Visualization of US Wildfire Data

CptS 475 Final Project
By Aldo Zepeda and Hongbo Wang

Abstract:

Our project is the visualization of US wildfire data from the United States Forest Service. We took the SQL database of wildfires from 1992-2015 and split them into six four year periods. We then mapped out all the wildfires according to their time period, while also displaying their size in relation to one another. Then we took the data and determined how much each fire size class contributes to both acres burned and the total number of fires. and found that while the majority of fires are class A-C they account for a very small percentage of acres burned. Also the inverse is true for class F and G fires. We then created several graphs displaying average fire size, acres burned and number of fires throughout the years. We determined that while the number of annual fires has remained steady the average number of acres burned has almost tripled. Finally we analyzed how the different causes of fires have changed through the years and determined which types of fires contribute most to acres burned annually, and found that a majority of total acres burned has been in fires caused by lightning.

Introduction:

Climate change has been one of the topics that has been at the forefront of many people's minds worldwide. From global warming to habitat destruction, it's no secret humans have been having an adverse effect on the environment for centuries. For decades many governments and organisations have been working to slow and prevent the adverse effects. One such problem is global warming and its leading cause is an over abundance of carbon dioxide being put into the atmosphere. Across the world many natural habitats have acted as carbon sinks, where biomass such as trees, shrubs and even moss have collected carbon from the atmosphere and stored it in the ground.

In the past year alone we have had many wildfires large enough to make global headlines, One such fire is the Australia bushfires, which burned over 27 million acres and destroyed more than 2,000 homes. Another, from a few months earlier, was the several Amazon Rainforest fires, which burned over 2.2 million acres. It is evident that these fires are growing larger and more destructive. Wildfires such as these destroy carbon sinks that carbon stored millennia worth of carbon. . This only accelerates global warming causing more fires, and so the cycle will continue if we don't find ways to prevent or stop wildfires. However that is not the only side effect of wildfires, they also destroy natural habitats, leave land susceptible to mudslides and endanger many human lives.

Therefore it is important to take a look at historic wildfire data and try to learn what we can about how wildfires have changed over the years. We intend to study their geography, size, causes, and more in order to gain new insights into how wildfires can be prevented and fought in the future.

To start with we plan to do some basic exploratory analysis on the 37 attributes of our data set and determine if there is any correlation between any of the data collected. After that we intend to create several mappings of the wildfires in the US over the 24 years the data represents. This can help us easily see and show any trends or interesting geographic features of the data. Wewill also create several basic graphs to show how the fires have changed, as well as the proportions of the different size classes of fires, as well as of the 13 types of reported wildfire causes.

Once we have gone through and analyzed the data we will report our findings in the following sections of this report.

Problem Definition:

In this project we intend to explore and hopefully address several questions. The first one is to confirm if in fact wildfires have been becoming larger and more frequent. This is important because we would not want ours or other peoples opinions on issues like these be made solely based on the word of sensationalist headlines or a few abnormally large fires.

Second, we would like to know using the dataset what percentage each size class of wildfires is of the total. similarly we would like to know what percentage of the total acres burned is burned by each size class. We would also like the same information in relation to the individual causes of wildfires. This would help us determine where efforts in wildfire fighting and management should be focused. Big or small fires? Man caused fires or natural ones?

Another thing we would like to analyze is if there are areas where, either large or small, wildfires are more frequent or larger. In the case where such areas do exist we would like to propose a larger focus on wildfire prevention and fighting in these areas. This would help the government and organisations fighting wildfires properly allocate their resources.

Finally, we would like to determine what the leading cause or causes of wildfires is in terms of acres burned. This would again help determine where and how funds and resources should be allocated to limit the damage caused by these kinds of fires.

Models/Algorithms/Measures:

In investigating and analysing the dataset we primarily relied on Tableau as our tool of choice. Using it we developed several visual representations of data features such as charts and graphs to represent the annual averages or totals of certain features such as acreage burned, or total fires. This allowed us to determine if we should focus on certain classes of fires or not.

Another thing we did was take the thirteen types of fire causes reported in this dataset and plot them, using a barchart, how the number of fires caused by each has fluctuated from 1992 to 2015. By doing this we were able to see trends and determine whether each individual cause has either grown,shrunk, or had no discernible change in how much fires they have caused.

In a similar fashion we also developed several pie charts that represent proportions of these same features, number of fires and acres burned by each size class. Again we took the same values but plotted the proportions according to the cause of the fire. Using this we could determine which types of fires, if any should be focused on in terms of study, investigation, prevention and fighting.

However, we believe the most important model we developed were the many maps that plotted fires according to their time, size, and location. Using Tableau we created several maps, each individual map represented one of two ranges of fires classes, then plotted every fire that occurred within one of six, four year ranges. The points were made translucent so that areas with many fires overlapping could more clearly be understood. Also the actual size of each fire determined how large its point or bubble was. By taking all six mapping of certain locations it provided a nice visual of how fires in different areas have evolved, whether in size or quantity, over the 24 years represented.

Finally once we learned about how large of a contributor to total acres burned lightning caused fires, we recreated the earlier annual bar chart representing that. However this time we split each year's bar chart and colored it according to the cause of the fires. Once that was done it was clear to see how much of an impact this type of fire has on total acres of land burned every year.

Implementation/Analysis:

The dataset we used in this project is a publicly available relational SQL Database compiled and cleaned by Karen C. Short for the United States Forest Service. The data was collected using the national Fire Program Analysis system which collects fire data from federal, state, and local fire departments. This dataset covers over 1.8 million wildfires from 1992 to 2015, and includes location data for each fire in the form of latitudes and longitudes, accurate to within one mile. The dataset includes 37 features, ranging from acres burned, location, size class, discovery and containment dates, and the cause of each fire.

In our analysis we used the longitude and latitude of each fire in order to map their location, and used its size to determine how large the points on our mappings should be. To create our mappings we also used the fire year, fire size class, state to determine what and where to map. Other useful data features we used were the statistical cause code, statistical cause description, and a few of our own generated sums and averages of these features.

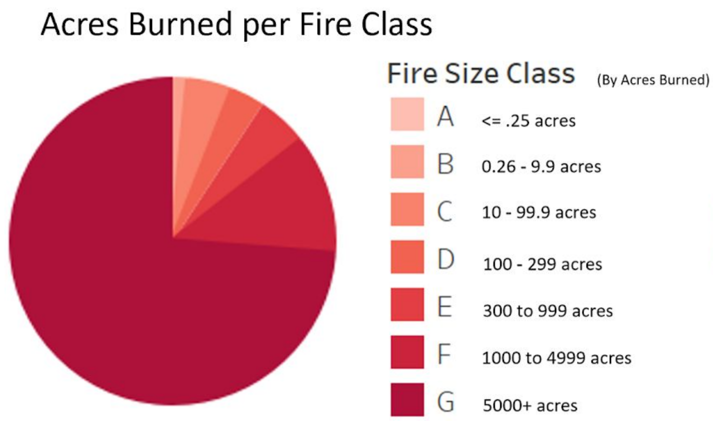
To analyze all this data we will be using a software package called Tableau. Tableau allows us to quickly and effectively go through the exploratory data analysis process with its built in features. We can then use it to create easily interpreted maps, charts and graphs that will allow us to find trends or interesting features of the data. Finally once we feel we have found meaningful information, it provides many tools for us to create effective visualizations that share the findings to our audiences.

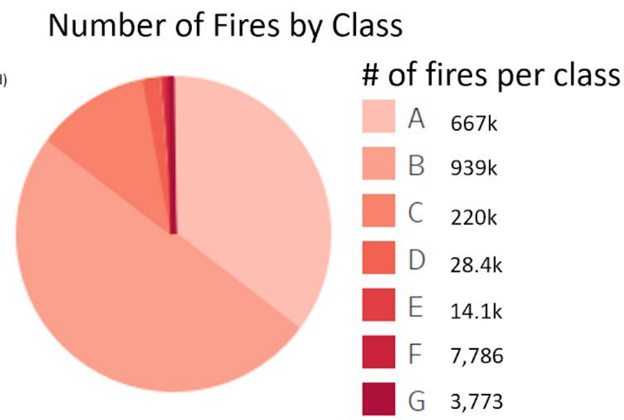
By using all the above data, graphs, and mappings we hope to discover any meaningful trends or correlations in the size of wildfires, their locations, and frequency. We hoped to perhaps find areas more prone to wildfires or where wildfires seem to be increasing in size or quantity. We also wished to determine through the proportions of the data features what the leading causes of wildfires and of total burned acreage are. This will point us to where further research and investigation should be targeted and hopefully towards more effective prevention, investigation and firefighting.

Finally although we did not develop it we did use a gif, created by Vivid Maps of the United State’s population density throughout the same years we are analyzing.This animation uses data from the United States Census Bureau to display the annual population density of the nation starting at 1790 and through 2010. However since we only needed information starting in 1992 we cut all the parts of the animation displaying years before 1992. We decided to include this in order to compare the mappings we create with the ones provided to see if any of our information or trends we discovered seemed to correlate with how densely populated those same areas are.

