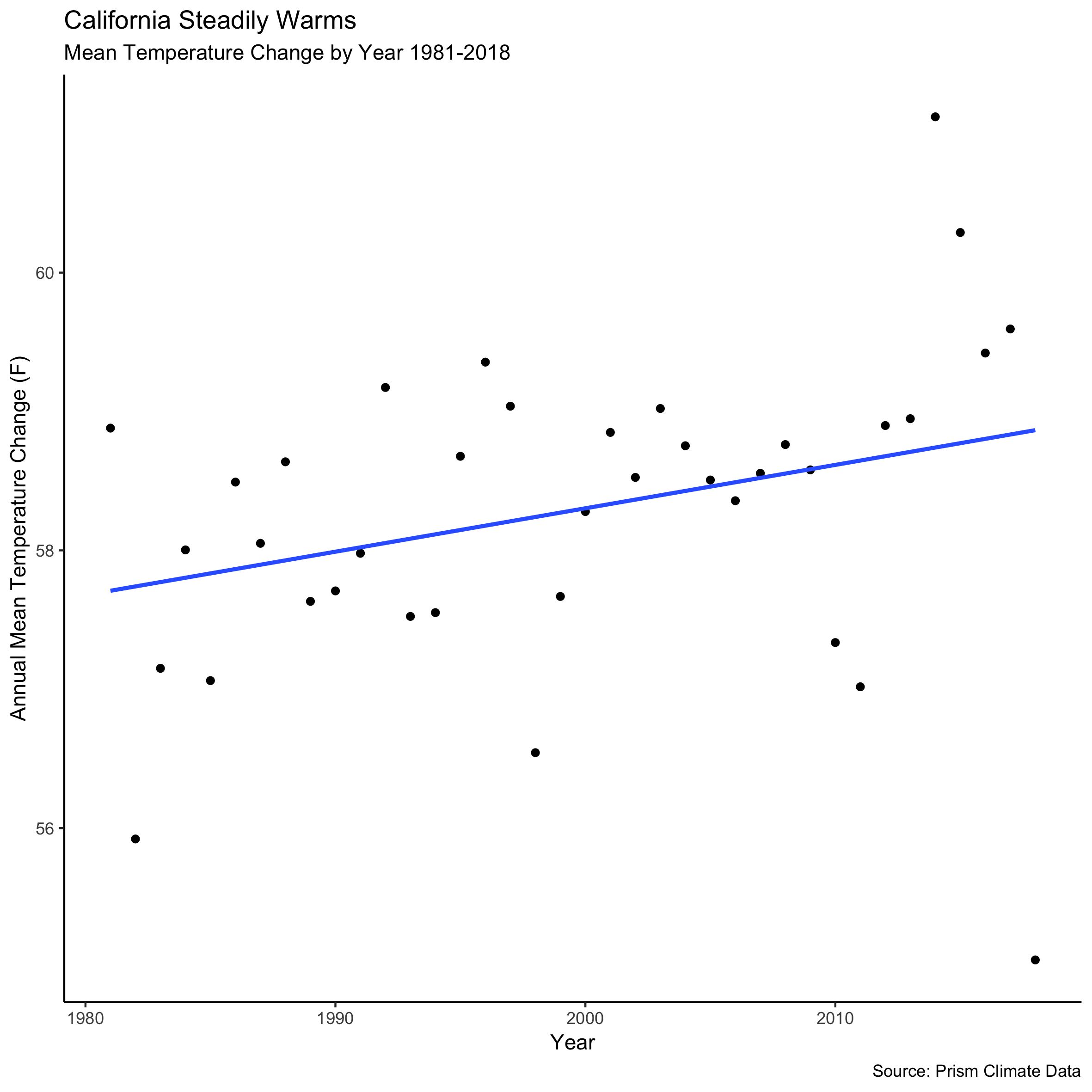
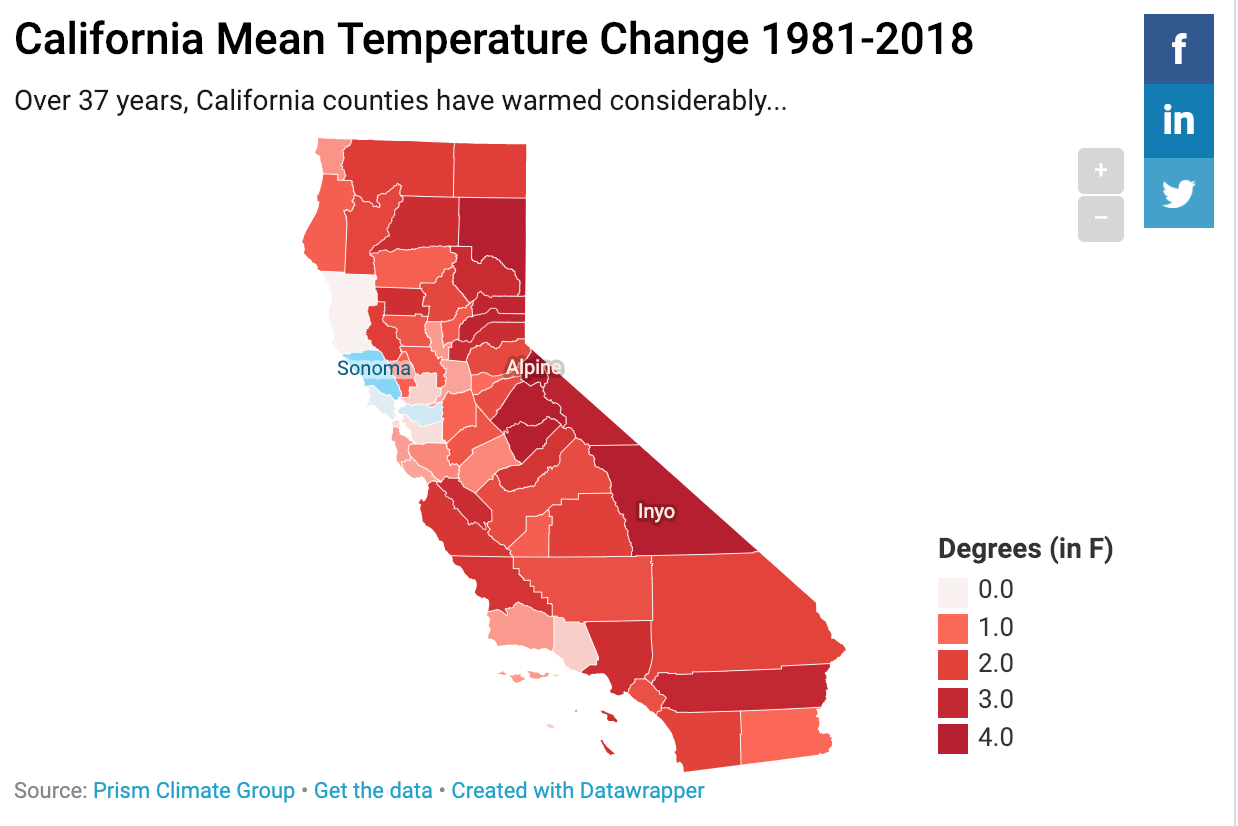
Climate Indicators

