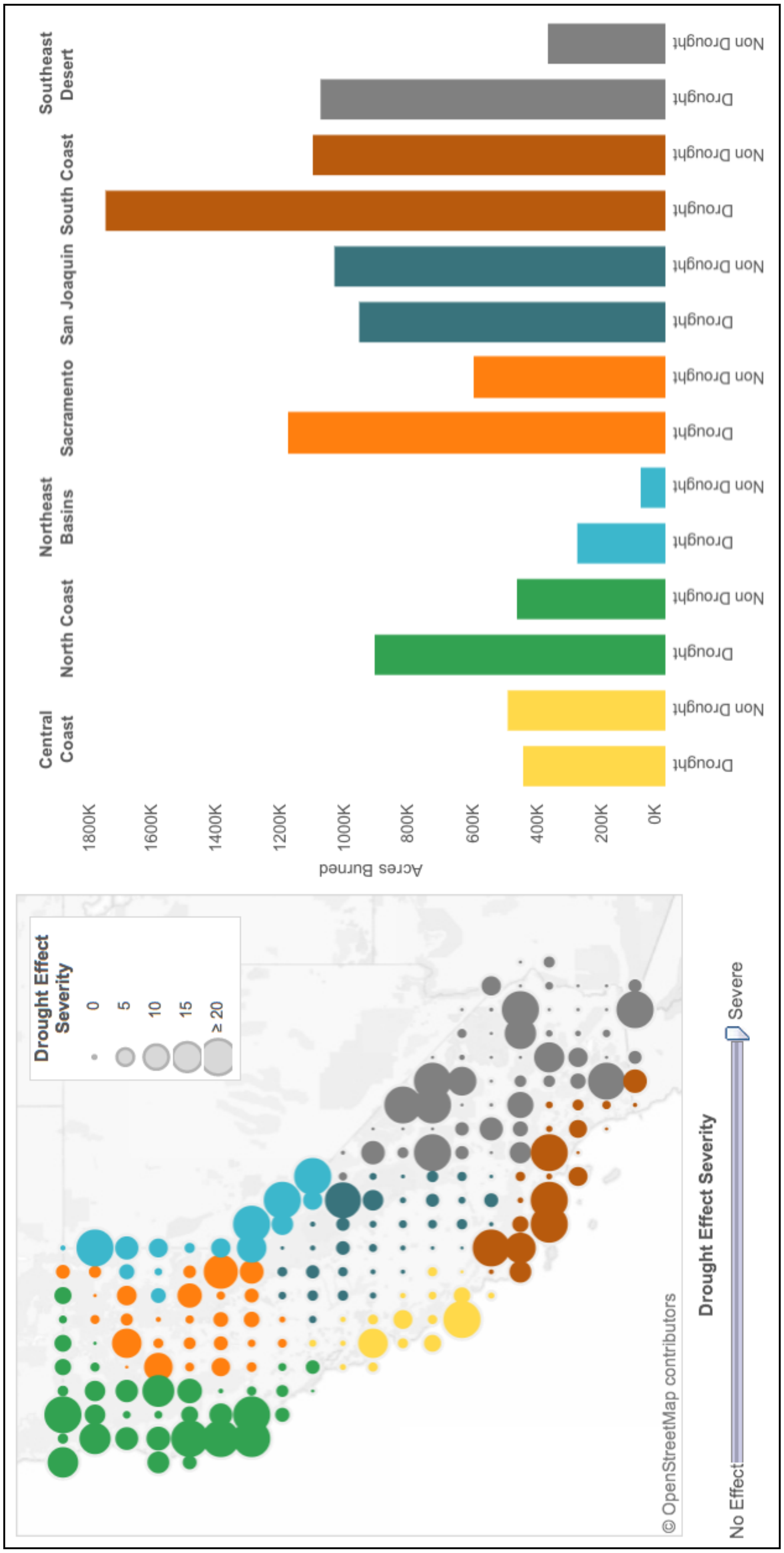


Dashboard 4B: Acres Burned in Drought Years and Non-Drought Years

Filter view on Southern California



Dashboard 5: Ratio of Acres Burned in Drought Years vs Non-Drought Years



Dashboard 6: Acres Burned in Fire Seasons (1993-2011)