Results and Discussion

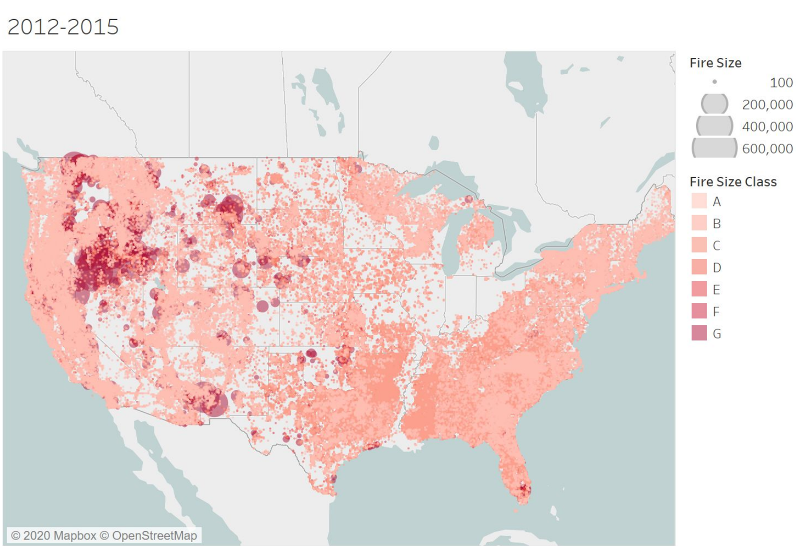
At the start of our research of the data after familiarizing ourselves with it and analyzing basic statics such as minimums, maximums, averages and sums we first decided to create two pie charts representing the proportions of wildfires and acres burned by fire size class where each size class has its own shade of red to distinguish it. Below are both of these charts.





By analyzing these pie charts we can see that Class G fires alone account for nearly seventy-five percent of all acreage burned, whereas they only represent less than a percent of the number of fires reported. And when combined with class E and F fires it's plain to see that most damage is done by a minority of fires. Therefore if we can learn more about these classes of fires we will be well on our way to accomplishing our goal.

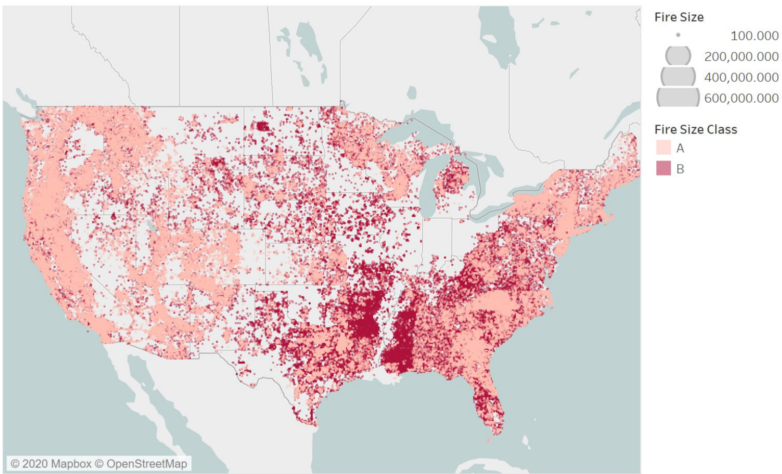
And by contrast we can see that while class A through C fires burned about ten percent of all acreage, however they account for nearly ninety-five percent of all fires reported. Class B fires alone account for about fifty percent while class As account for about a third of all reported wildfires. We can hypothesise that perhaps most wildfires are contained before they get a chance to grow beyond these sizes, and by analyzing these smaller class fires we can learn more about how or why these fires are contained before growing even more out of control into larger class fires.



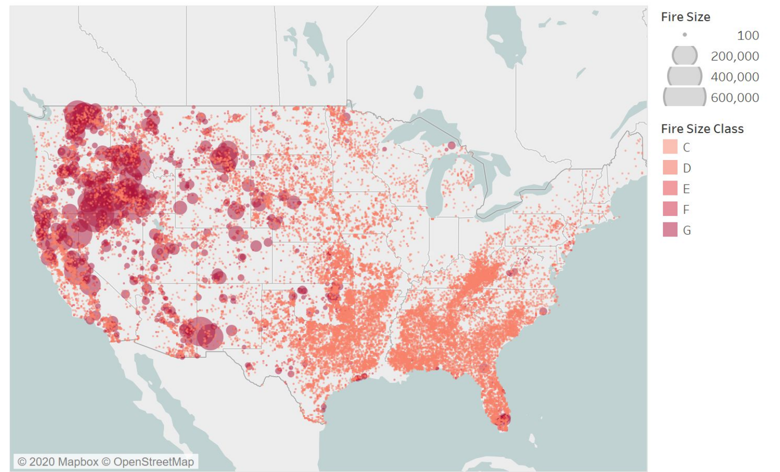
Armed with this knowledge we then decided to begin the mapping of the fires. We started by simply mapping each point over a map of the contiguous United States, while color coding each fire by its fire size class attribute. For each size class from A to G, we mapped a shade of red to it starting with a nearly white red all the way up to a deep blood red. After noticing just how many fires there were and how cluttered the map was, we decided that would be a good time to make copies of the maps and filter them by year into six segments. From there to help make areas with many overlapping fires more clear we make each point translucent. The above map is an example of one of the resulting maps representing wildfires from 2011-2015.

As can be seen this map is still very messy and hard on the eyes. Therefore since we know from the previous charts that smaller class fires represent the majority but only burn a very small percentage of fires, and vice-versa for large class fires, we decided to split each four year segment into two maps. One to represent class A and B fires, which are fires smaller than 10 acres. And another to represent class C through G fires, which are all fires 10 acres and larger. This proved to be much easier to analyze and present. Below is an example of the above 2011-2015 map but separated into aforementioned maps.

2012-2015 A-B



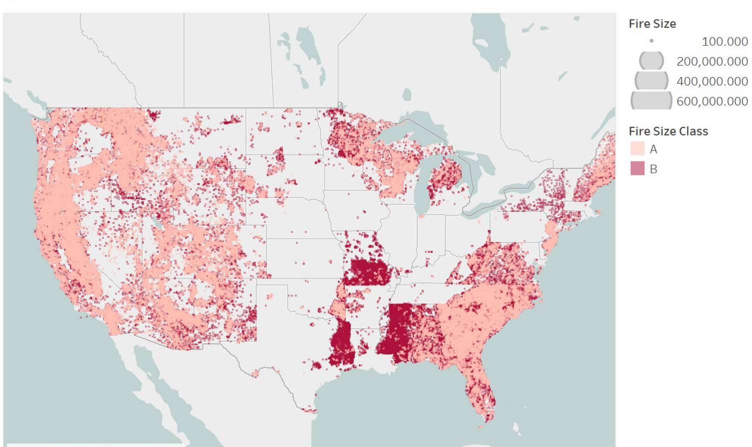
2012-2015 C-G

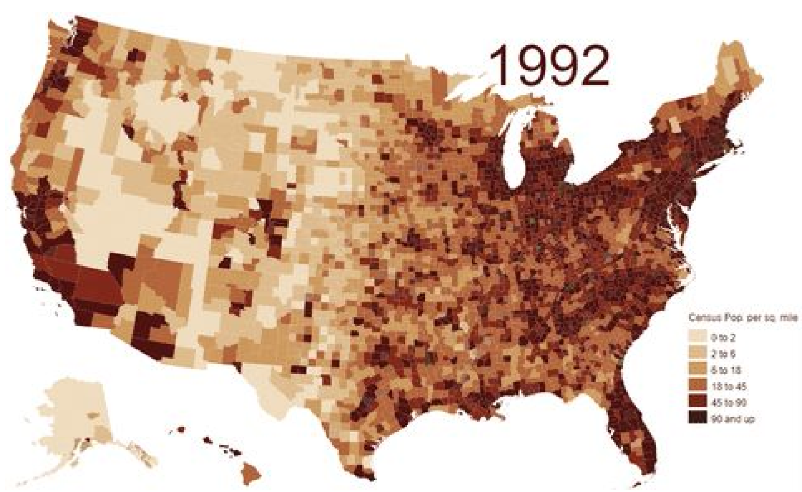


By doing this we can now clearly see the distribution of both sets of fires. Also in the case of A and B fires their shades were so similar before that it was difficult to distinguish them, so now that they are separate maps we can reset their color palette to easily distinguish between the two.

Now once we saw the distribution of the class A and B fires we saw that it was quite similar to the population density of the United States. For example both coasts have high densities of fires with much more fires in the eastern half of the US. Also the bible belt has a smaller density of fires. The larger fires also follow a similar trend but when compared to population density maps from Vivid Maps it's plain to see that smaller fires definitely correlated to population densities. Below is an example of this for the year 1992.

1992



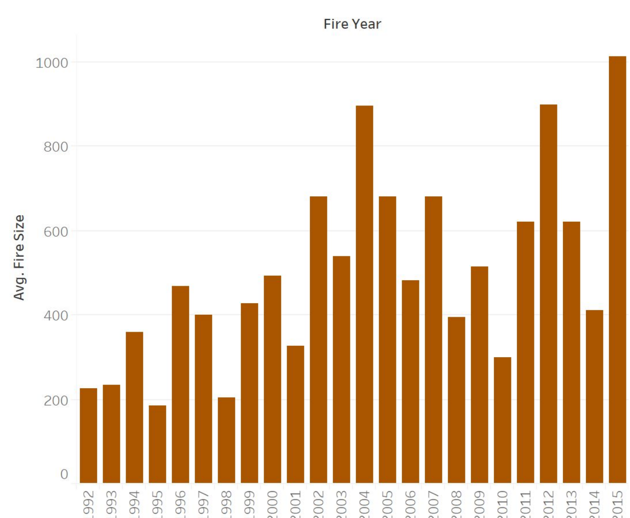


We created similar maps for all six four year groups, which will be attached in appendix A. Also attached there will be the same maps for the states of Alaska and Hawaii, however since these maps were not as cluttered as the others we did not separate each four year group into small and large fires. This way there were less maps and everything still remains clearly visible.We also took all of these six periods sequentially and created gifs to more clearly show the changes throughout the years, this will be in a powerpoint which will be in a github repository that is linked in appendix B.

