

INFO 247 Final Project Report

20 Years of Climate Change in California

Wadzanai Makomva and Soorya Narayan Satheesh | May 8, 2023

Table of Contents	1
Project Goals	2
Related Work	3
Visualization Description	10
California Counties Temperature Trends	10
Anomaly Temperature Heat Map	11
Average Precipitation Over Time	12
Precipitation Heat Map	13
Approach	19
Temperature and Precipitation Dataset	19
Tools	21
Steps	21
User Testing and Results	22
Introduction	22
Method	22
Participants	23
Scenarios/Tasks	23
Screenshots	24
Test Measures & Results	28
Discussion	29
Link to Code and Visualization	30
Division of Labor	30
Appendix	31
Tableau Workbooks	31
Usability Test Script	31
Feedback Question Responses	33
References	34

Project Goals

Spurred on by an experience by one of our team members who was a first respondent to the scene of a wildfire in India, the purpose of this project was to provide knowledge and understanding to users regarding key indicators of climate change and changes occurring therein. For the purpose of this project we chose to focus on California and a limited number of key indicators of climate change. Our goal for this website was to educate users about the impacts of climate change and provide visitors with a clear understanding of how climate change is affecting the environment. Our second goal was to build awareness about the urgency of the issue by emphasizing the importance of climate change as an urgent global issue, not just in California. As students, we were also interested in exploring visualization techniques with a topic that was of interest to all team members.

Overall, it is our goal that our visualizations will accomplish the following:

- **Stir curiosity in the user:** We appreciate that the information provided along with our visualizations is highly summarized, however, we hope that this simplification of the relationship between temperature and precipitation changes, wildfire occurrences and climate change will prompt the users to seek more information on the matter and join the fight.
- **Bring awareness:** We believe that building awareness through education is an essential component of addressing climate change. An increase in awareness of this issue will undoubtedly lead to more support for policies and initiatives aimed at mitigating and adapting to climate change.

With these goals in mind we conducted our usability tests and tested for our participants' understanding before and after interacting with our visualizations.

Related Work

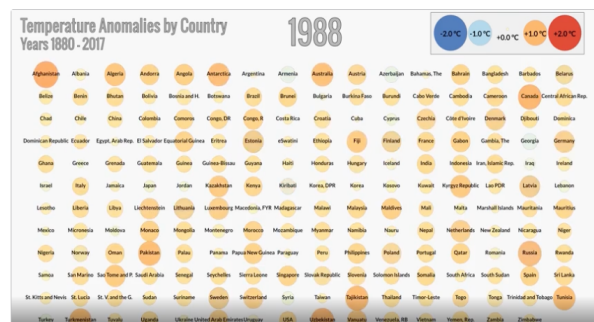
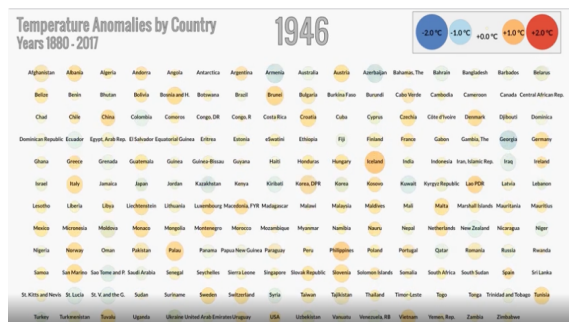
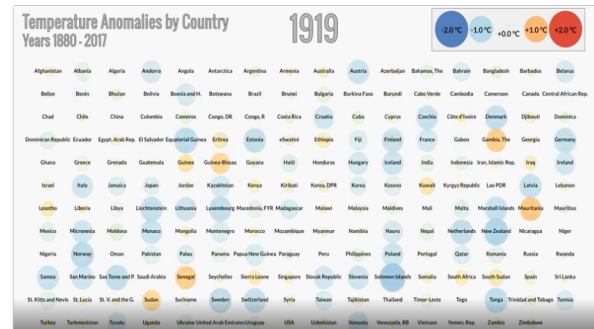
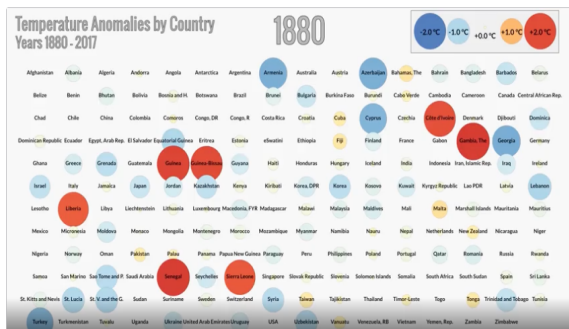
1. [Climate Change 'switchboard' visualization shows every country on the planet turning red-hot](#)

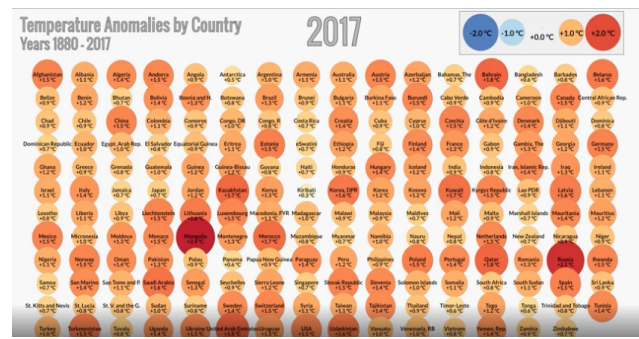
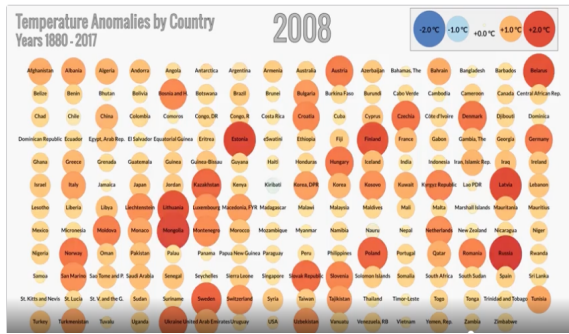
Both categories of visualizations, temperature and precipitation changes over the past 20 years were inspired by this New York Times article. The article discusses a new interactive map called the "Climate Change Switchboard" which shows the average temperature changes in each country over the past 100 years. The map depicts each country as a circle that changes color

based on the average temperature increase or decrease over the past century. The warmer the country has become, the redder the circle appears. Analysis of the map reveals that every country on Earth has experienced some level of warming over the past century, with some areas experiencing more significant increases in temperature than others. The article also highlights the importance of visualizations like the [Climate Change Switchboard](#) in raising awareness about the impacts of climate change and the urgent need to address the issue. The reaction that this visualization sparked in its users inspired us to attain the same goal. The Climate Change Switchboard was and is a powerful tool for visualizing the impacts of climate change on a global scale. Looking at the visualization inspires certain feelings of panic because it appears as if the world is really on fire.

The article notes that the warming trend shown by the map is consistent with the scientific consensus on climate change, which states that global temperatures have risen by around 1 degree Celsius since pre-industrial times, primarily due to human activities such as burning fossil fuels and deforestation. This is the same spirit in which we have approached this project with the same intent to show the user the changes of temperature and precipitation over a 20 year period.

Below are just 6 screenshots of different years within the 100 year period shown in the visualization by Antti Lipponen. The screenshots show the progression of average temperature increases over the world as the colors of circles turn from a blue to yellow to a light orange to red depicting an increase in average temperatures over the 100 years.



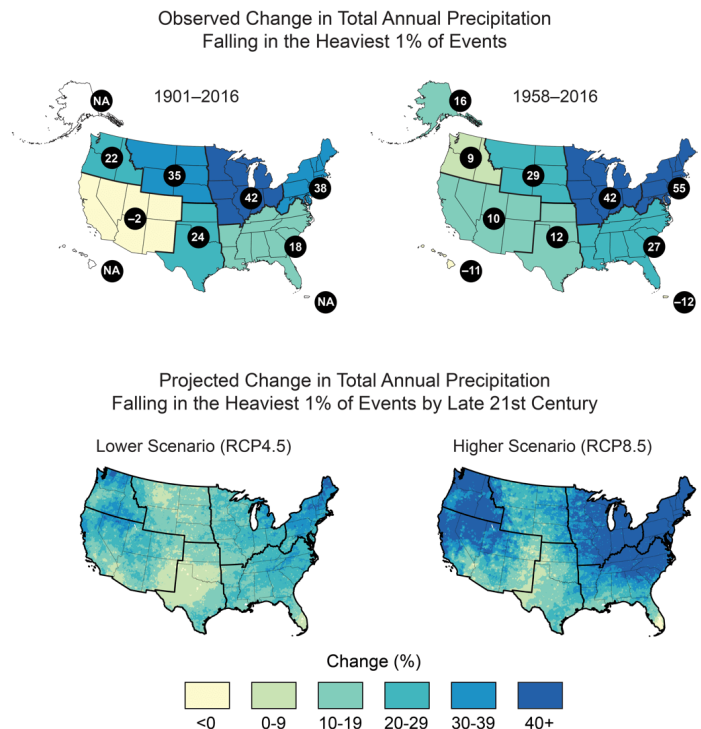


2. [Center for Climate and Energy Solutions: Extreme Precipitation and Climate Change](#)

This source provided us with an overview of how climate change is affecting precipitation patterns and increasing the frequency and severity of extreme precipitation events. We found some key scientific findings on the relationship between climate change and extreme precipitation, including:

- Global warming is causing more frequent and severe heat waves, droughts, and wildfires, which can exacerbate the impacts of extreme precipitation events.
- Human activities, such as deforestation and urbanization, can further intensify the impacts of extreme precipitation events by increasing runoff and flooding.
- Climate models project that extreme precipitation events will continue to become more frequent and intense in the future, leading to increased risk of flooding, landslides, and other impacts.