* Visual: Overall change -
  + Text: What is the change

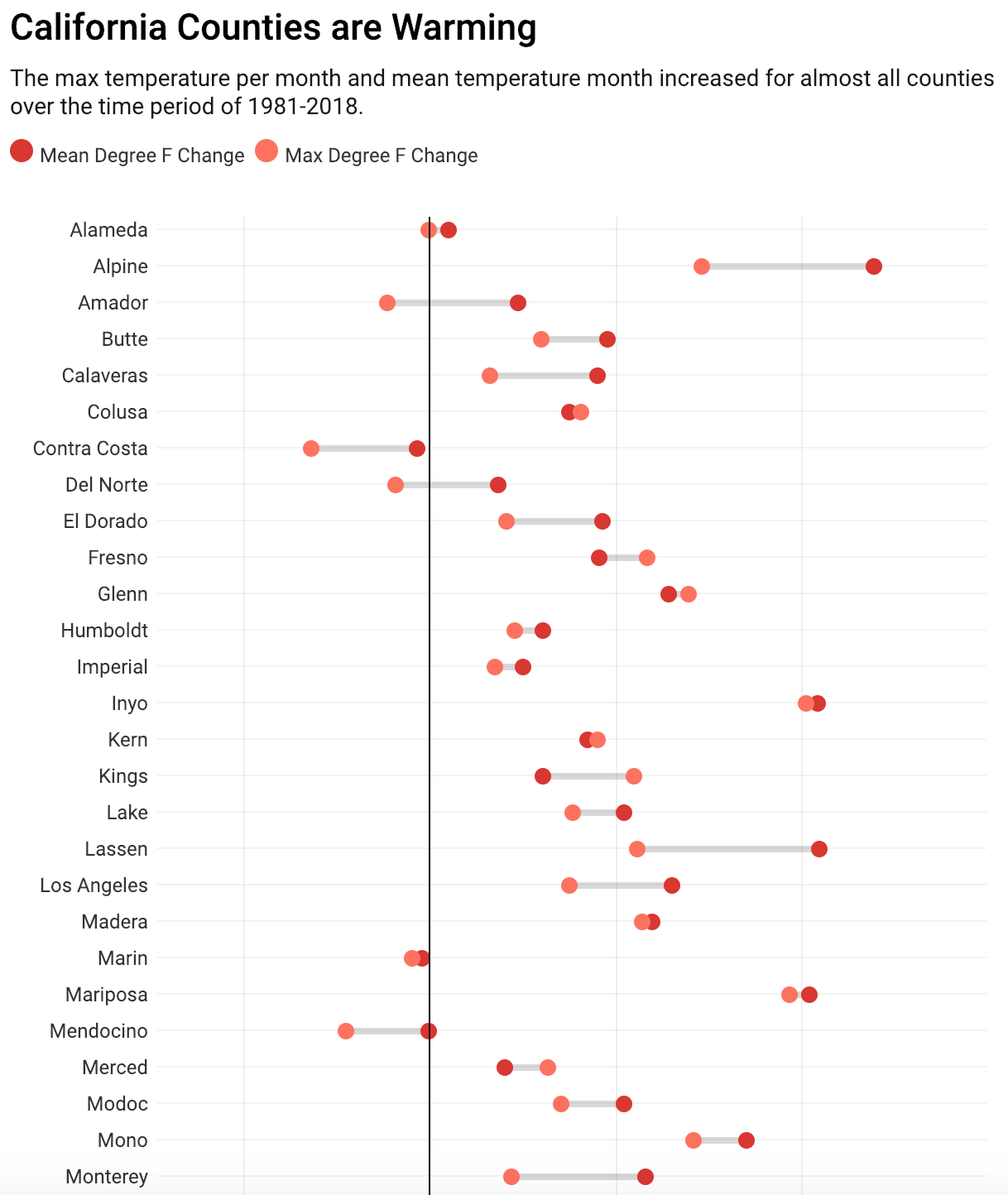
California is getting warmer overall. While we only have a snapshot of data, we can see the gradual change. Normalizing this by county, we observe a mean change of 1.16 degrees F for the time period. This effect is statistically significant at the 10% level.