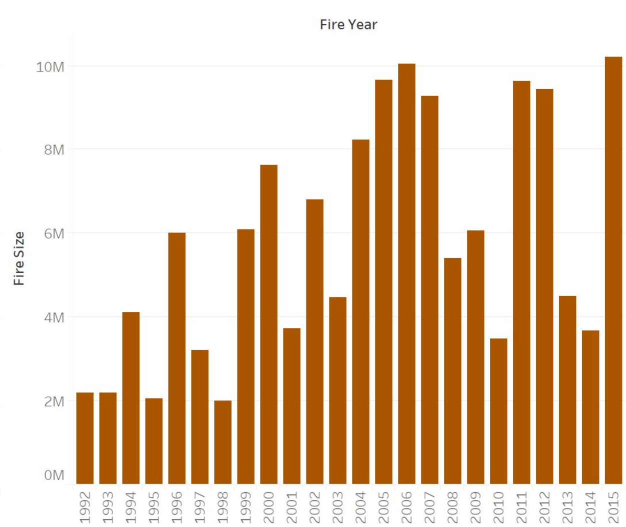
After analyzing the changes over the years we saw that there was a clear trend of increasing sizes and quantity of very large fires. So we decided to graph the average fire size of each year from 1992 to 2015. Again a clear upward trend can be seen in the graph down below. Analyzing this we see that the average did in fact grow from 37.6 acres in 1992-95, to 97.6 acres in 2011-15.

We created and once again saw similar trends for the total number of acres burned annually. That of an upward trend. When checked we see that the average number of acres burned annually grew from 2.6 million acres in 1992-95, to 6.9 million in 2011-15.

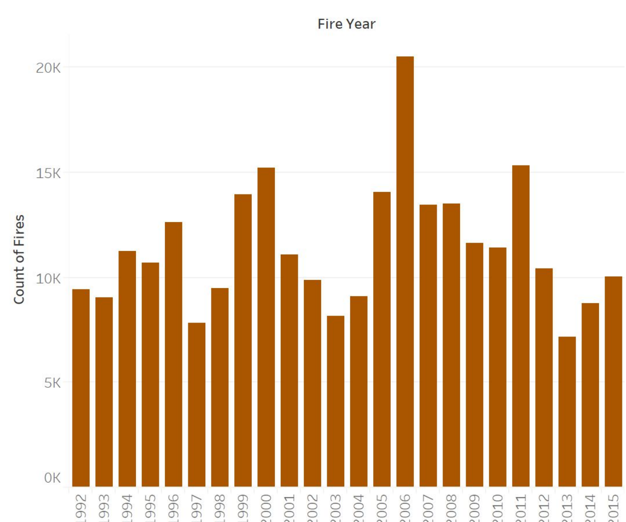
avg fire size per year



acres burned annually

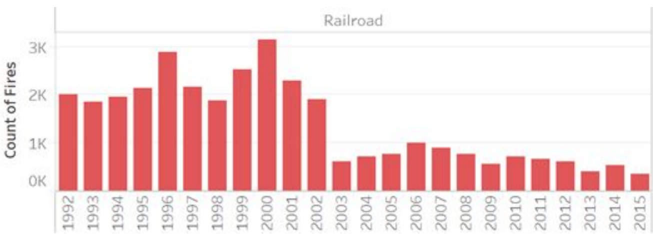


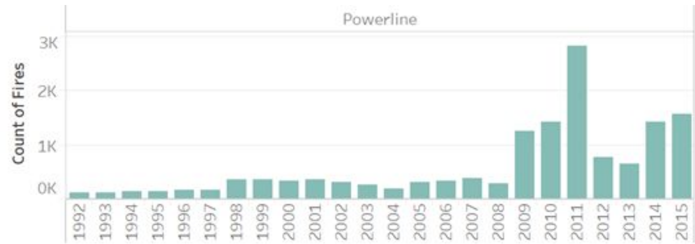
Number of annual fires

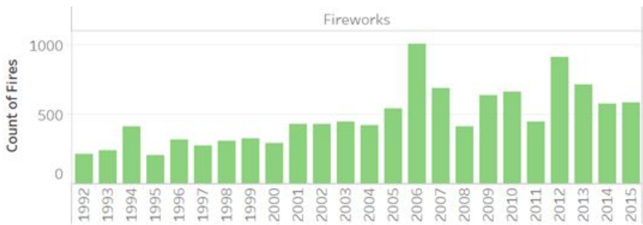


However, when we create a similar chart for the number of annual fires we can see that there is no clear trend and it has remained relatively steady. Because the average number of acres burned annually has increased steadily but the number of fires hasn’t, that either the average fire size has increased or the number of large fires has increased annually.

Next we decided to analyze each of the thirteen causes of fires by the number of fires, and acres burned, again annually. While there were some relatively clear trends like the dip in railroad caused fires in the early 2000's and increases in both powerline and fireworks caused fires (shown below) most changes in the grand scheme of things were on the order of thousands and the trends were big enough to lead us to believe they are the cause of the increase in acres burned.

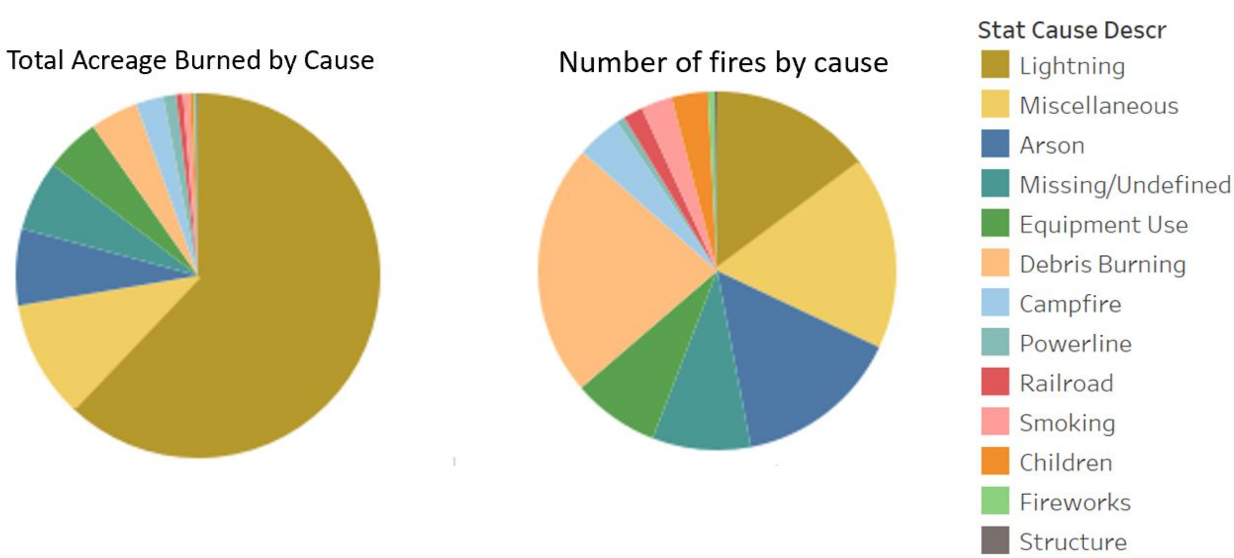




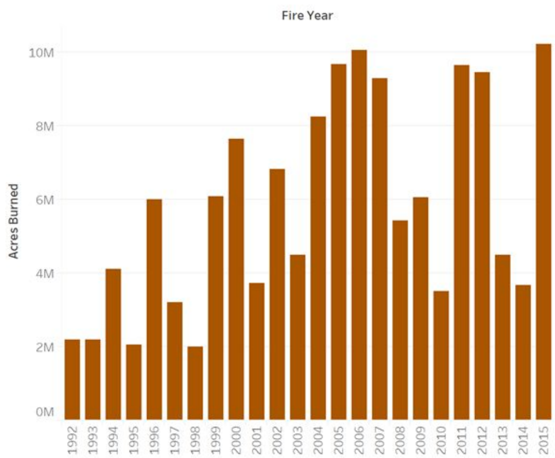


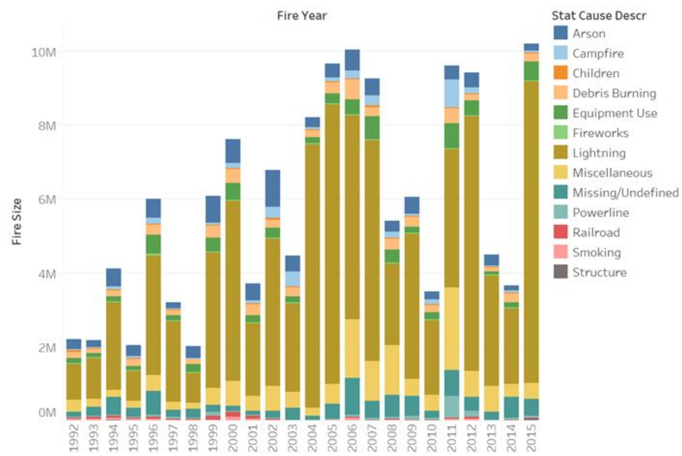
However when we again create color coded pie charts to represent the proportions of total fires and of acres burned according to the thirteen types of fire causes. We were surprised by just how much lightning was the leading cause for acres burned by wildfires. In those charts seen below, we see that lightning is responsible for nearly two-thirds of acres burned while only being the cause of about thirteen percent of reported fires. Two of the next three leading causes in both the number of fires and acres burned are miscellaneous and undefined which only further goes to show how big of an impact lightning really is.

From this we can posit that lightning from storms presumably in rural areas leads to fires which go longer without being reported, and which take much longer to bring firefighting resources too. This gives the wildfire much more time to grow and become difficult to control. This information is valuable and if we had more time, would be where we would next focus our research, to determine from these larger fires how many are in rural areas and what the conditions, for example temperature and humidity, leading up to the fires could have told us.