This resource was instrumental in inspiring the visualizations of choice for our final project. We observed the use of choropleth maps to show the change in the percentage of precipitation falling in the heaviest 1% of events for each state over time for both observed and projected events. The shades of color on the maps are used to indicate the magnitude of the change, with darker colors indicating larger increases in the percentage of precipitation falling in the heaviest 1% of events. The use of



choropleth maps of the United States allows viewers to identify which states have experienced or are projected to experience the smallest and largest changes in extreme precipitation events over time. The use of color shades also helps to draw attention to the areas where the changes have been the most significant.

3. [Fourth National Climate Assessment](#)

The National Climate Assessment report was instrumental in our team gaining a better understanding of the current state of changes in climate through its comprehensive analysis of the impacts of climate change on the United States. The report, which is a result of collaboration between the U.S. The Global Change Research Program and a team of more than 300 experts was used in the identification of the impacts of the key indicators of climate change on people and the environment.

Key findings from the report included the insight that climate change is affecting all regions of the United States, including California, and that the impacts are expected to continue to worsen over time. The report highlights the impacts of climate change on various sectors, such as health, infrastructure, agriculture, and ecosystems. The website also provided us with access to interactive tools and resources for exploring the data and findings. For the purposes of this project we explored the data and maps related to climate impacts, adaptation, and mitigation strategies, as well as downloaded data and graphics for use in our research and understanding. Overall, the NCA website was a comprehensive and accessible resource aiding in the understanding of the impacts of climate change on the United States and the urgent need for action to address the issue. We aimed to achieve the same goals with our own website that we created for this project.

☰	SUMMARY FINDINGS
	Background
	1. Communities
	2. Economy
	3. Interconnected Impacts
	4. Actions to Reduce Risks
	5. Water
	6. Health
	7. Indigenous Peoples
	8. Ecosystems & Services
	9. Agriculture
	10. Infrastructure
	11. Oceans & Coasts
	12. Tourism & Recreation

4. [Temporal Trends in Exceptional Fires](#)

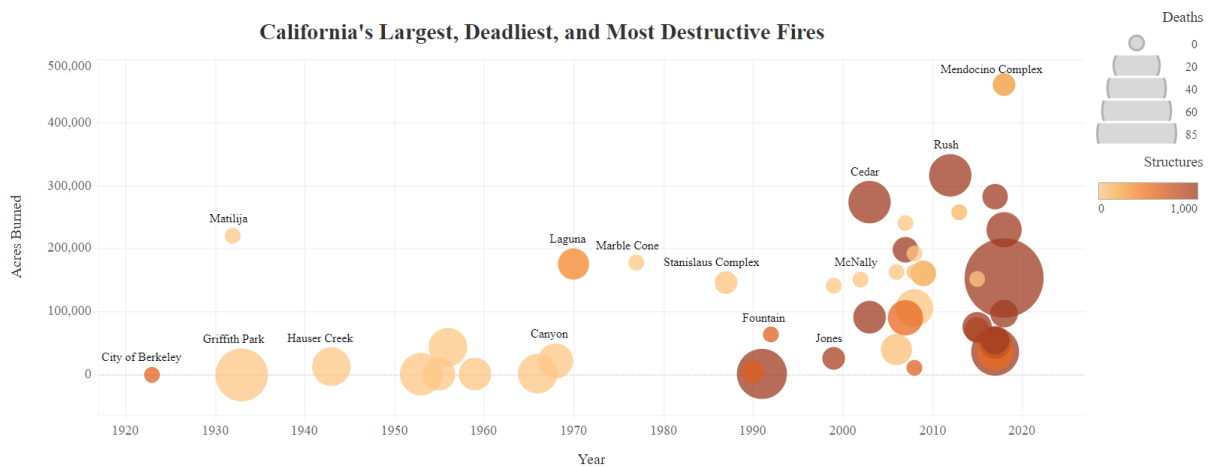
The California Wildfire Interactive Tool is an online resource that provides information on the relationship between wildfires, climate change, and public health in California. The tool was developed by PSE, in collaboration with the University of California, Berkeley School of Public Health and the University of California, San Francisco, and is based on the latest scientific research on the impacts of wildfires on air quality and health.

The tool allows users to explore the risks associated with wildfire smoke exposure in their area by entering their ZIP code or city name. It provides information on the estimated number of people who are at risk of exposure to wildfire smoke, the length of time that people are exposed to smoke, and the estimated health impacts of smoke exposure, including the number of asthma attacks, hospitalizations, and premature deaths.

The tool also provides recommendations for protecting health during wildfire events. These include staying indoors, using air filters, wearing masks, and avoiding outdoor physical activities. It also provides links to resources for more information on wildfire prevention, emergency preparedness, and community resilience.

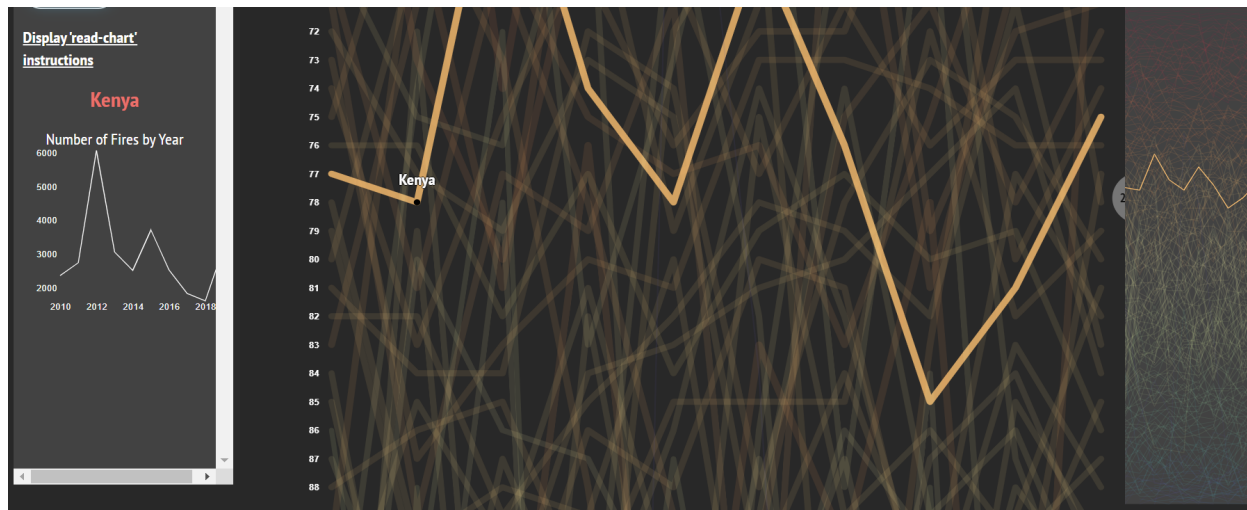
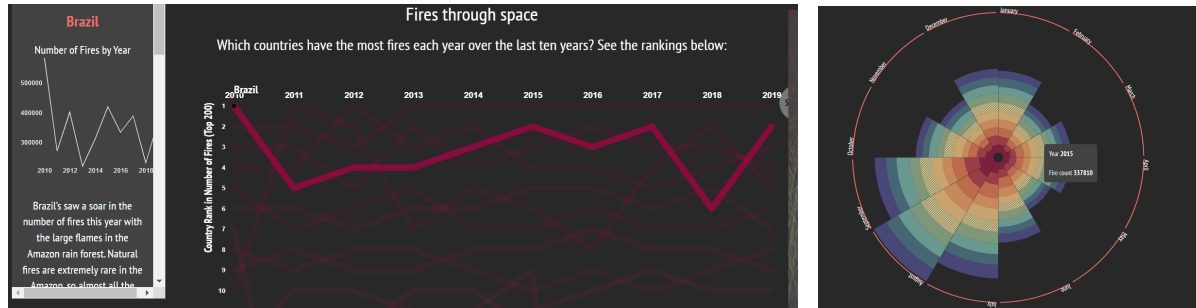
The California Wildfire Interactive Tool is intended to help individuals, health professionals, and policymakers make informed decisions about wildfire preparedness and response. It is free and publicly available on the PSE website.

Our project explores the impact of wildfires in California, and from this project, the visualization on “Temporal Trends in Exceptional Fires” inspired us to make the visualization to understand the relationship between wildfires, population densities and area and a county.



California's 20 largest fires by acreage burned, structures damaged or destroyed, and total fatalities according to *Top 20 Largest California Wildfires, Top 20 Most Destructive California Wildfires, and Top 20 Deadliest California Wildfires* (CAL FIRE 2019).

5. [Visualizing Fire Patterns 2010-2019](#)



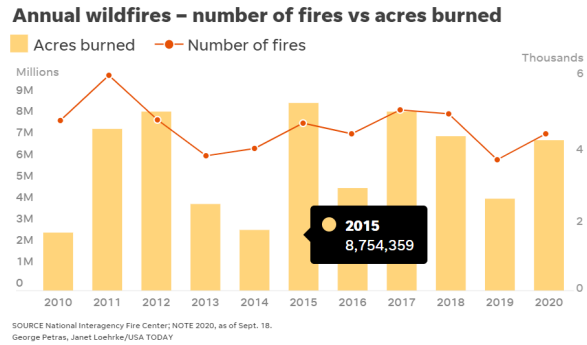
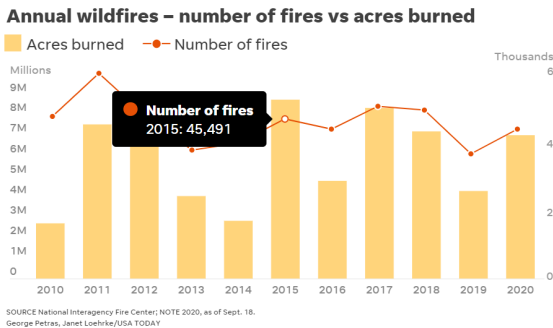
This project by Julia Janicki is to visualize the information on fires detected by NASA's MODIS Satellite and compiled at NASA's FIRMS (Fire Information For Resource Management System). The creator downloaded data in CSV format from NASA's FIRMS (Fire Information For Resource Management System), having information on all fires across the globe from 2010 to 2019. This includes information from all countries of the world. It has information on over 40 million fires detected. The creator used Python scripting for the categorization of data points by country and month.