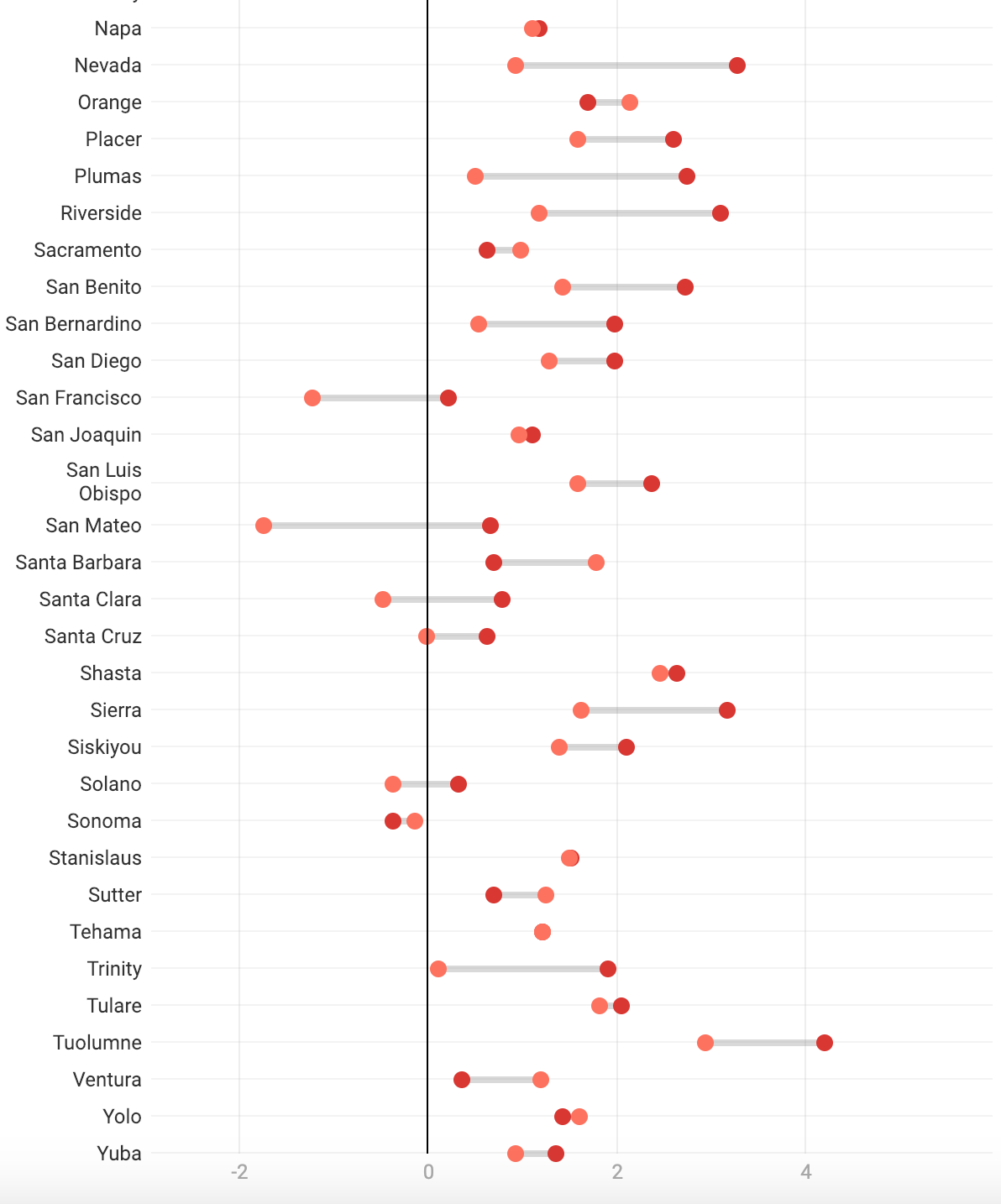


However, this overall trend hides substantial heterogeneity as some counties and some months have seen larger changes than others.