From this information we wanted to see if we could further emphasize just how big of an impact lightning caused fires were to acres burned annually. So we decided to revisit one of our previous visualizations. Specifically the bar graph representing total acres annually, we decided to take this graph and modify it to show the annual proportions of burned acres.When we did so, as can be seen below it became evident that the fluctuations and overall upward trend of acres burned by wildfires was primarily affected by lightning and the fires they cause. The amount of acres burned by every other cause remains pretty steady other than some slight deviations in miscellaneous and undefined fires. And for every spike or dip in the graph it can be almost completely accounted for by increases or dip in acres burned by lightning caused wildfires.





Related Work:

There are several related works in this field that we have looked at. The first of which was done by the NASA funded Rehabilitation Capability Convergence for Ecosystem Recovery (RECOVER). This group's study, which as with all others discussed will be linked in appendix B, focused solely on wildfires near the west coast. Their data consisted of 40,000 fires between 1950 and 2017. They found clear increases in both number and size of fires and as we began to suspect after we began to review our work, that in some areas these large fires repeatedly burned the same areas.

The second related work was done by Anthonly LeRoy Westerling of the US National Center for Biotechnology Information. In his study he again focuses on fires on the west coast and has found that there was a sharp increase in wildfires in the mid 1980s. He also found that a large portion of these fires were not in forested areas but in other habitats and that they are strongly associated with an earlier melting of snow in spring months.

And Finally the last related work we found was a study called Wildfire Statistics done by the US Congressional Research Service. In this study it was found that while the number of annual wildfires decreased from the 1990’s into the 2000s the average acreage burned more than doubled. and like our research showed lightning caused fires were the largest contributor to burned acres. However they found that the leading cause of individual wildfires was human caused.

Conclusion:

To conclude, we studied a publicly available dataset provided by the US Forest Service, and created many charts, graphs, and maps based on the data. Through the analysis of the data and these visuals we found that there has been an overall increase in the size of the average wildfire, while the number of wildfires has remained steady. We followed this information to find that most of the acres burned in Wildfires were done by a very small minority of wildfires.

By then analyzing the proportions of the causes of wildfires we found that the overwhelming majority of acres burned by lightning caused wildfires. And that while several of the causes for wildfires did have increases or decreases those of lightning remained pretty steady. This suggests that the increase of damage done by these fires is not because there are more of them but that they are growing much larger than before. Perhaps due to environmental climate changes. If we were to continue our project we would have liked to see how far these larger fires were from population centers and if that could be another leading factor.

One final thing we noticed and wanted to mention after reviewing our work and wanted to note was that it appears that most large class fires tend to be in the same areas. For example where most large fires tend to occur Alaska and near the west coast, the large wildfires are repeatedly in the same areas. This is something else that, had we more time, would warrant continued investigation and analysis. If for example the majority of the acreage burned is in a few concentrated areas, it would be a huge help as too where funding and prevention projects should be focused.

If we were asked to provide some ideas for watching out for future wildfires we came up with three possibilities for fire detection systems, so that fires can be detected and fought much earlier. The first is a satellite watching areas prone to wildfires like Alaska and the West Coast. By having satellite images of the geography we would be able to detect abrupt changes in areas and perhaps detect smoke going up into the sky. However this method would be susceptible to being blocked by clouds. Our second idea which is less expensive is having thermal cameras set up on posts throughout fire prone areas to detect the heat of fires and again report them as soon as possible. Our last idea was similar but even cheaper would be using similar posts but instead placing smoke detectors on them, however in this case wind could affect the ability to detect fires so more posts placed throughout would be required .

Bibliography:

Short, Karen C. 2017. Spatial wildfire occurrence data for the United States, 1992-2015 [FPA_FOD_20170508]. 4th Edition. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2013-0009.4

Desjardins, Jeff. “Visualizing 200 Years of U.S. Population Density.” Visual Capitalist, 12 Mar. 2019, www.visualcapitalist.com/visualizing-200-years-of-u-s-population-density/.

Alex. “US Population Growth Mapped.” *Vividmaps.com*, 2017, vividmaps.com/us-map-population-density/.

Andela, N., Morton, D.C., Giglio, L., Chen, Y., Van Der Werf, G.R., Kasibhatla, P.S., DeFries, R.S., Collatz, G.J., Hantson, S., Kloster, S. and Bachelet, D., 2017. A human-driven decline in global burned area. Science, 356(6345), pp.1356-1362.

Colombaroli, Daniele, and Sophie C Lewis. “Search By Topic.” Future Earth, 1 Dec. 2020, futureearth.org/publications/issue-briefs-2/global-fires/.

CBS News. “Brazil's Bolsonaro Says He Will Accept Aid to Fight Amazon Fires.” *CBS News*, CBS Interactive, 28 Aug. 2019, www.cbsnews.com/news/amazon-wildfires-brazil-spurns-20-million-aid-offer-from-g-7-nations-today-2019-08-27/.

N/A, N/A. “Wildfire Statistics.” *Fas.org*, 4 Dec. 2020, fas.org/sgp/crs/misc/IF10244.pdf.

Patel, Kasha. “Six Trends to Know about Fire Season in the Western U.S. – Climate Change: Vital Signs of the Planet.” *NASA*, NASA's Earth Observatory, 5 Dec. 2018, climate.nasa.gov/blog/2830/six-trends-to-know-about-fire-season-in-the-western-us/.

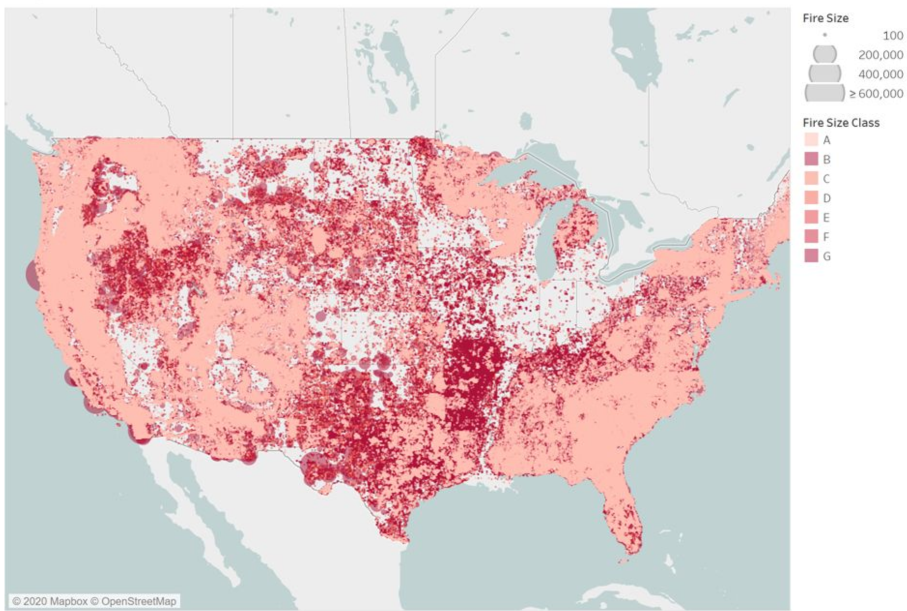
Westerling, Anthony LeRoy. “Increasing Western US Forest Wildfire Activity: Sensitivity to Changes in the Timing of Spring.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, The Royal Society, 5 June 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC4874415/.

Appendix A:

GitHub Link for Powerpoint and Tableau files.:
https://github.com/zepedaaldo/Cpt_S-475-Final-Project.git

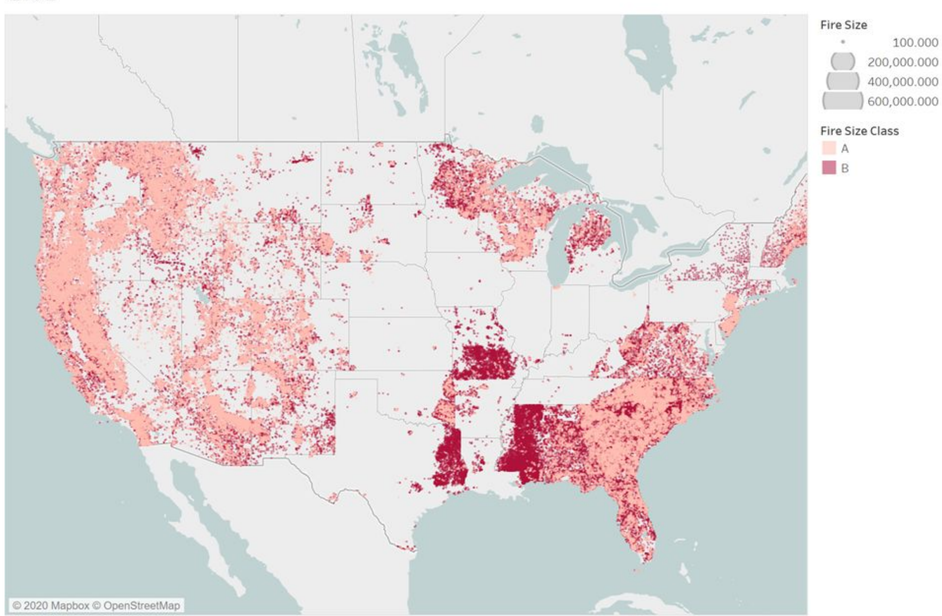
Appendix B:

original



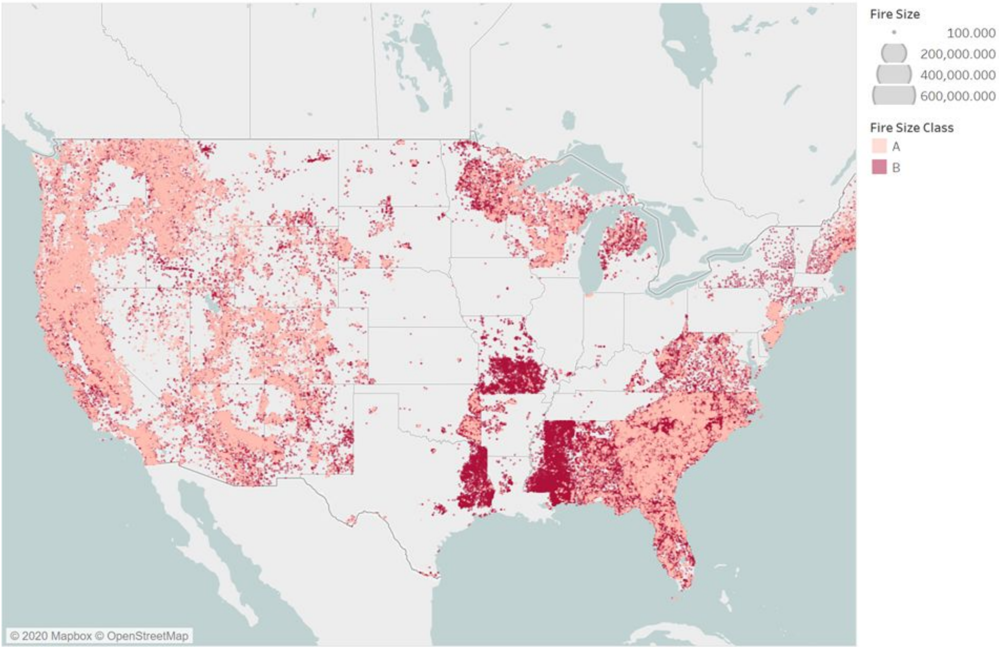
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute.

1992



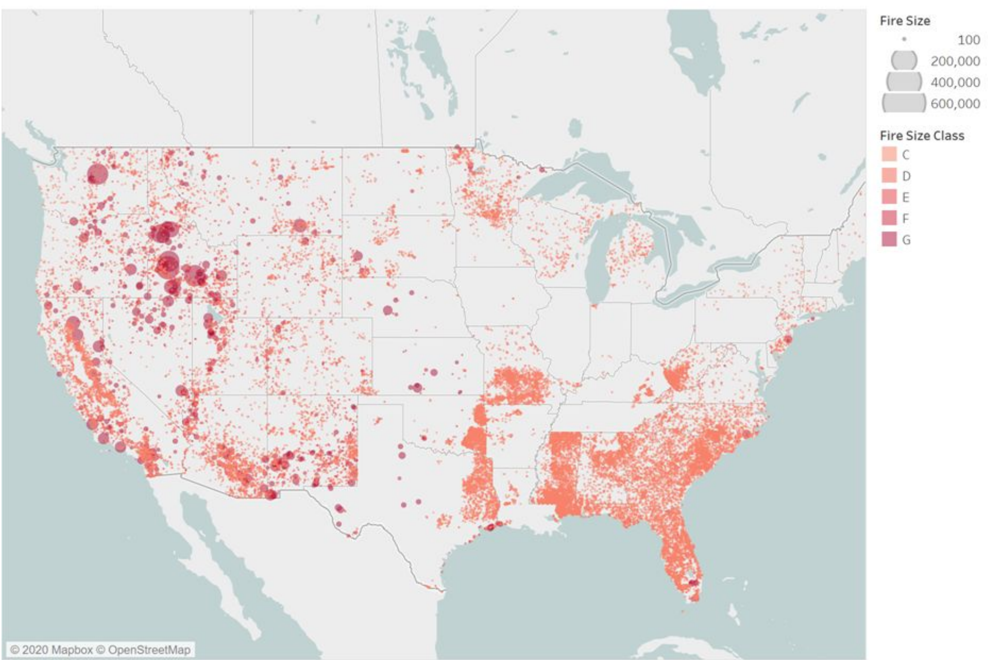
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 1992, 1993, 1994 and 1995. The view is filtered on Fire Size Class, which excludes C, D, E, F and G.

1992-1995 A-B



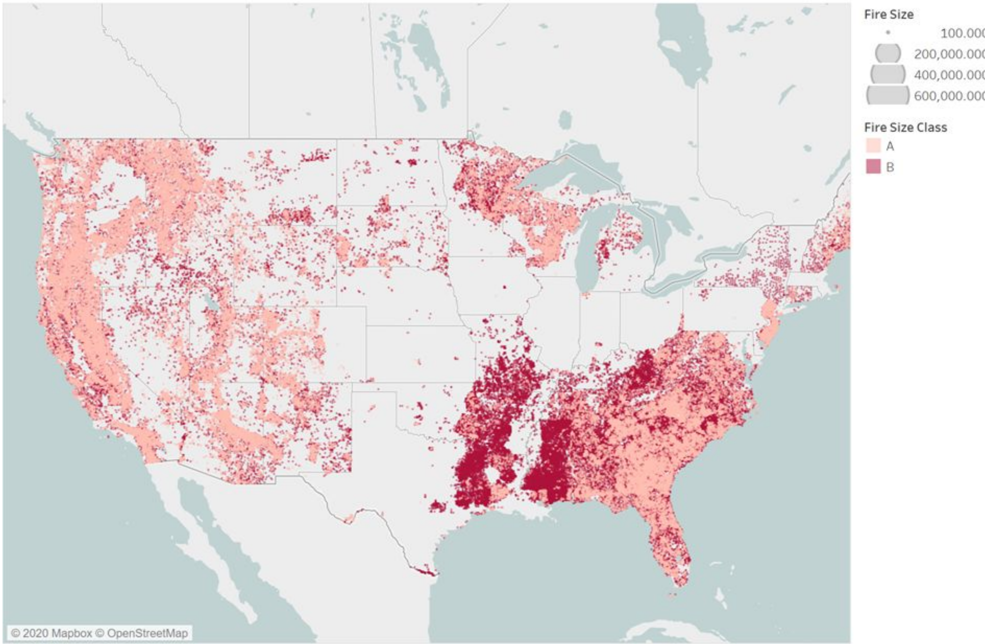
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 1992, 1993, 1994 and 1995. The view is filtered on Fire Size Class, which excludes C, D, E, F and G.

1992-1995 C-G



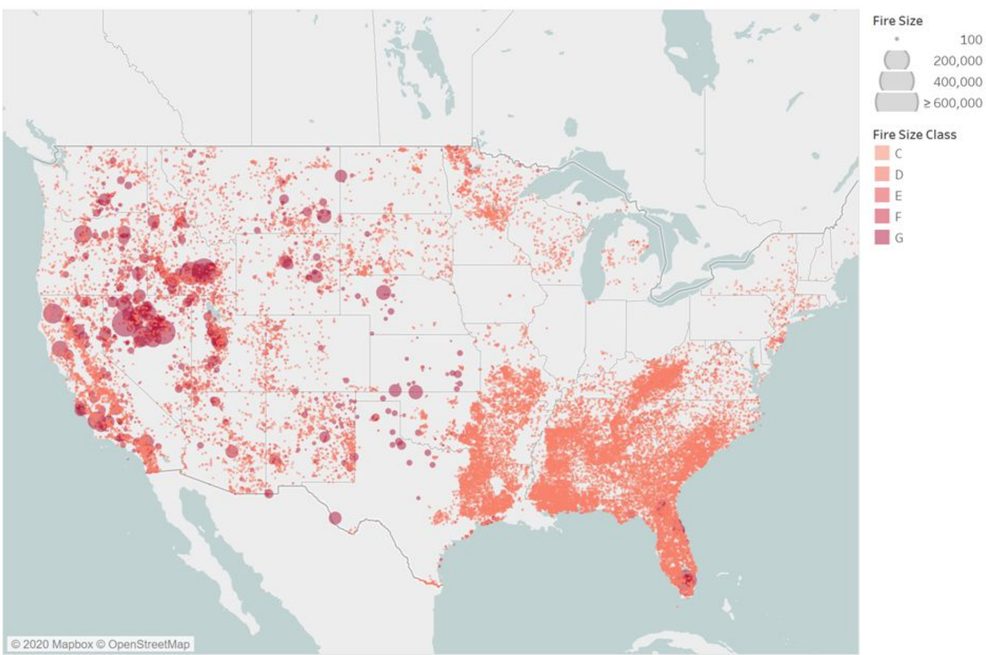
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 1992, 1993, 1994 and 1995. The view is filtered on Fire Size Class, which keeps C, D, E, F and G.