Our project also obtained information on wildfires from NASA's MODIS Satellite and compiled at NASA's FIRMS (Fire Information For Resource Management System). Similarly, we also used Python coding for data cleaning and categorization.

This project ranked all countries based on the number of fires detected across the study period. We can see that the top-ranked countries were displayed at the top of the scroll page, while the countries with the lowest numbers of forest fires were displayed at the base of the scroll page. It

helps users in understanding the spatial variation in the occurrence of wildfires. We were inspired by this mode of stacking of different countries and chose to depict the top/bottom five counties of California with respect to the number of wildfires detected in 2000-2020 period.

6. [Wildfires burned millions of acres across the West. See what that looks like.](#)



This report by USAtoday.com covers the wildfires that ravaged the West Coast of the US in the summer of 2020. It has a large number of visualizations that impresses upon an uninformed viewer about the unimaginable scale of destruction that has occurred due to the wildfires in California, Oregon and Seattle.

The report pedantically enlightens the reader on the relative spatial dimension of a football ground then the New York’s Central park and slowly takes the user to imagine the magnitude of land that has been destroyed by forest fires. It helps a person from New York comprehend and empathize with the people of the West Coast much better. Wildfires are such a cruel calamity destroying communities, which needs empathy and support from all corners of the country to overcome.

Our project also highlights how wildfires have been affecting California from 2000 to 2020. The visualization depicting the annual wildfires inspired us to make the bar chart visualization depicting the California forest fires from 2000 to 2020.

Visualization Description

The purpose of the visualizations that we included in the final website was in line with our overall goals for this project, to aid the user in learning more about these indicators of climate change with the ultimate goal of educating and bringing awareness. We chose the following 3 categories of visualizations:

- Temperature Changes in California
- Precipitation Changes in California
- Wildfires in California

Overall, we utilized color, labels, annotations, and interactivity in the visualizations to effectively communicate the data and provide a clear understanding of the trends and patterns in California's climate change indicators over time.

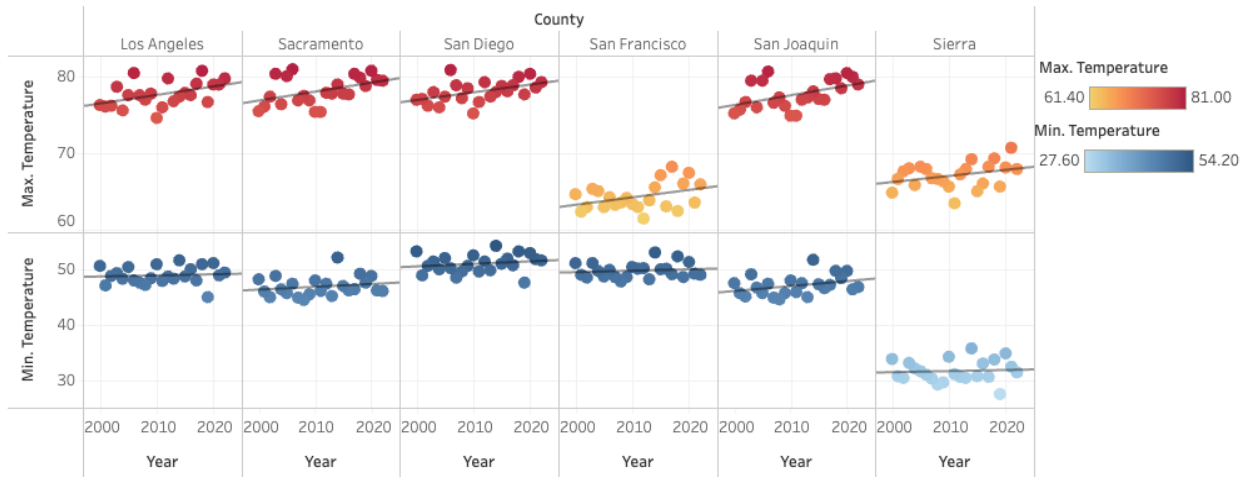
California Counties Temperature Trends

Although we created a number of visualizations using our temperature data, we could not utilize all of them as this would undoubtedly overwhelm the user. We chose the visualizations that we found to be most effective in achieving the above mentioned goals. The charts show the maximum and minimum temperature trends over time for different counties in California, providing the user with a way to compare temperature changes over time and understand how they vary by location.

Visual Techniques

From the charts the user is able to see that the scatter plot for the 6 California counties shows that temperatures have been consistently above the baseline over the last 20 years with some years showing particularly high temperature spikes. The visualizations within the dashboard use a range of techniques to convey information about minimum and maximum temperature changes in California and its counties over time. The scatterplot is a useful way to show how temperatures have changed over time in California and its counties. The use of a trend line allows the user to easily see trends and patterns in the data.

County Trend Charts



Anomaly Temperature Heat Map

The anomaly temperature heat map in the temperature dashboard provides a visual representation of the temperature changes across different counties in California over time, highlighting the specific insights of where the temperature changes have been most significant.

Some of the specific insights that can be gleaned from the anomaly temperature heat map for California include:

- The map shows that temperatures have been consistently above the historical baseline for the past 2 decades, with some regions experiencing particularly high temperature spikes.
- The temperature changes have been more significant in some regions of California than in others. The heat map highlights that some areas of the state, such as the Central Valley, have experienced more rapid temperature increases than others.
- The heat map also shows that temperature changes have become more frequent and more intense in recent years, with many areas of the state experiencing more frequent and longer heatwaves.

Visual Techniques

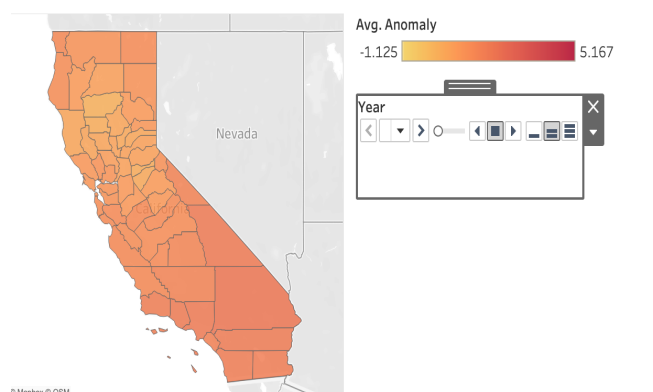
Color Encoding: The choropleth uses a color scale to represent the degree of temperature anomaly, with red indicating a higher-than-average temperature and lighter yellow indicating a lower-than-average temperature.

County boundaries: The map includes clear boundaries for each California county, which helps the viewer to easily identify the location of each temperature anomaly.

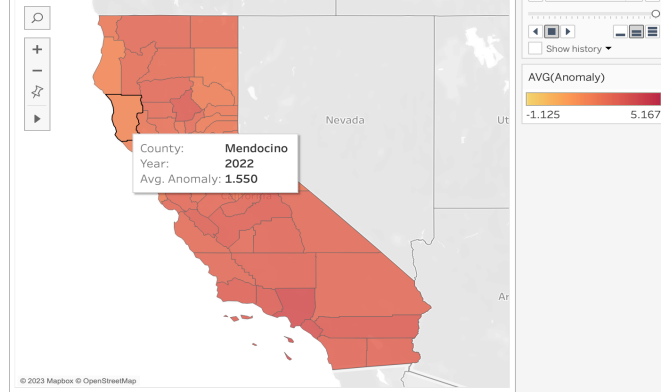
Interactivity: The map is interactive, which means that the viewer can hover over specific counties to see more detailed information about the temperature anomaly for that area. This technique helps to provide more specific insights and engage the viewer in the data.

- Time slider: We provided a time slider that allows the viewer to see how the temperature anomalies have changed over the 2 decades. We hoped that this technique would help to convey the overall trend of temperature change in California over time.

Anomaly Temp Heat Map - 2000



Anomaly Temp Heat Map - 2022



Average Precipitation Over Time

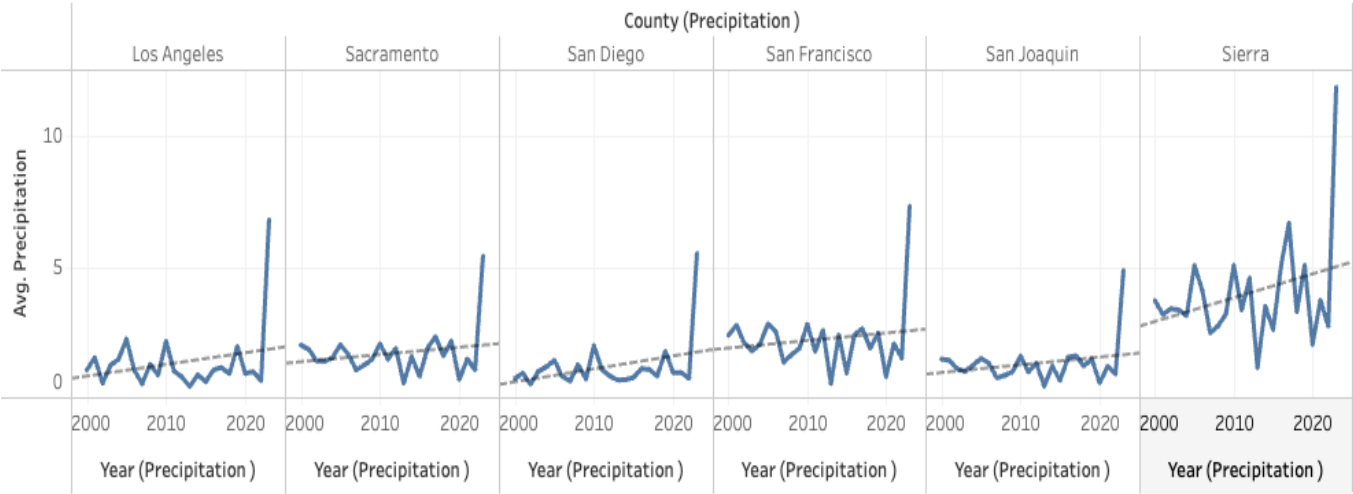
We chose to focus on 6 counties in California to show average precipitation over the past 20 years. We observe that these 6 states have experienced significant variability in precipitation levels over time. The line charts show that precipitation levels have varied widely over the past few decades, with some counties showing much higher or lower average levels than others.