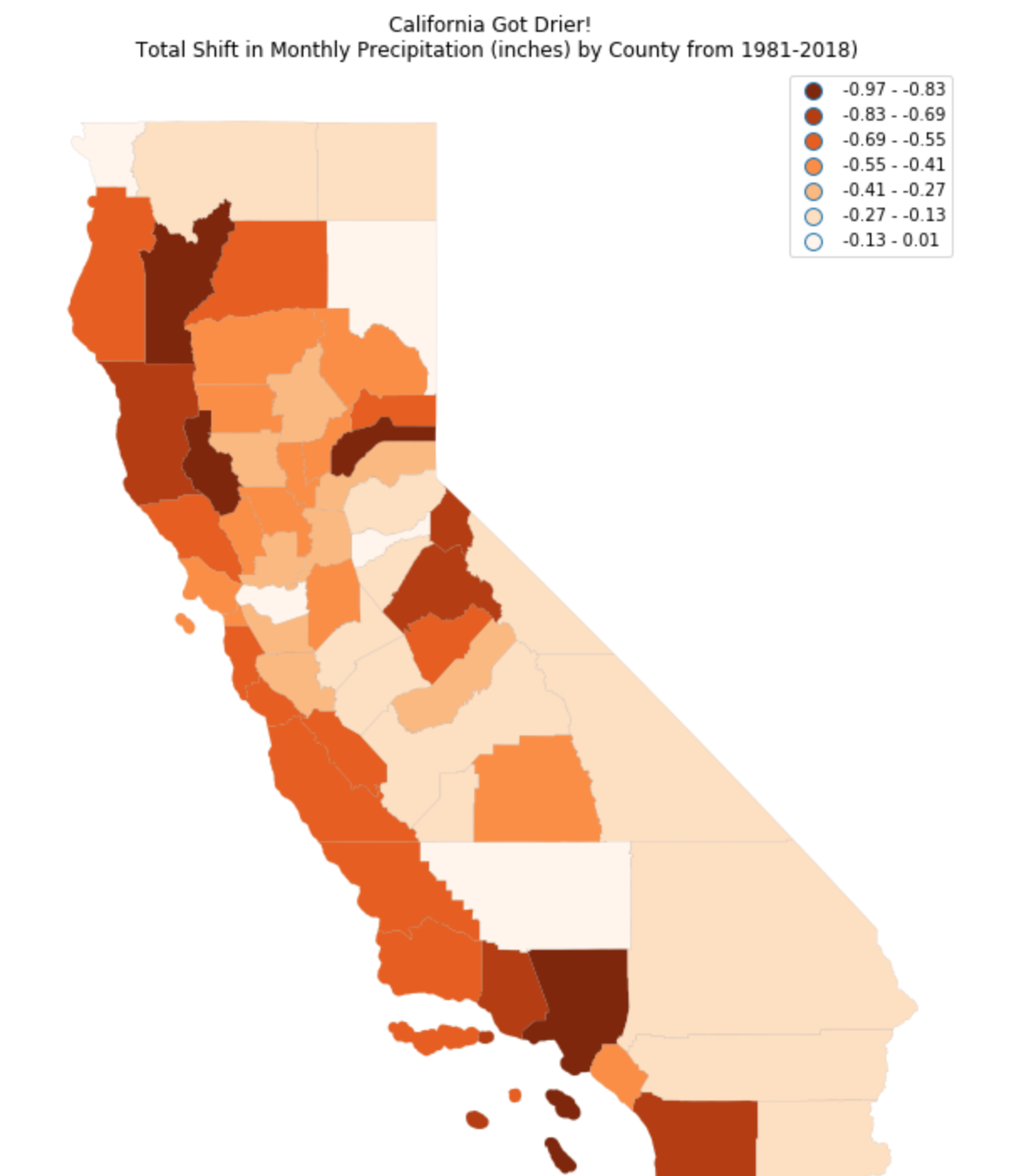
This interactive map shows how some counties, like Inyo have seen over 4 degree average temperature changes in this time period, other counties like Sonoma have seen a small decrease in average temperature.