1996-1999 A-B



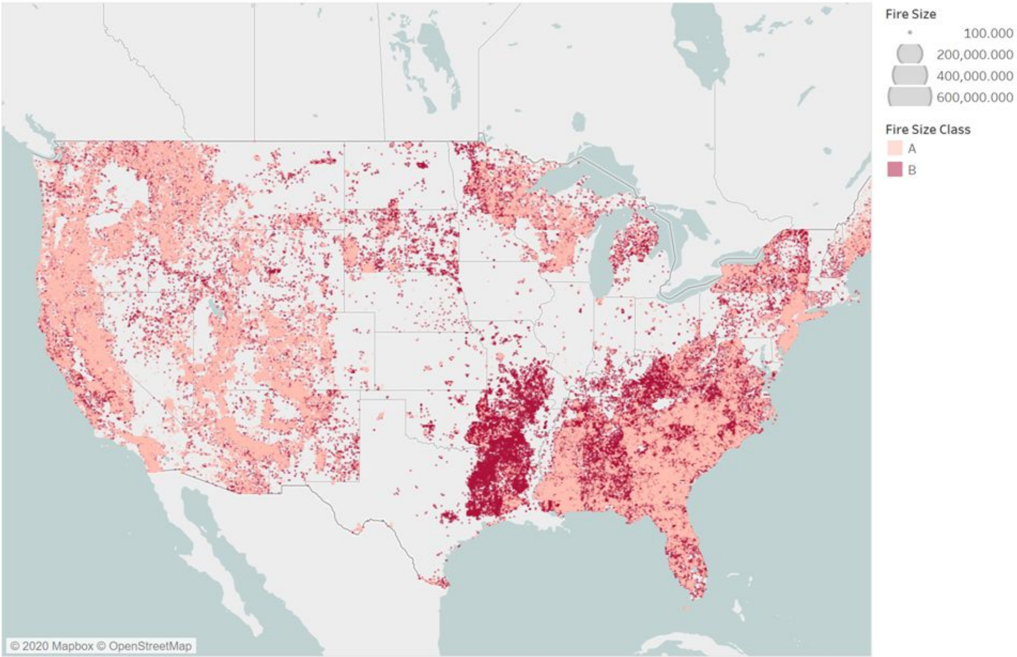
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 1996, 1997, 1998 and 1999. The view is filtered on Fire Size Class, which excludes C, D, E, F and G.

1996-1999 C-G



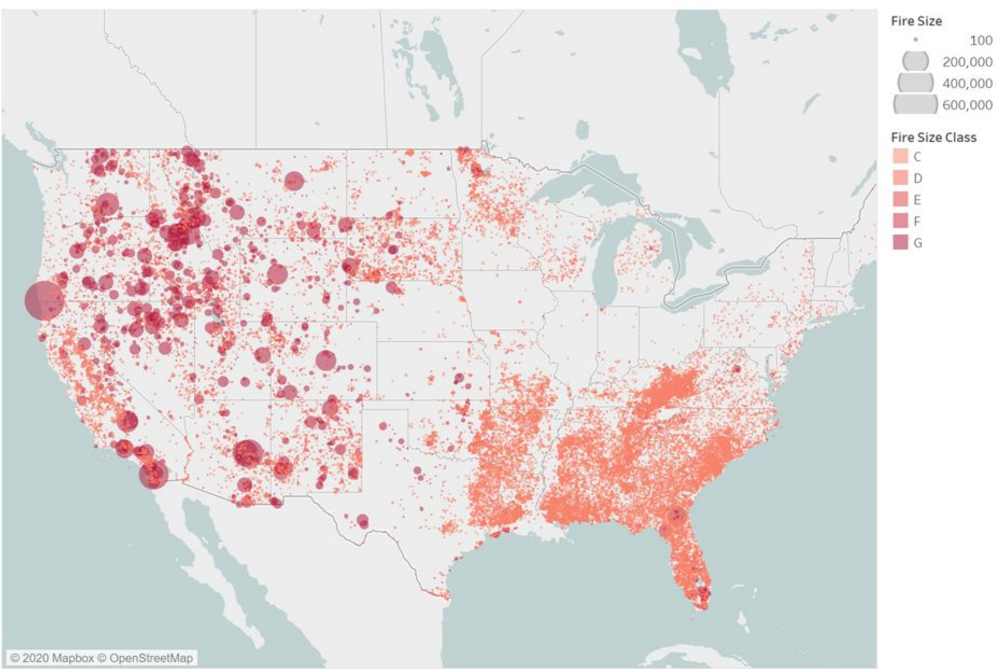
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 1996, 1997, 1998 and 1999. The view is filtered on Fire Size Class, which keeps C, D, E, F and G.

2000-2003 A-B



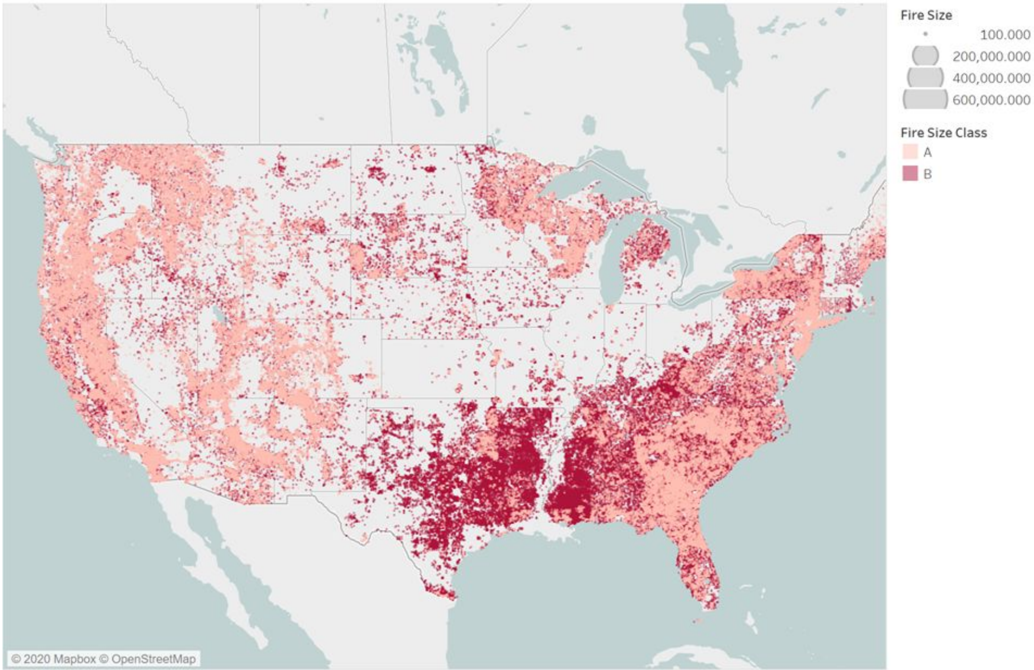
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 2000, 2001, 2002 and 2003. The view is filtered on Fire Size Class, which excludes C, D, E, F and G.

2000-2003 C-G



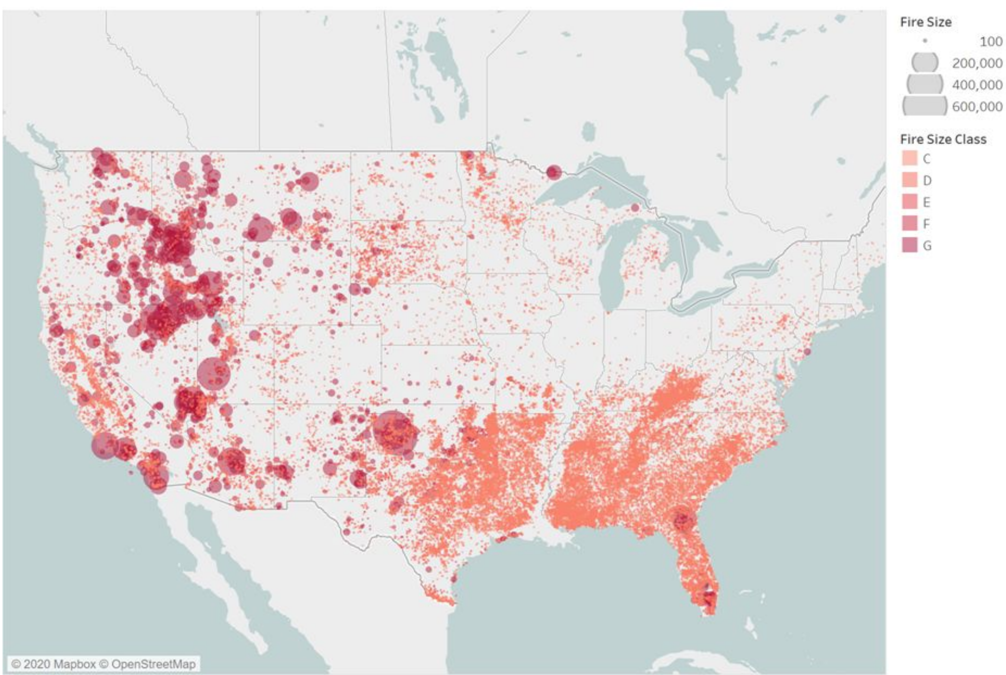
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 2000, 2001, 2002 and 2003. The view is filtered on Fire Size Class, which keeps C, D, E, F and G.

2004-2007 A-B



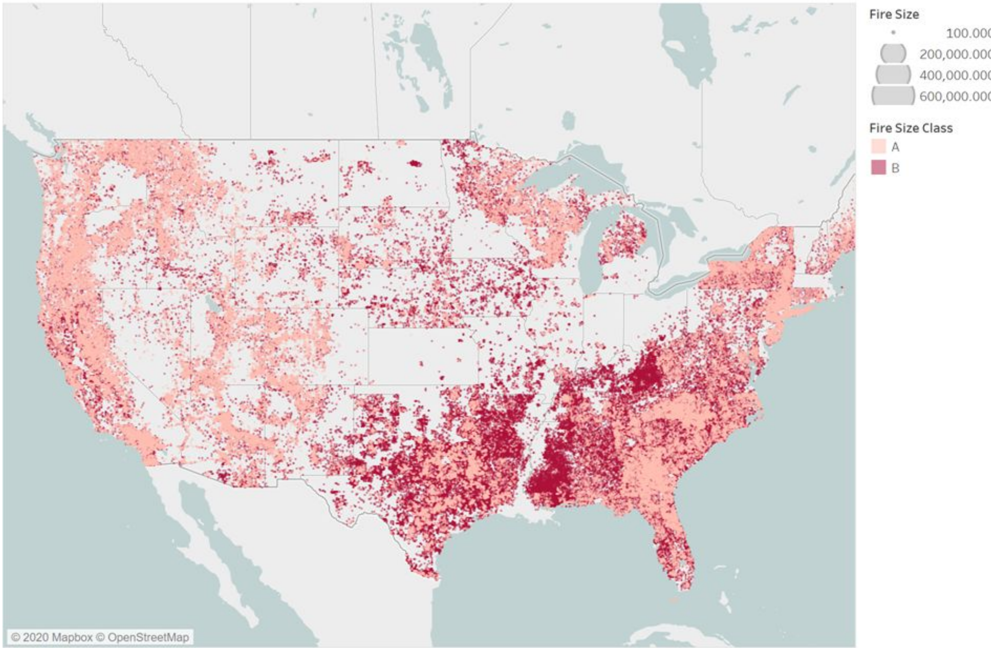
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 2004, 2005, 2006 and 2007. The view is filtered on Fire Size Class, which excludes C, D, E, F and G.