Visual Techniques

We used the line charts to represent the annual average precipitation and blue to represent the actual data, while the gray line represents the trendline. This helps to differentiate the actual data from the trendline and makes it easier to follow the trend over time.

We used labels and annotations to provide context for the data. The y-axis is labeled with the units of measurement (inches) and includes a range that makes it easy to interpret the data. The x-axis is labeled with the years being represented in the data. Finally, the chart includes interactive elements that allow the user to hover over specific points in the chart to view the actual data for that year. Our goal was that interactivity would make it easier to explore the data and understand the specific changes in precipitation over time.

Average Precipitation



Precipitation Heat Map

The precipitation heat map visualization shows changes in precipitation levels in California counties over time. One observes that levels of precipitation across California have been highly variable over the past two decades. The heat map shows that precipitation levels have fluctuated widely with some counties and years showing much higher or lower precipitation levels than others. A study of the precipitation levels in the region by state shows that coastal areas tend to receive higher levels of precipitation, while inland areas tend to be drier.

Visual Techniques

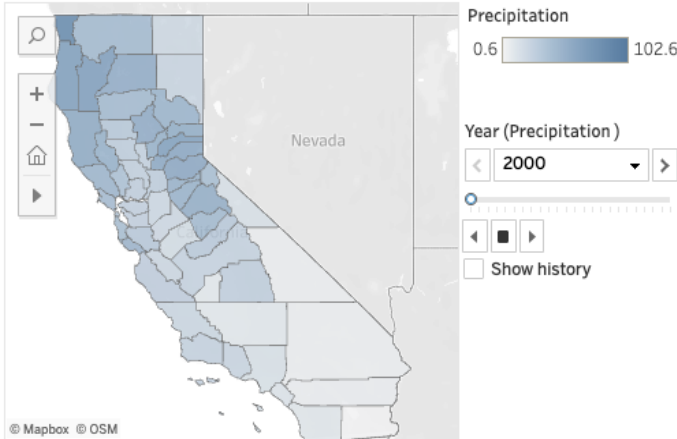
Color encoding: The heat map uses a visual encoding method to represent precipitation levels in the form of a color gradient from light blue to dark blue to represent lower to higher precipitation levels. The light-to-dark blue gradient helps users quickly identify these variations in precipitation.

Legend: We provided a legend on the right-hand side of the heat map to provide users with a clear understanding of the color encoding used in the map. The legend includes a color scale and numerical values that correspond to the precipitation levels represented by each color.

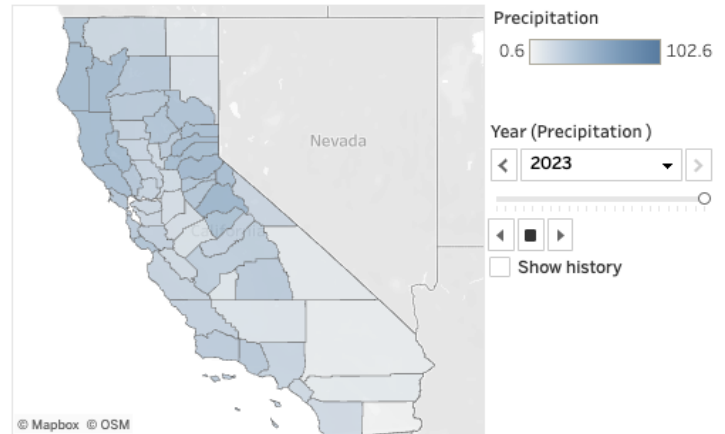
Interactivity: Users are also able to interact with the heat map by hovering over counties to view the exact precipitation level for each year. This provides a more detailed view of the data to allow the viewers to explore the information in more depth.

- Time slider: We also included a time slider beside the screen to allow users to view precipitation levels for different years over the past 20 years. This provides a temporal dimension to the data and allows users to see how precipitation levels have changed over that time.

Precipitation Heat Map - 2000



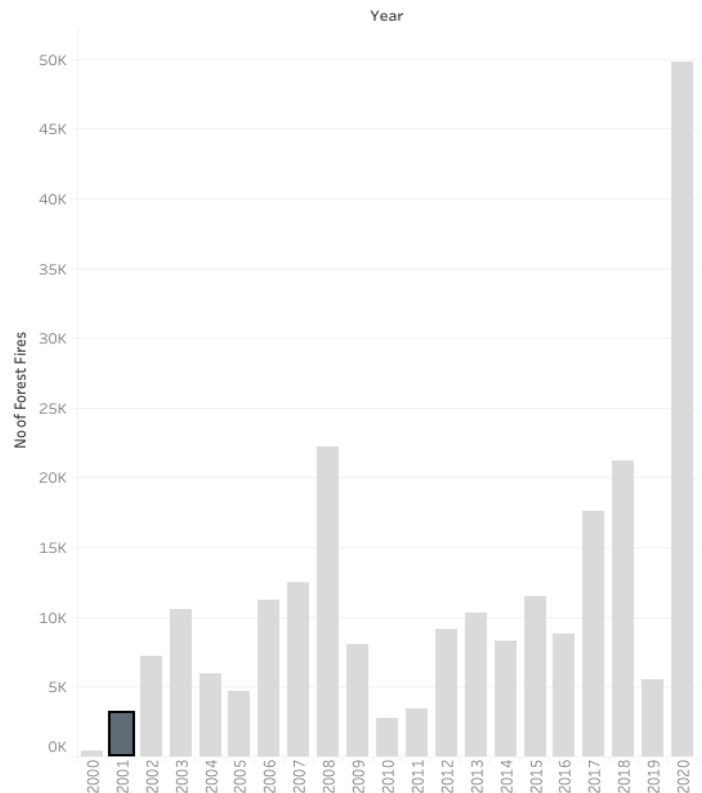
Precipitation Heat Map - 2023



Annual Forest fires from 2000 to 2020 in California

The Annual Forest Fires chart depicts the amount of annual forest fires detected in California from 2000 to 2020 based on NASA's MODIS satellite. It highlights the variations in the numbers of forest fires detected. A user who is new to this topic can easily observe that there is an increase in the number of forest fires from 2000 to 2020. The total number of forest fires in 2000 is around 3,000 whereas the number in 2020 is close to 50,000.

Annual Forest Fires (2000-2020)



Visual Techniques

We decided to use the bar charts for depicting this information. Each bar represents a year and on the X-axis the labels of each year are rotated for space optimization. Spacing between the bars is ensured to enable the users to distinguish between different years of information. The two axes are labeled and the horizontal grid lines make it easier for users to understand the relative differences between years. The chart is interactive and provides users with visual confirmation feedback on mouse over. The dark gray color was chosen for this chart since it can remind users of the rising smoke plumes of a forest fire.

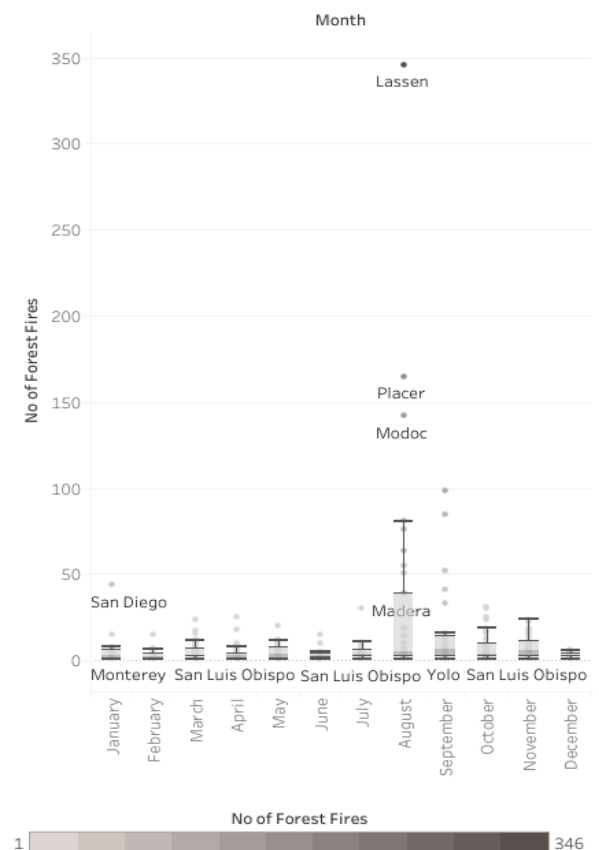
Monthly Variation in Forest Fire Data

Understanding the relative differences between different counties and also the monthly variation in forest fires is crucial for every user. This chart aims to provide the user with an understanding of how the forest fires incidences vary temporally in a year and also depict which counties have the highest numbers of forest fires in a particular month of a particular year. This chart is interactively linked to the previous chart so that when a user clicks a bar of a specific year the information in this chart also changes simultaneously.

Visual Techniques

We decided to use a box and whisker plot for this chart to help users understand the range of information and specially which counties are the outliers in each month. We decided not to use multi-color shades for the data points since similar colors were already used in previous charts. The month information is plotted on the X-axis and the number of forest fires is plotted on the Y-axis. The legend provides information on the number of forest fires in each county.