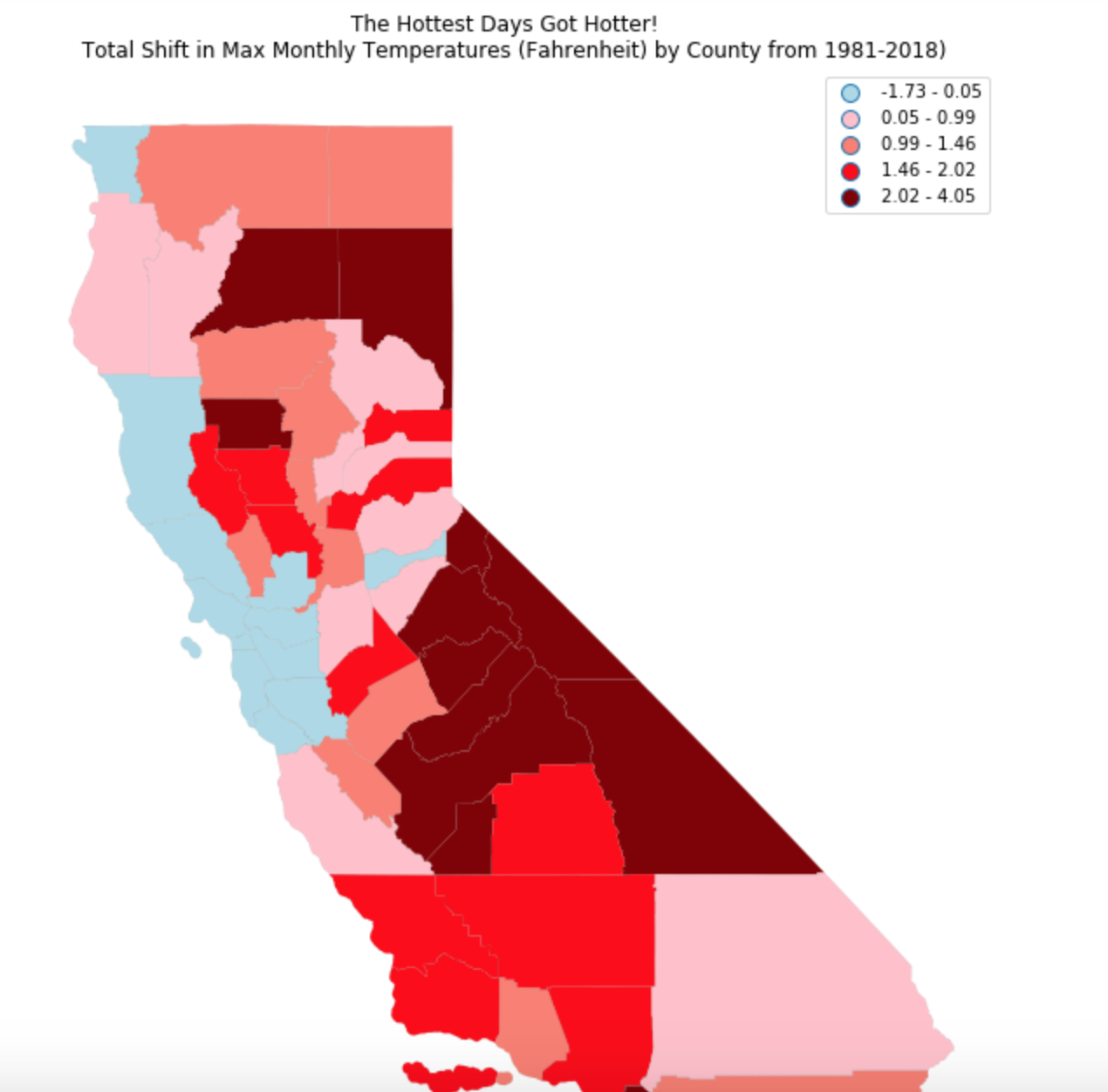
This visual shows even more clearly how different counties have experience different levels of temperature changes. Climate change research indicates that extreme temperatures (very hot days) will become more common as well as the mean temperature increasing. This graphic shows that the hottest day per month for different counties also got hotter on average across the time period, but this shift was not in perfect lockstep with the shift in the mean temperature. Other visuals can help us make sense of this finding.

* Visual: Counties on map interactive
* Visual: Other indicators on map



While California got drier during the time period, this trend is less noticable than the change in heat (though almost universal across counties). However, one can see that places like Los Angeles county, which receives very little rainfall (**get yearly average for the county)**

Climate change does not inherently led to drier climates, but hotter temperatures means more evaporation and that the air can store more water vapor before it must release the water in precipitation events. These means that storms are often more intense when they happen. In addition, the change to the climate has interacted with other climate phenomena in a way that has exacerbated drought in some prone areas. For example, California has suffered recent droughts that have been linked to climate change (**cite**).