2004-2007 C-G



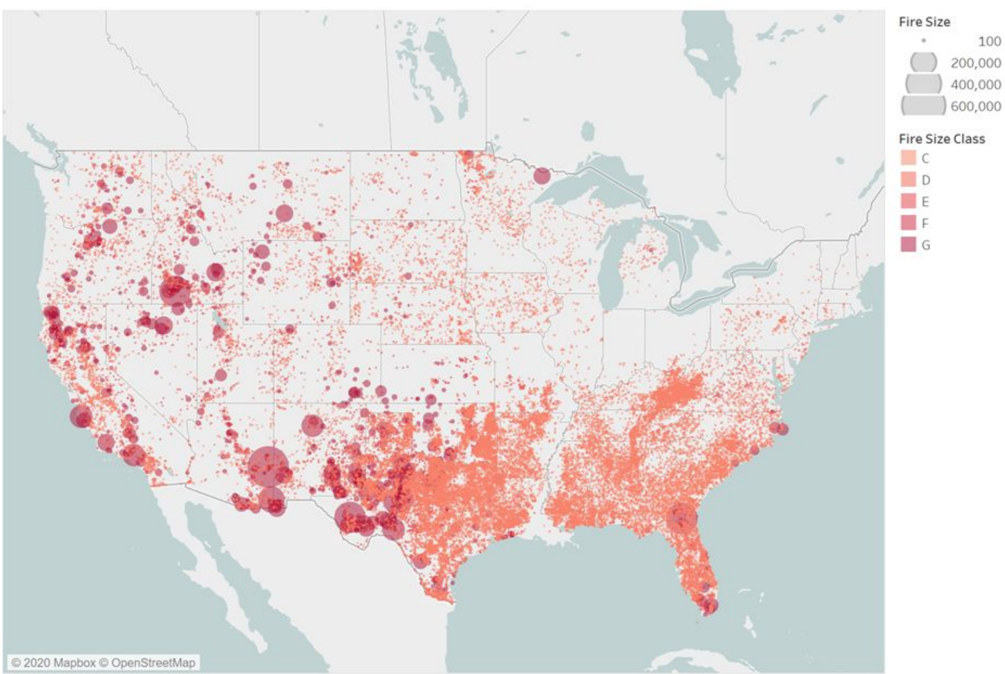
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 2004, 2005, 2006 and 2007. The view is filtered on Fire Size Class, which keeps C, D, E, F and G.

2008-2011 A-B



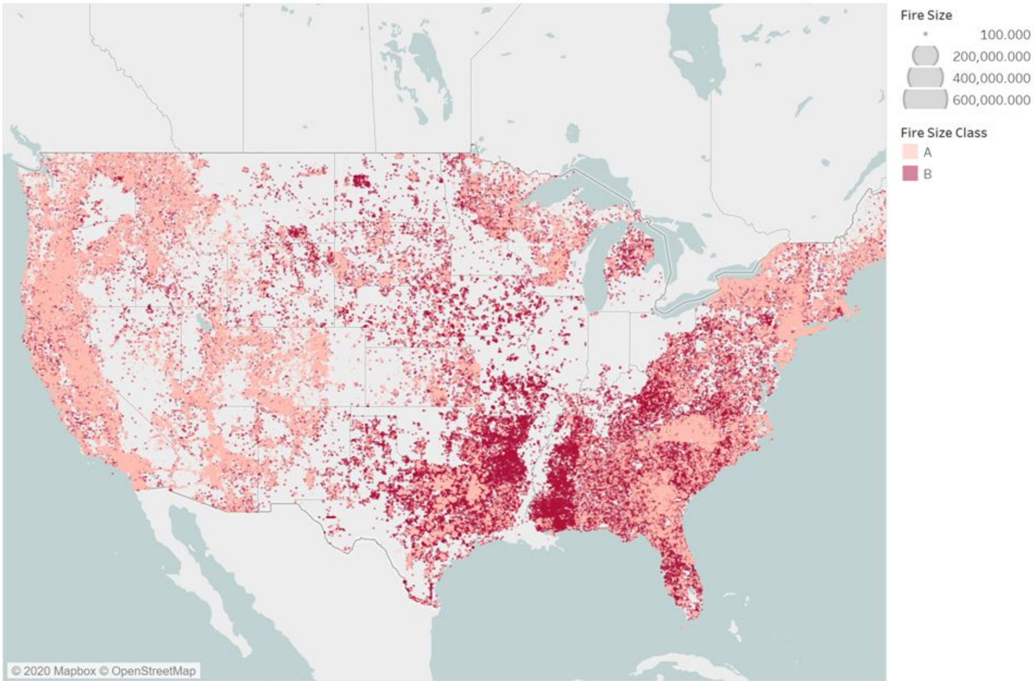
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 2008, 2009, 2010 and 2011. The view is filtered on Fire Size Class, which excludes C, D, E, F and G.

2008-2011 C-G



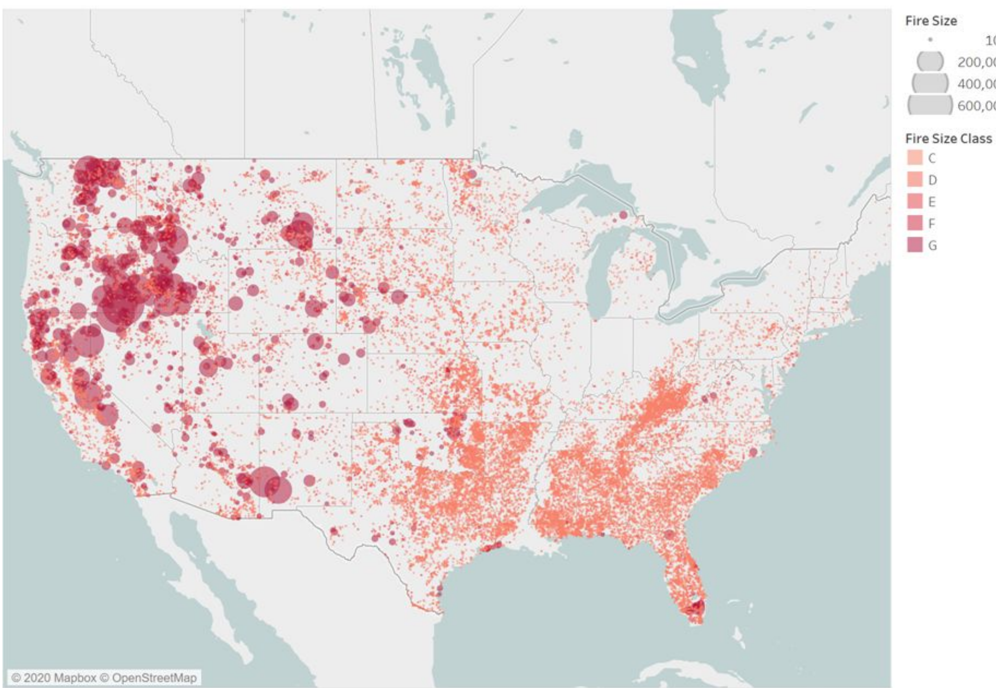
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 2008, 2009, 2010 and 2011. The view is filtered on Fire Size Class, which keeps C, D, E, F and G.

2012-2015 A-B



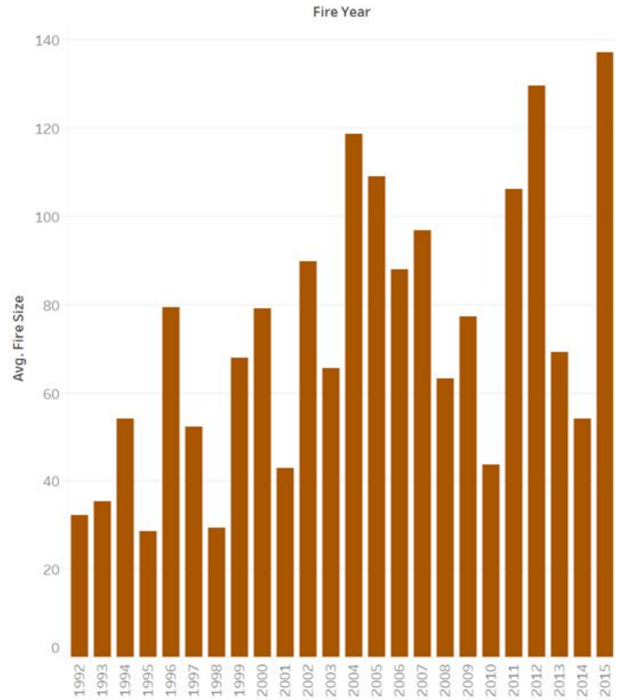
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 2012, 2013, 2014 and 2015. The view is filtered on Fire Size Class, which excludes C, D, E, F and G.