Month wise variation in Forest Fires in all counties - 2001



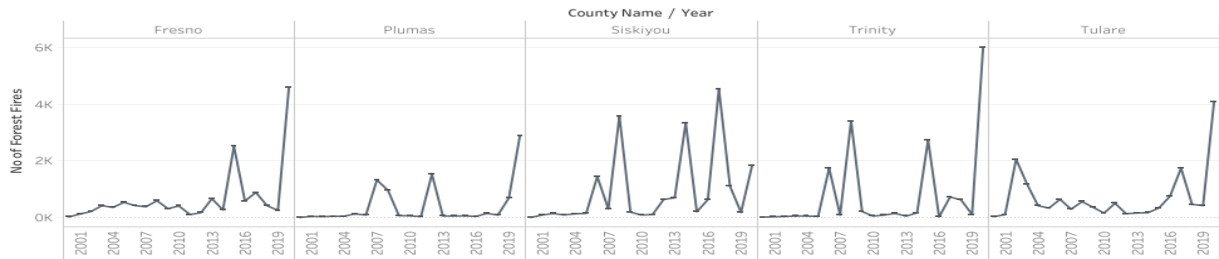
Top/bottom five counties forest fire data (2000-2020)

The top/bottom five counties forest fire data (2000-2020) helps the user quickly understand which counties are badly affected by forest fires and which are least affected. In the media, it is fashionable to say that the entire west coast is affected by forest fires. This unfortunately sends a wrong signal to people that there is uniformity in destruction wrecked by forest fires across the west coast. This needs to be corrected and it can lead to the generation of wrong narratives about this issue causing misinformed decisions. Preventing such misinformation is crucial and these two charts help users understand the drastic differences between different counties of the same state.

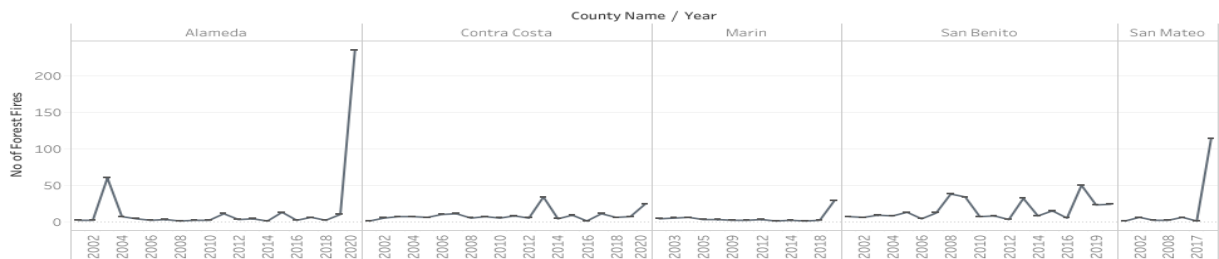
Visual Techniques

We decided to utilize the line graph for depicting the variation in number of forest fires in each county from 2000 to 2020. The axes are the same for each category. Horizontal stacking of the different charts helps a user to also study the temporal variations in forest fire between different counties. The vertical stacking of the top counties above the bottom five counties helps the users quickly compare and contrast at the y-axis values of highest number of forest fires. In the top row the user can see that the highest value of Trinity county is close to 6000 but in the bottom row the highest value is Alameda with just 250 odd values.

Top five counties in Forest Fire count (2000-2020)



Bottom five counties in Forest Fire count (2000-2020)



The Relationship between Forest Fires, Population Density and County Areas

Forest fires sweeping across the forests spread devastation. But they gain notoriety when more people are affected by death, smoke pollution and destruction of structures caused by the forest fires. Therefore we decided to help the users to explore the relationship between the number of forest fires, the population density and area of a county. We are able to observe that the population density of counties with more forest fires is generally low. Los Angeles is one of the few counties with high population density and high number of forest fires as seen in 2009.

Visual Techniques

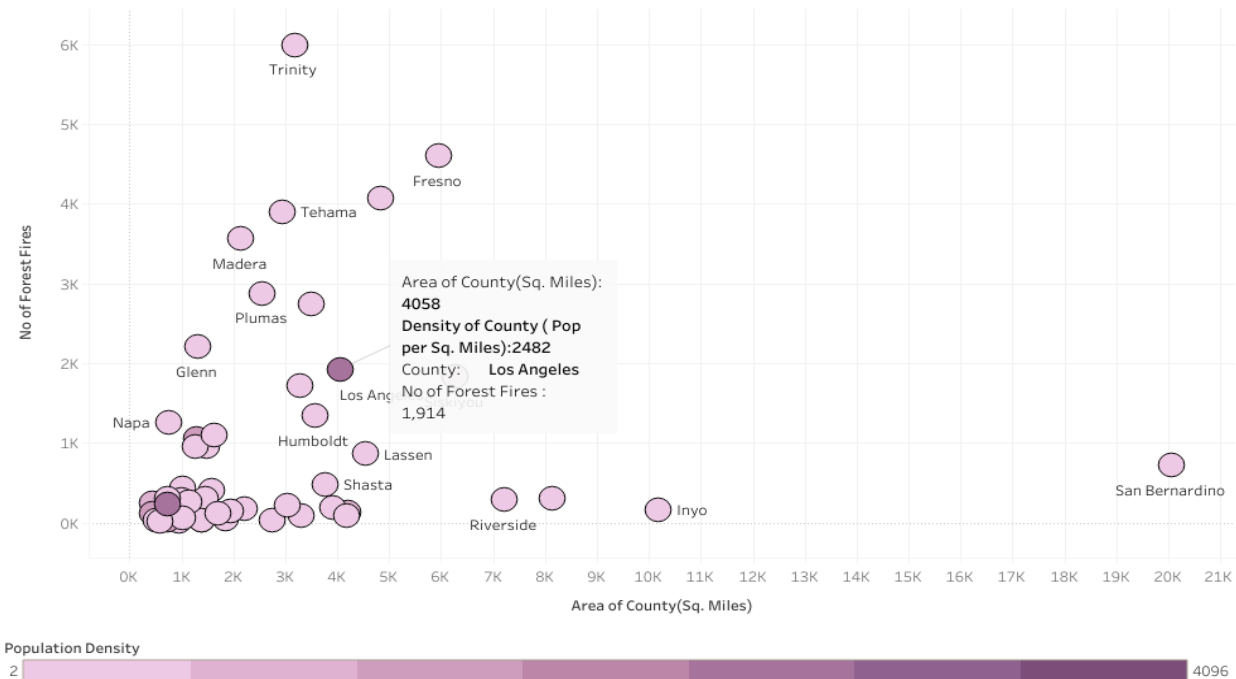
On the x-axis the area of the county is plotted while on the y-axis the number of forest fires is plotted. Each county's population density is represented through color encoding as depicted in the legend.

Slider - The chart is interactive and using the slider the users can view the forest fire data from 2000 to 2020.

Year
2020

Interactivity guide- Please move the slider to view how the data on forest fires in different counties varies from 2000 to 2020. At the left end of the slider you can view the cumulative data for the entire period. Please observe how Los Angeles county varies. The intensity of the color of each circular marker is directly proportional to the density of the county.

Understanding relationship between County area, population density and forest fires -2020



Approach

Temperature and Precipitation Dataset

Given the copious amount of data available relating to changes in climate indicators, we were intentional about finding data sources that were verifiable, accredited and trustworthy whilst also being easy to understand. We downloaded our data from the National Centers for Environmental Information which provided us with simplified data on temperature and precipitation changes over time.

The screenshot shows the NOAA National Centers for Environmental Information website. The header includes the NOAA logo and the text 'National Centers for Environmental Information' and 'Climate at a Glance'. A search bar is present with the text 'Search Monitoring Products'. Below the header, there is a navigation menu with options: Home, Climate Monitoring, Climate at a Glance. A date stamp reads 'April U.S. Release: Mon, 8 May 2023, 11:00 AM EDT'. The main content area features a navigation bar with options: Globe, Nation (selected), Region, State, Division, County, City. Below this, there are five main data categories, each with an 'Access Now' button: Mapping (National Mapping), Time Series (National Time Series), Rankings (National Rankings), Haywoods (National Haywoods), and Data Info (National Data Info). A sidebar on the right lists various climate products: All Products, Monthly Climate Reports, Climate at a Glance, Billion-Dollar Disasters, Temperature Precipitation, Snow and Ice, Drought and Wildfire, Storms and Wind Extremes, Weather Patterns, Key Datasets, and Other Monitoring.

<https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/county/time-series>

We downloaded csv files for each of the 58 counties separately. These files were then cleaned, processed and merged to form a single dataset using Python coding. In the raw file the year and month were fused together which prevented creation of any visualization. This was separated to form two separate fields. Additionally , a new field for county names was created.

Wildfire Dataset

The dataset is a collection of over 283k data points containing information of Forest fires in the different National Park of California from 2000 to 2022 detected through the NASA Land, Atmosphere Near real-time Capability for EOS (LANCE) Fire Information for Resource Management System (FIRMS) using the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Terra/Aqua satellites.

[Dataset source website](#)

Dataset was downloaded as a csv file.

Summary of dataset download-

Download Id : 343028

Data Source : MODIS C6.1

Start Date : 2000-11-01

End Date : 2023-04-10

Output Format : csv

Area of Interest: -124.8,32.6,-114.4,42.1

Transformations of the dataset

1. The dataset includes fire locations that may be forest fire, other terrestrial fire, fire in sea and fire from volcanic activity. It is a set of x/y locations with details of map. The dataset was filtered using tools in ArcMap to include only the forest fires.
2. Using QGIS, two new fields were created in the attribute table for year and month of the forest fire.
3. The dataset was processed in ArcMap using the intersection tool with a layer of California counties to filter the fire points on the basis of California counties (Source - <https://gis.data.ca.gov/datasets/CALFIRE-Forestry::california-county-boundaries/about>). This provided every datapoint with a new field having the name of the respective California county.
4. The csv file was then processed in Python using Pandas library to convert the numeric values of months to text data. Irrelevant fields in the dataset were dropped.

California Counties Demographics Dataset

We used information on demographics of the different counties of California including population density, population growth, 2023 population and area from the World Population review website. (<https://worldpopulationreview.com/states/california/counties>). The data was downloaded as a CSV file. The county names were cleaned in Excel so as to match the county names in the other two datasets.

Tools

For all three datasets, data was cleaned and sorted using Python/QGIS/ArcMap/Excel whilst visualization generation was conducted in Tableau. Other visualizations included in the screenshots were also created using d3 in Observable plot, however, these were not included in our final website. The webpage was made using HTML, CSS and Visual Studio Code for coding. Other illustrations were also made using Figma but were not used in the final website. Other visualizations and code lists are linked at the end of this document and all used resources are referenced at the bottom of the page.