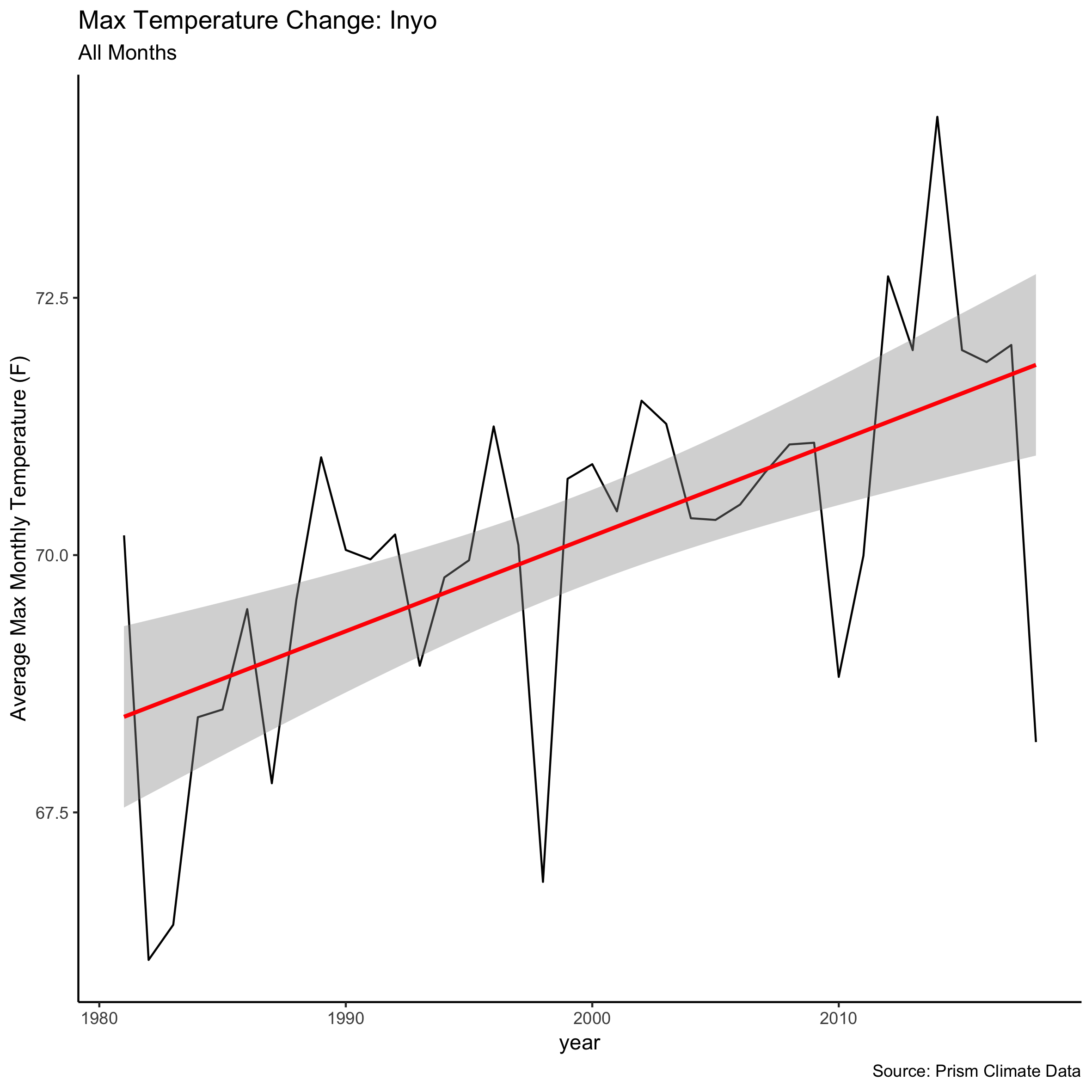


The extreme heat map perhaps best exposed the substantial heterogeneity across California. Inland and southern counties experienced much hotter maximum monthly temperatures, while some counties around the bay area actually saw their hottest temperatures decrease. Following the Mark Twain quote “The coldest winter I ever spent was a summer in San Francisco”, one reason for this surprising trend could be due to fog. The Bay Area receives a lot of fog during high temperatures which is a natural safeguard against high temperatures. Purely speculatively, one could imagine that with a warmer ocean and more extreme heat, that this area is seeing the cooling mechanism of fog get set off more often on would be max heat days.