2012-2015 C-G



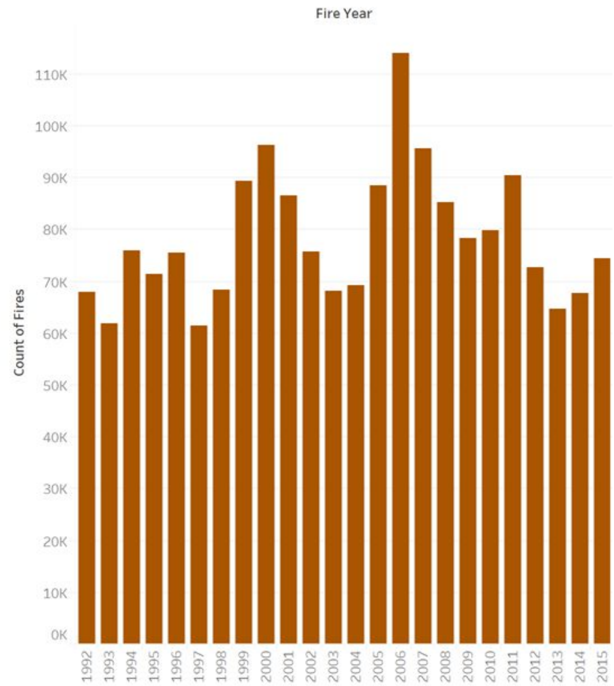
Map based on Longitude and Latitude. Color shows details about Fire Size Class. Size shows Fire Size as an attribute. The data is filtered on Fire Year, which keeps 2012, 2013, 2014 and 2015. The view is filtered on Fire Size Class, which keeps C, D, E, F and G.

avg fire size per year



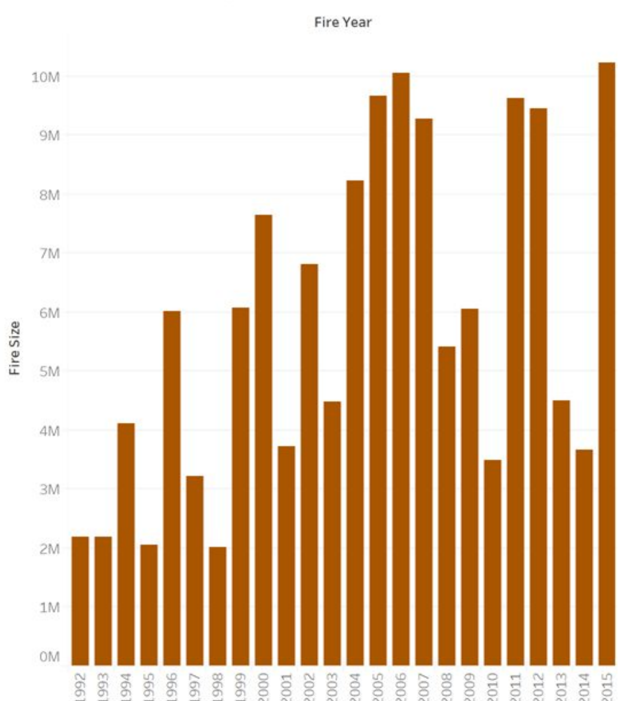
Average of Fire Size for each Fire Year.

Number of annual fires



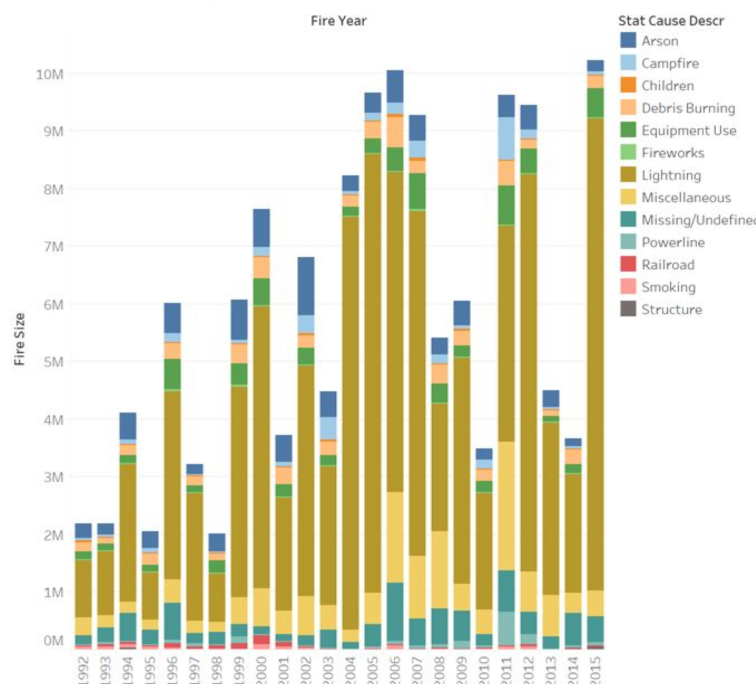
Count of Fires for each Fire Year.

acres burned annually



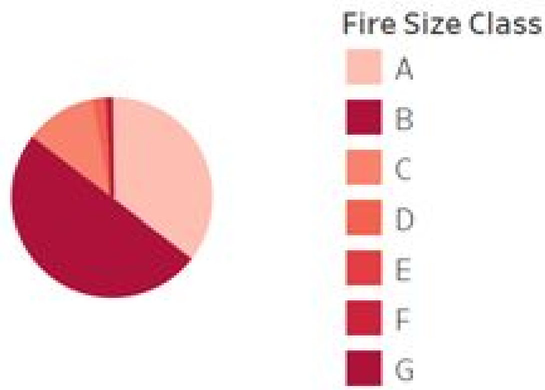
Sum of Fire Size for each Fire Year.

Acre burned annualy with causes



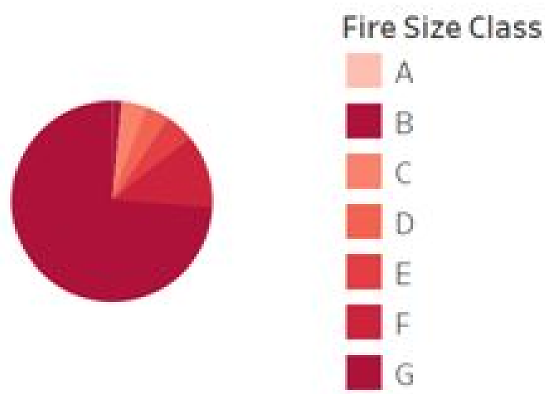
Sum of Fire Size for each Fire Year. Color shows details about Stat Cause Descr.

number of fires
Pie Chart



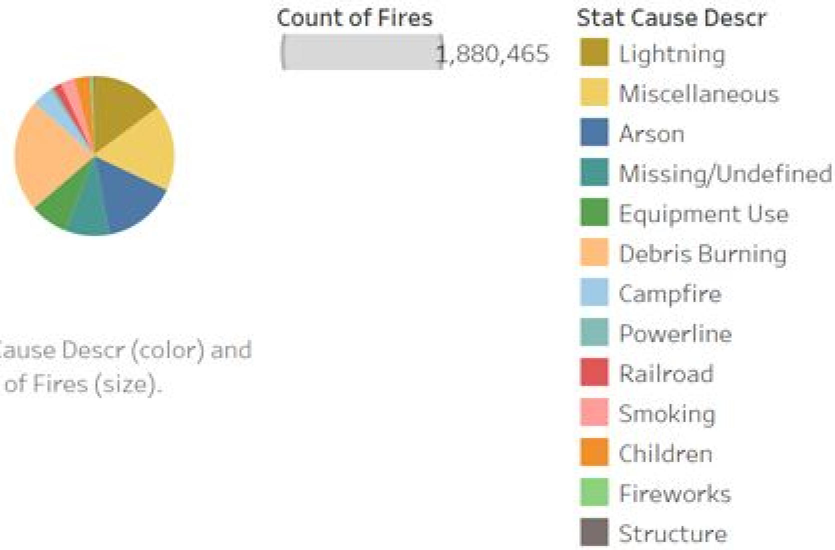
Fire Size Class (color).

Acres Burned
Pie Chart



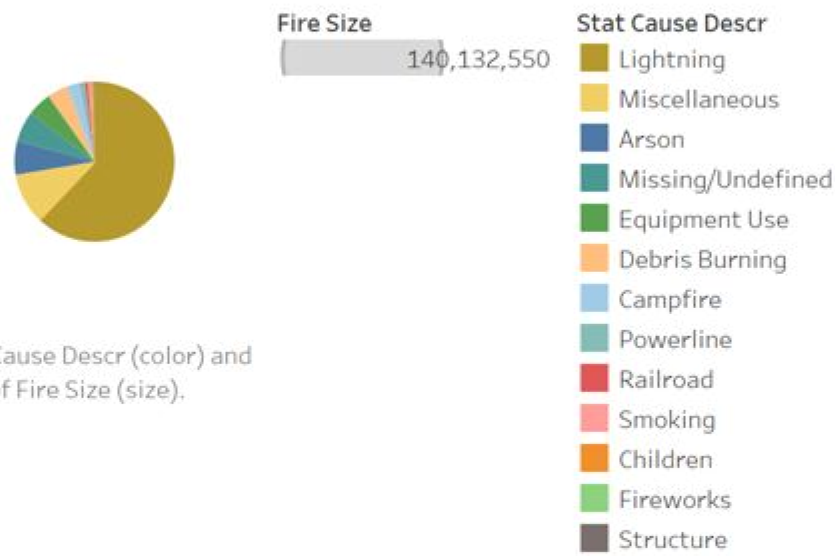
Fire Size Class (color).

of fires by cause



Stat Cause Descr (color) and count of Fires (size).

acres burned by cause



Stat Cause Descr (color) and sum of Fire Size (size).