Steps

1. Climate Change in California Research

We searched for and agreed on the temperature, precipitation and wildfire datasets that were needed in order to tell our story as planned during our research and design phase. The following steps were taken in regards to this.

- **Temperature and Precipitation dataset**

We downloaded csv files for each of the 58 counties separately. These files were then cleaned, processed and merged to form a single dataset using Python coding. In the raw file the year and month were fused together which prevented creation of any visualization. This was separated to form two separate fields. In addition a new field for county names was created.

- **Wildfire Dataset -**

The dataset includes fire locations that may be forest fire, other terrestrial fire, fire in sea and fire from volcanic activity. It is a set of x/y locations with details of a map. The dataset was filtered using tools in ArcMap to include only the forest fires.

- Using QGIS, two new fields were created in the attribute table for year and month of the forest fire.
- The dataset was processed in ArcMap using the intersection tool with a layer of California counties to filter the fire points on the basis of California counties

(Source -

<https://gis.data.ca.gov/datasets/CALFIRE-Forestry::california-county-boun>

[daries/about](#)). This provided every datapoint with a new field having the name of the respective California county.

- The csv file was then processed in Python using Pandas library to convert the numeric values of months to text data. Irrelevant fields in the dataset were dropped.
- **California Counties Demographics Dataset**
The data was downloaded as a CSV file. The county names were cleaned in Excel so as to match the county names in the other two datasets.

2. Initial Visualizations

After cleaning the data and agreeing on the story that we were trying to tell, we created some prototype visualizations using d3 in Observable plot and Tableau. This allowed us to discuss the vision of the final project deliverable and ideate around the other possible visualizations that could be added in or removed to best meet our goals of creating awareness around the subject matter whilst also presenting our understanding of data visualization techniques.

3. Defining Our Story

We divided the work between ourselves with one team member focusing on temperature and precipitation storytelling and the other focusing on wildfires and how the prior mentioned indicators relate to wildfires in California. We ideated over this process numerous times as our story telling and vision for the project began to improve. We were now ready to share our story with a few possible users.

4. User Testing and Refinement

The user testing experience was an incredibly helpful experience for us as we received useful feedback and suggestions on how we could improve our visualizations to better communicate the information we were attempting to convey. We tested the design with 3 students that had some knowledge about the interpretation of data visualizations but not a lot of knowledge about Climate Change in California. Once this was complete, we were able to incorporate the feedback and get started with building our website and refining our story.

User Testing and Results

Introduction

Our user testing was conducted with 3 prospective users who were all UC Berkeley masters students that expressed an interest in wanting to understand the changing climate within California and the possible impacts that these changes are having. None of the participants communicated having any additional interest or background in climate change over and above that of the average human being. The goal of our study was to evaluate the effectiveness and user-friendliness of our data visualization tools whilst also identifying areas in need of improvement.

Method

Our method for the user testing was to recruit 3 willing participants to interact with prototypes of our data visualization charts and provide feedback on their experience and understanding. Prior to conducting the tests, the team developed testing protocols and materials which were used during the testing process which was conducted in person. We asked participants to “think aloud’ in order to assist us in better witnessing their thought process and understanding if the tools were confusing or well received as intended. At the end of each test we asked participants to fill out a short questionnaire that among other things was based on their understanding of the information being relayed through the visualizations.

Participants

Our Participants

UC Berkeley Masters students that expressed an interest in wanting to understand the changing climate within California and the possible impacts that these changes are having

Selection method

Random selection of three participants from a list of twelve who expressed interest in participation

Scenarios/Tasks

Each participant was given the same scenario and results were analyzed upon completion.

Participants were each provided with the opportunity to interact with 4 Climate change related data visualizations. As they were interactive with the visualizations they were asked to adhere to

the “think out loud” method and explain their understanding of the visualizations as they understood them.

The testing process then progressed to asking the participants to each answer very specific questions which were predetermined by the interviewees: The questions are as follows:

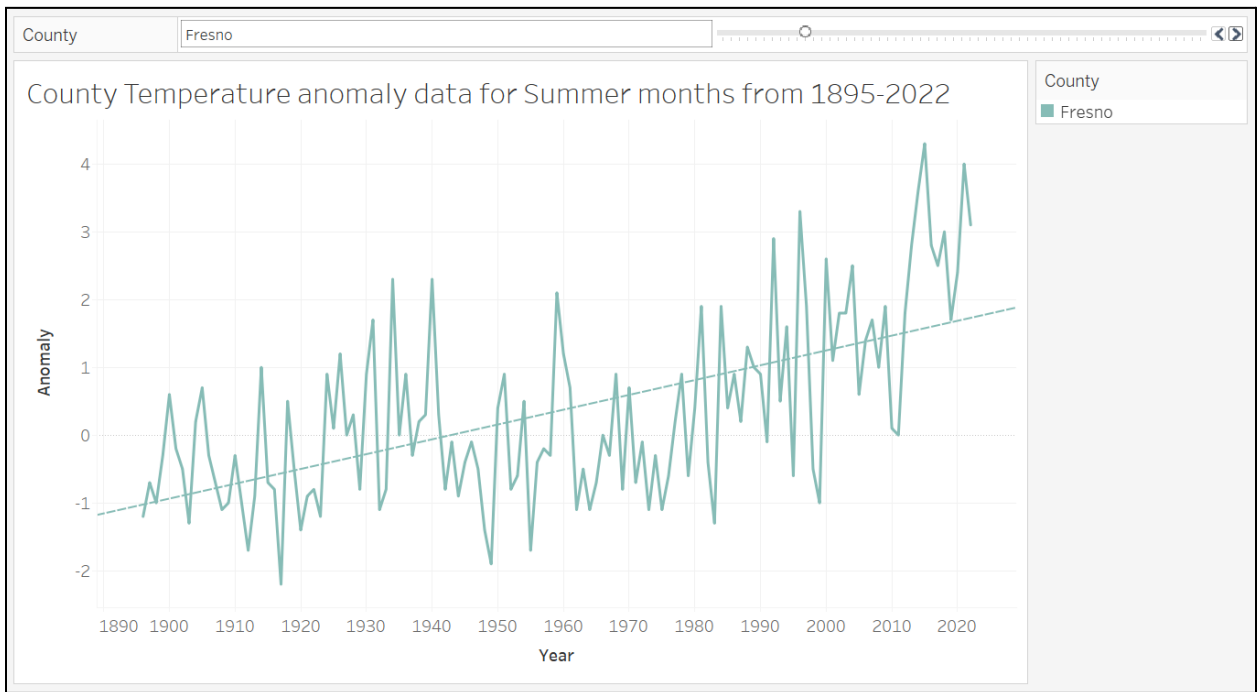
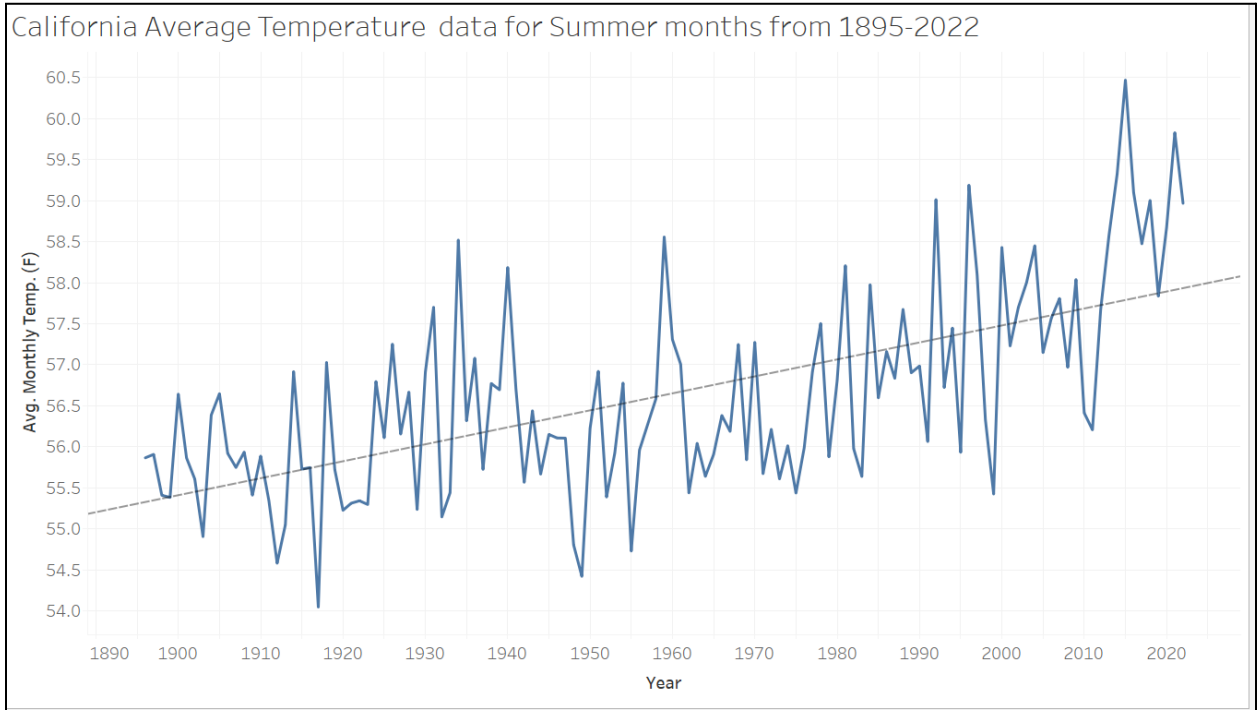
- In what year was the average temperature in California 56.6 degrees Fahrenheit?
- What year has the lowest average temperature in California?
- Name a country in California whose average temperature is higher than that of another in August from 1895 - 2022?
- Name the county which had the highest number of wildfires in the year 2020?
- Name the National forest which had the highest number of wildfires in August 2019?
- Name the California Vegetation Zone that had the highest numbers of Wildfires in 2007