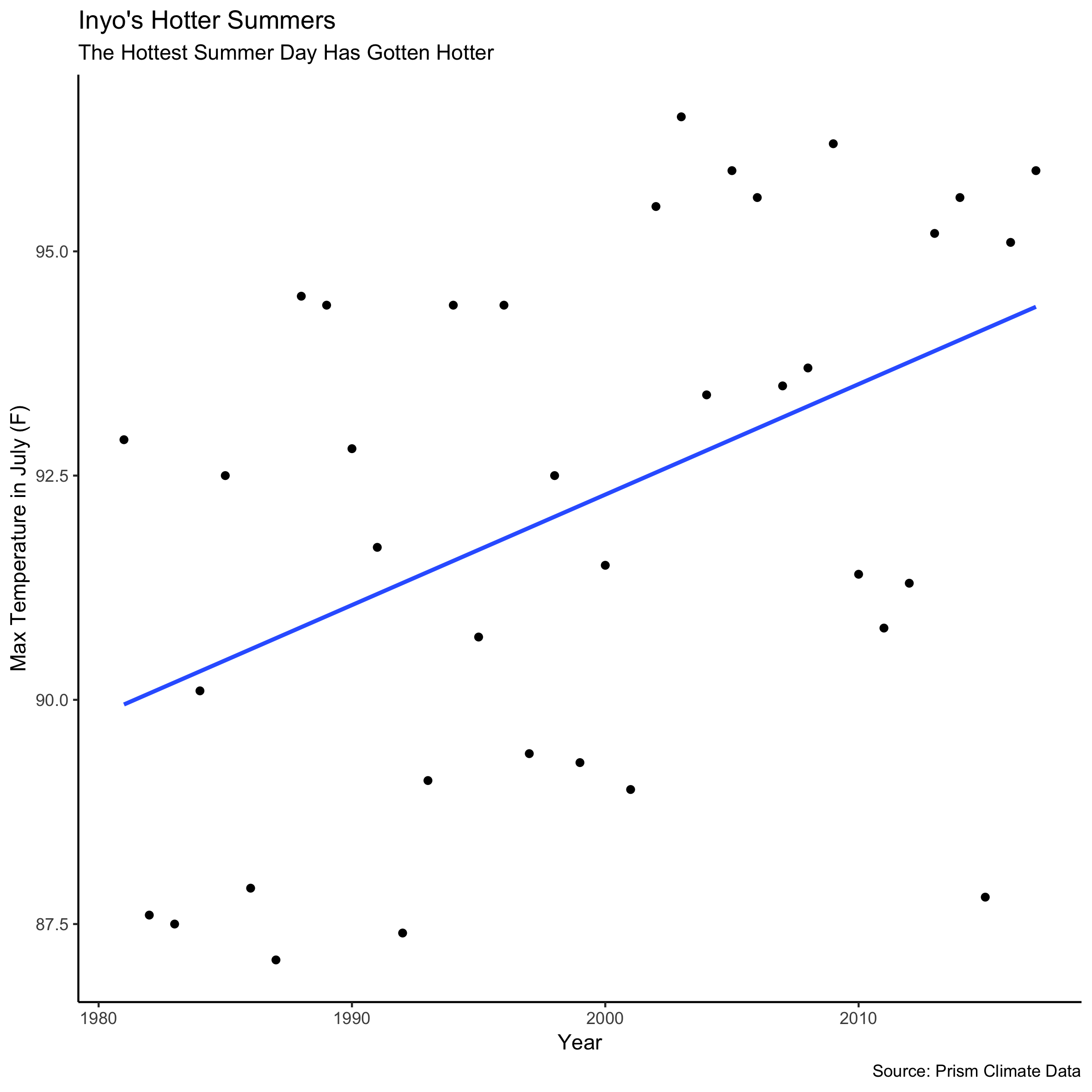
However, San Francisco, where we have seen slight cooling of max temperatures, has still warmed a little on average and is not one of the counties that ever had max temperatures in the summer that put its citizens at risk. To understand the impact of the change to max temperature it is worth exploring a case study.

* + Text: Variance of change
  + Text: Where is change happening and heterogeneity
* Visual: Highlighted Case - pretty D3
  + Text: Why it matters

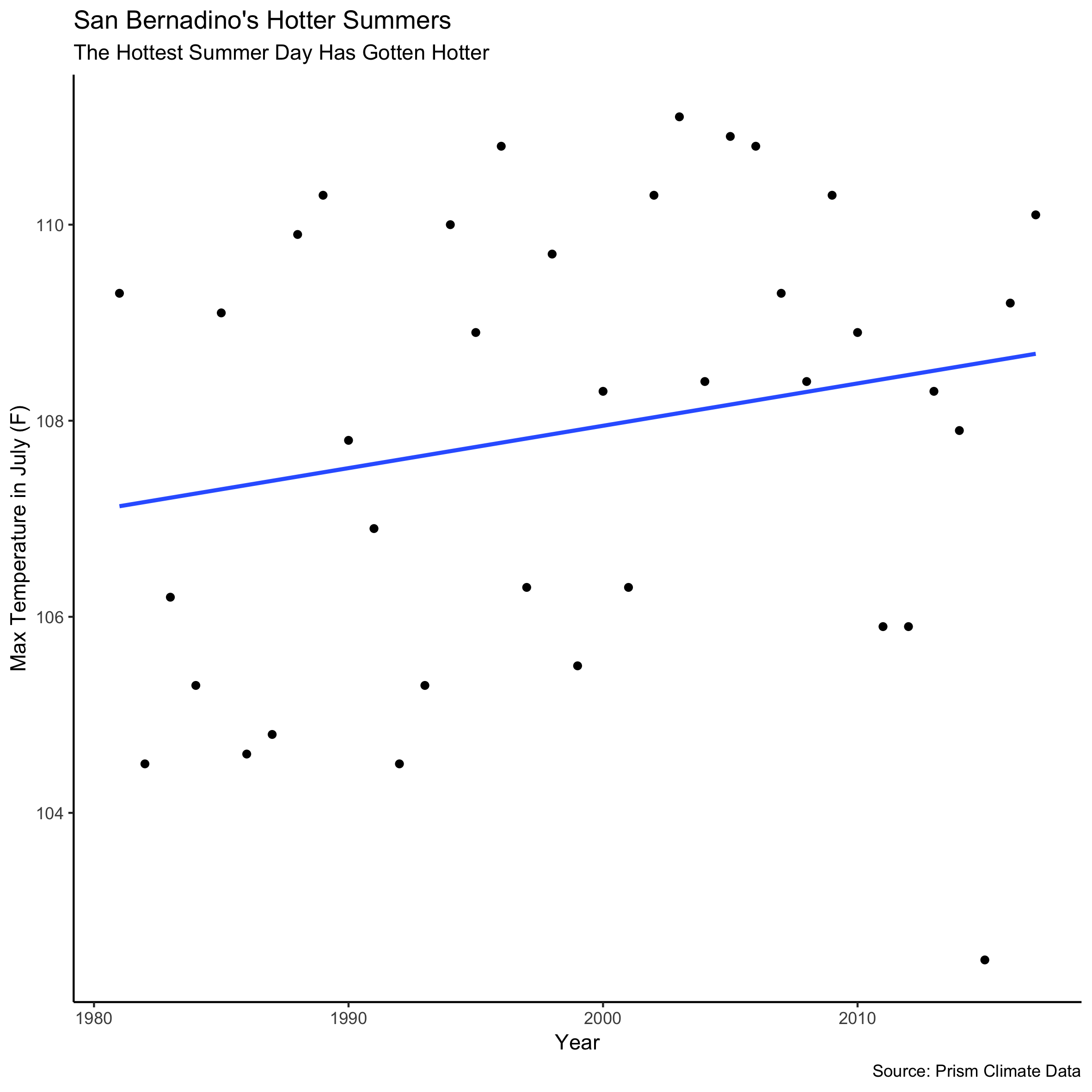
‘



The max temperature in Inyo increased by about 4 degrees Fahrenheit over the time period. At first glance, this may look like a minor change from the high 60s to low 70s, but this has also led to hotter summers.



In July, for instance, the max temperature which hovered around 90 degrees now regularly sits closer to 95 degrees. Already, 95 degree temperatures could pose problems for less healthy individuals, but this change may



README.txt for climate indicators section

I found the data from PRISM climate group at Oregon State during my search for high quality climate data. I saw that it had data on max temperature, mean temperature, min temperature, and precipitation down to the county level in California for every month since 1981. However, I also observed that it was not easy to download the data in bulk.

The first step I took was to see if I could use wget from the command line to download the data and append it together. I built the getprism.sh file to do that, but realized that there was a middle part of the url that changed for each data download before the csv such that using special characters {} to download multiple counties no longer worked. For that reason, we divided the 58 counties to our group members and each downloaded some. Once the files were downloaded, I was able to append them using the command line (using getprism.sh), but found later that it was better to use STATA. The main reason for this choice was that the data for each county came with latitude and longitude coordinates but did not list the county name so I had to back into that link later (I also tried to spatially join the latitude and longitude to shapefiles, but had unresolvable issues related to the projection – more on that later).

Append and Data Cleaning with STATA:

* Once I had 58 csvs (1 per county) with the raw climate indicators (mean temp in month, max temp in month, precipitation in inches for month, and min temp in month) for all months over the time period of 1981-2018, I needed to read them into a single file
* I created a loop to read each of the 58 csv files and used a local variable to attach the number in which the file was read as an idnum column (the first file had an idnum equal to 1, second file idnum ==2, etc.)
  + This loop created a STATA data file for each of the input csvs
* Given how we divided the download of files using wget between the team members, I had a list of the file names that was (if order could be preserved) in alphabetical order
* Next, I found an alphabetical list of California counties online and created a matching key of ascending numbers
* In the STATA file I created a county variable and I filled that county with appropriate name based on the matching list. I used the alt trick you showed us in class in Jupyter Notebook cells to create STATA code (replacing county if idnum == 1, etc.). There isn’t a specific record of this because in the STATA code this results in 58 lines of “if idnum==1, replace = “county name”, but this wasn’t as repetitive as it looks.
* This was the main cleaning issue, but I also renamed variables and changed the date to the appropriate date filetype

Create the change variables:

* The dataset had mean and max temperatures, but it did not tell us how these things changed over time
* I consulted a professor and learned that I could get a measure of the monthly change by county by running a regression where I would regress time (year) on the various indicators (max\_temp, mean\_temp, precipitation) for each county and month
  + This would tell me how these indicators changed each year for this locality and month
* At this point, I ran into an issue with my matrix size (my version of STATA was not powerful enough to run this operation), so I divided this process into 6 stata files where I ran the analysis code on 10 counties each
* I created a file called prism\_runall.do and appended the resulting datasets together after running the analysis 10 counties at a time and outputting 6 intermediate files

Outputs from STATA files:

* From the first cleaning effort (inelegantly named prism\_cleaning\_0\_10v3.do), I saved a long data format, which had the raw indicators and all months over the 37 years called appended\_prism.dta
* Then after getting the indicator change measure and appending the files (in prism\_runall.do), I created an output called prism\_all.dta, which has one line per county with its change for each indicator (this change is a measure of yearly change so I multiplied by 37 in many cases to show total change)
* I made a post-analysis file to load data in ready-to-visualize formats in R in another STATA script file called prism\_viz.do
  + This created csvs with annual averages (across the months) for the indicators,

Visualizing data:

* I conducted the visualization in Jupyter Notebooks (with Python), Rstudio, and Datawrapper