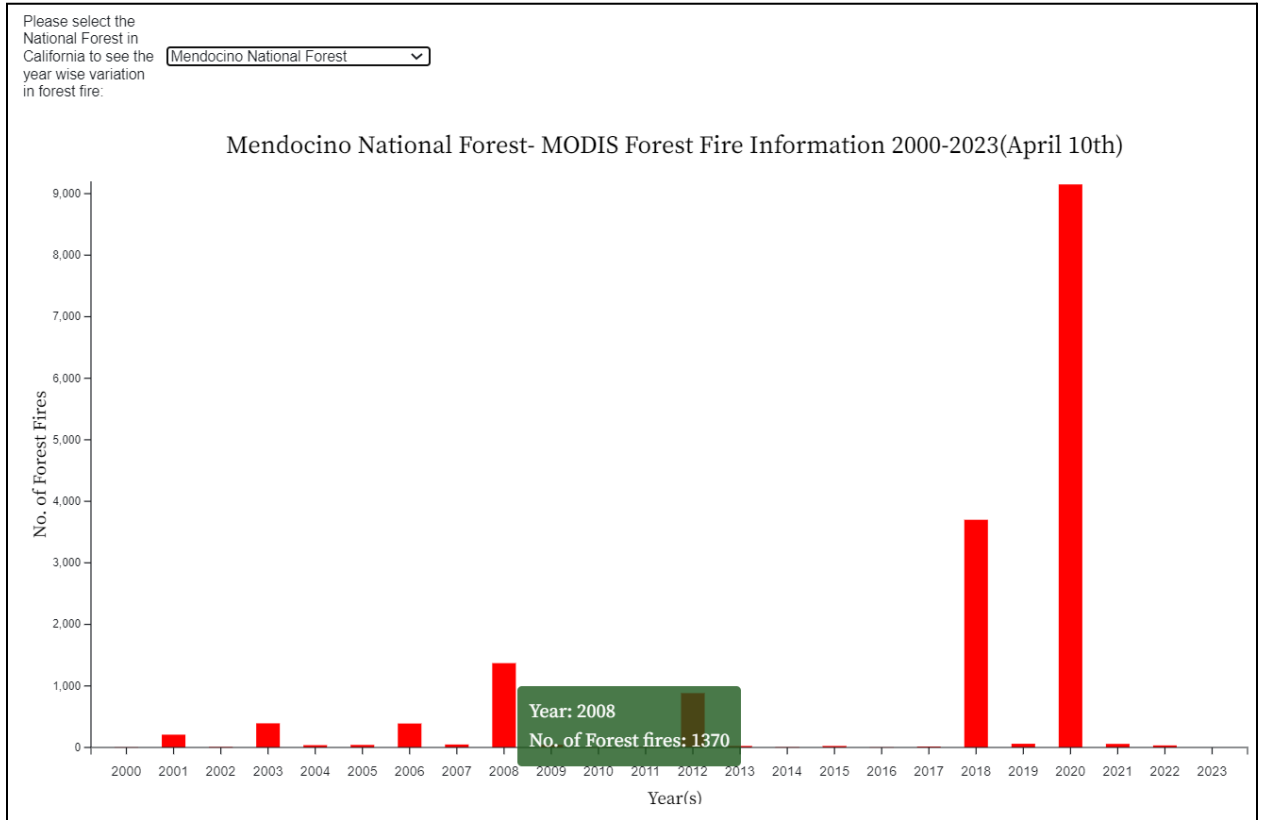
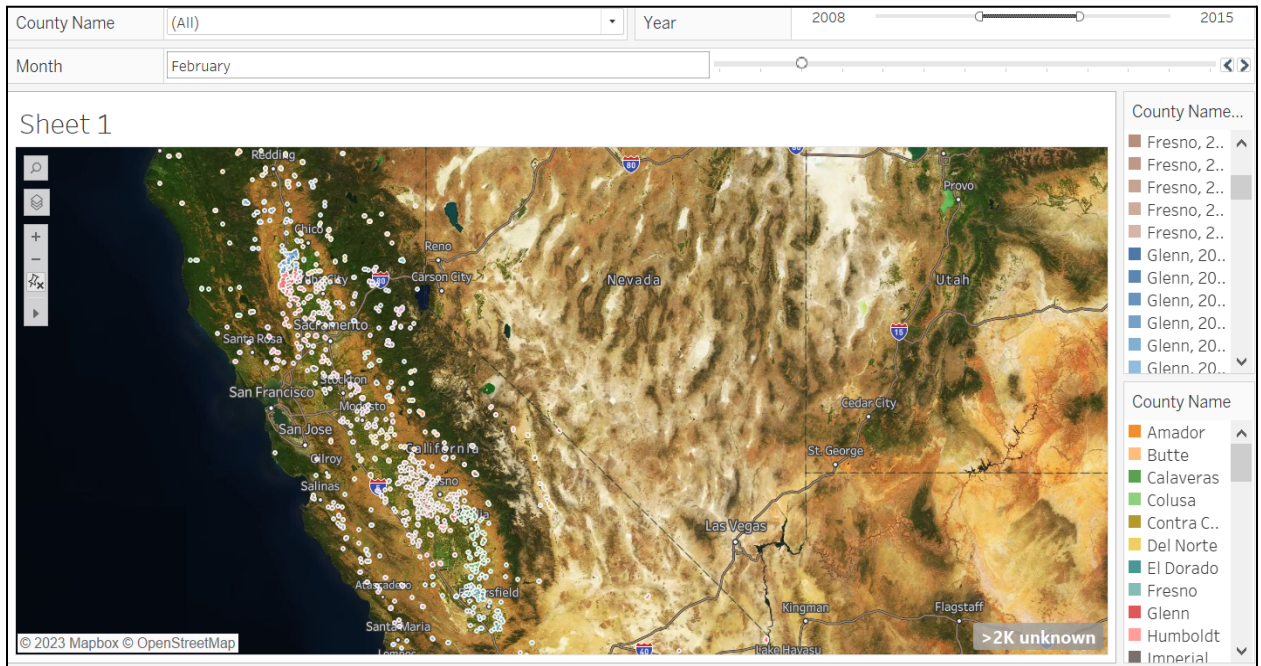
The questions were designed to test the users understanding of the data visualizations. Correct answers to the questions above indicated a higher level of usability and ultimately clarity of information relayed. During this stage of the testing, interviews were interested in and recording the following response variables:

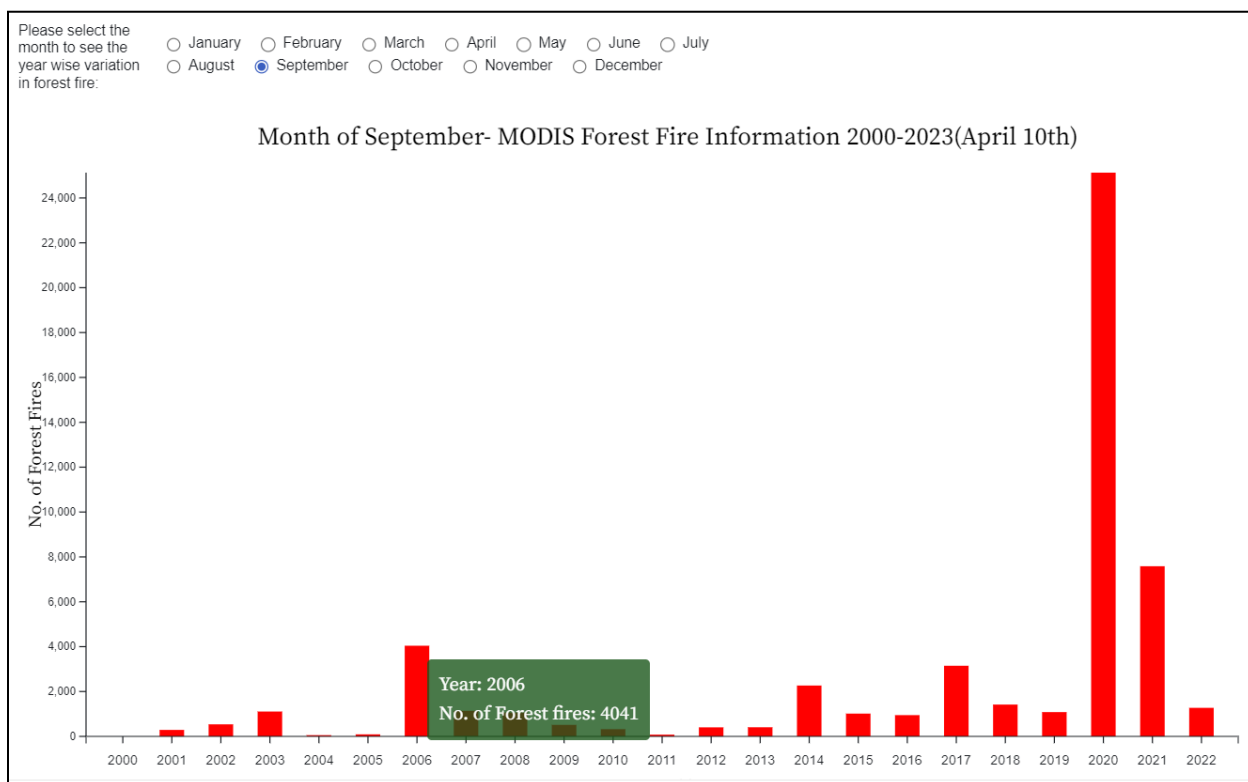
- Task Completion Time
- Number of Errors
- Questions Indicating Confusion/ Uncertainty

The next stage of the usability test was to ask participants generalized task and feedback related questions that were designed to assess and understand the users overall experience. With the use of Likert scale type questions, users were asked to rate the difficulty or ease of the tasks they were previously asked to complete.

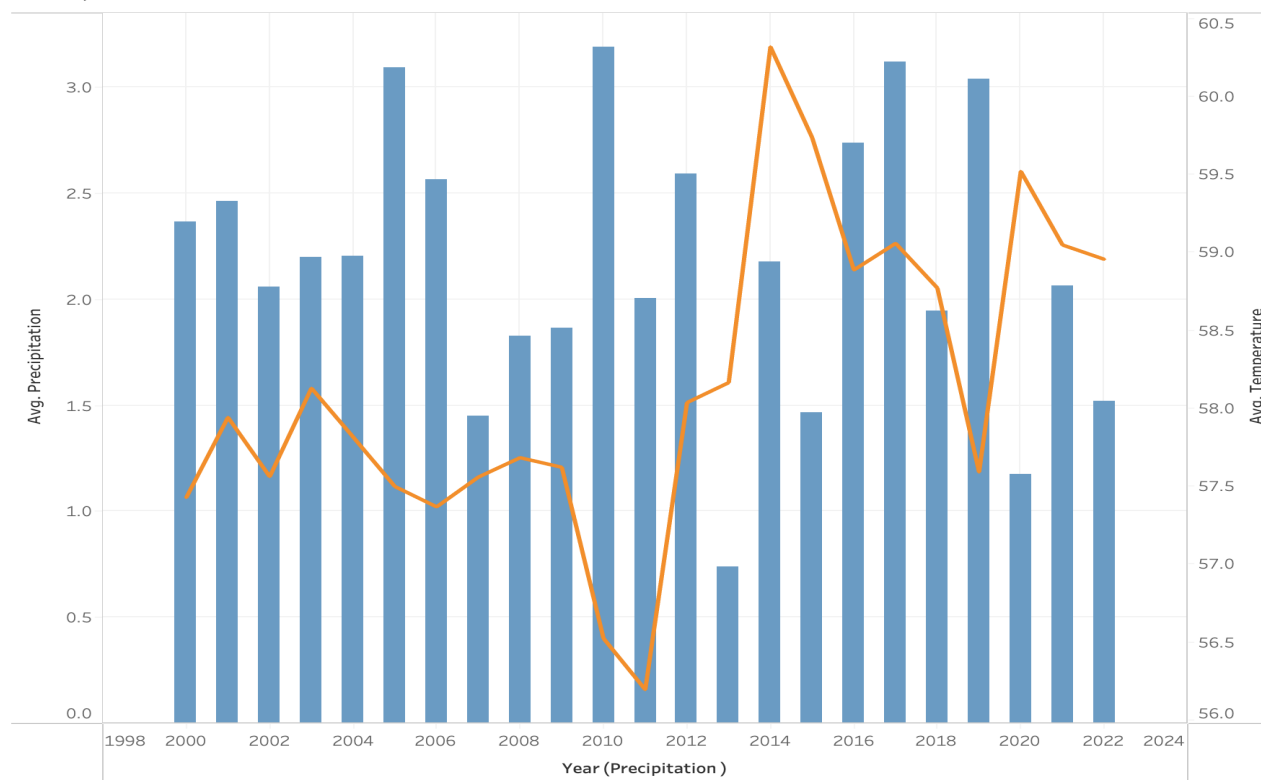
Screenshots







Temp/Precipitation



Test Measures & Results

Usability Testing	User1	User2	User3
Task ✓ - Correct ✗ - Incorrect			
In what year was the average temperature in California 56.6 degrees Fahrenheit?	✓	✓	✓
What year has the lowest average temperature in California?	✓	✓	✓
Name a country in California whose average temperature is higher than that of another in August from 1895 - 2022?	✓	✓	✓
Name the county which had the highest number of wildfires in the year 2020?	✓	✓	✓
Name the National forest which had the highest number of wildfires in August 2019?	✓	✓	✓
Name the California Vegetation Zone that had the highest numbers of Wildfires in 2007	✓	✓	✓
Task Related Feedback Likert Scale 1 -10 1 - Difficult, 5 - Neutral, 10 - Easy			
1. On a scale of 1- 10, how easy or difficult was it to complete the task?	9	10	10

2. On a scale of 1 - 10, was the data visualization tool intuitive and easy to use?	8	9	10
3. On a scale of 1 - 10, were you able to understand the data presented in the visualization?	10	10	10
4. Overall, how would you rate the usability of the data visualization tool on a scale of 1 to 10?	10	10	10
Feedback Questions (See Appendix)			
1. Did you encounter any challenges or issues while completing the task?	Yes	No	No
2. Were there any features or functionality that you found particularly confusing?	Yes	No	Not Really
3. What changes or improvements would you suggest for the tool?	Change filter	No	No
4. Are there any features or functionality that you would like to see added to the tool?	No	No	No

Discussion

The pilot usability study was a great learning opportunity for us. It enabled us to understand the viewpoints of the ultimate users of the visualization. In designing visualizations it is possible that the personal preferences/biases of the designer dominates the design process. Due to which it is possible that the design may have features that are incompatible with the needs and requirements of the final users of the visualization. The aim of any visualization is to help the ultimate user gain valuable insights and form connections between the different data. A designer of a visualization fails unless these key goals are attained.

In our usability study the respondents examined the visualization thoroughly and they completed the tasks. The feedback we received was mostly positive. One crucial feedback we obtained

was regarding the different filters used. Our respondent wanted improvements to be made to the filters to make them more intuitive and understandable.

On the basis of the results of this study we plan to improve our filters and make further improvements to meet the needs of the ultimate user.

Link to Code and Visualization

- [Link to our visualization of 20 Year of Climate Change in California](#)
- [Github Repository containing Source code](#)

Division of Labor

The table below is an approximate indication of the proportion of labor division between team members for the purposes of this project.

Project Component	Sub-Category	Wadzi	Soorya
Research	Problem Statement Definition	20%	80%
	Literature Review	50%	50%
	Research Design	50%	50%
Data Collection	Data Sourcing	40%	60%
	Data Cleaning	20%	80%
Visualizations	California Counties Temperature Trends	100%	
	Anomaly Temperature Heat Maps	100%	
	Average Precipitation Over Time	100%	
	Precipitation Heat Map	100%	
	Annual Forest fires from 2000 to 2020 in		100%

	California		
	Monthly variation in Forest Fire Data		100%
	Top/bottom five counties forest fire data (2000-2020)		100%
	The Relationship between Forest Fires, Population Density and County Areas		100%
Design	Website Text Writeup	50%	50%
	Website Layout Design	60%	40%
User Testing and Others	User Testing	50%	50%
	Report Writing	60%	40%
Final Paper	Report Writing	60%	40%

Appendix

Tableau Workbooks

- [2000 - 2020 Temperature Dashboard](#)
- [2000-2020 Precipitation Dashboard](#)
- [Annual Forest fires from 2000 to 2020 in California & Monthly variation in Forest Fire Data](#)
- [Top/bottom five counties forest fire data \(2000-2020\)](#)
- [The Relationship between Forest Fires, Population Density and County Areas](#)

Usability Test Script

Hi there. Thank you for taking the time to participate in our user testing study.

Today, we will be testing some data visualizations and gathering your feedback on their usability.

First, we want to assure you that there are no right or wrong answers, and we are simply interested in your honest feedback.

We will start with some background questions to better understand your data viz experiences.

- Can you tell us a bit about your background and experience with data visualization tools?
- Have you related with any data viz tools in the past? If so, for what purposes?

Now, let's get started with the testing. We will be asking you to complete a few tasks using the data visualization tool we have provided.

You will be given a scenario and asked to perform certain tasks, such as finding specific data points or changing the visualization type.

Please feel free to take your time and ask questions if anything is unclear.

Tasks

- In what year was the average temperature in California 56.6 degrees Fahrenheit?
- What year has the lowest average temperature in California?
- Name a country in California whose average temperature is higher than that of another in August from 1895 - 2022?
- Name the county which had the highest number of wildfires in the year 2020?
- Name the National forest which had the highest number of wildfires in a August 2019?
- Name the California Vegetation Zone that had the highest numbers of Wildfires in 2007

Task questions:

1 - easy, 5 - neutral, 10 - difficult

5. On a scale of 1- 10, how easy or difficult was it to complete the task?
6. Did you encounter any challenges or issues while completing the task?
7. Were there any features or functionality that you found particularly helpful or confusing?
8. On a scale of 1 - 10, was the data visualization tool intuitive and easy to use?
9. On a scale of 1 10, were you able to understand the data presented in the visualization?

Feedback questions:

1. Overall, how would you rate the usability of the data visualization tool on a scale of 1 to 10?
2. What changes or improvements would you suggest for the tool?
3. Are there any features or functionality that you would like to see added to the tool?

Thank you again for your time and feedback. Your feedback has been insightful and will help us improve the usability of our data viz tool.

If you have any further questions or feedback, please don't hesitate to contact us.

Feedback Question Responses

Feedback Questions (See Appendix)
<p>5. Did you encounter any challenges or issues while completing the task?</p> <ul style="list-style-type: none">● Yes, I encountered many challenges● Yes, I encountered some challenges● "No, I did not encounter any challenges"
<p>6. Were there any features or functionality that you found particularly confusing?</p> <ul style="list-style-type: none">● Yes, I found all features to be challenging● Yes, I found a few features to be challenging● No, I did not encounter any challenges
<p>7. What changes or improvements would you suggest for the tool?</p> <ul style="list-style-type: none">○ Yes,I will suggest complete overhaul of the tool○ Yes,I will suggest minor changes to the tool○ No,I will not suggest any changes to the tool
<p>8. Are there any features or functionality that you would like to see added to the tool?</p> <ul style="list-style-type: none">○ Change filters○ Change font○ Change marks

- Change color of marks
- Change scale
- None

Acknowledgments

- This research benefited from the support and services of UC Berkeley's Geospatial Innovation Facility (GIF), gif.berkeley.edu.
- We acknowledge the use of data and/or imagery from NASA's Fire Information for Resource Management System (FIRMS) (<https://earthdata.nasa.gov/firms>), part of NASA's Earth Observing System Data and Information System (EOSDIS).
- We acknowledge the use of information from NOAA website for this project <https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/county/time-series>