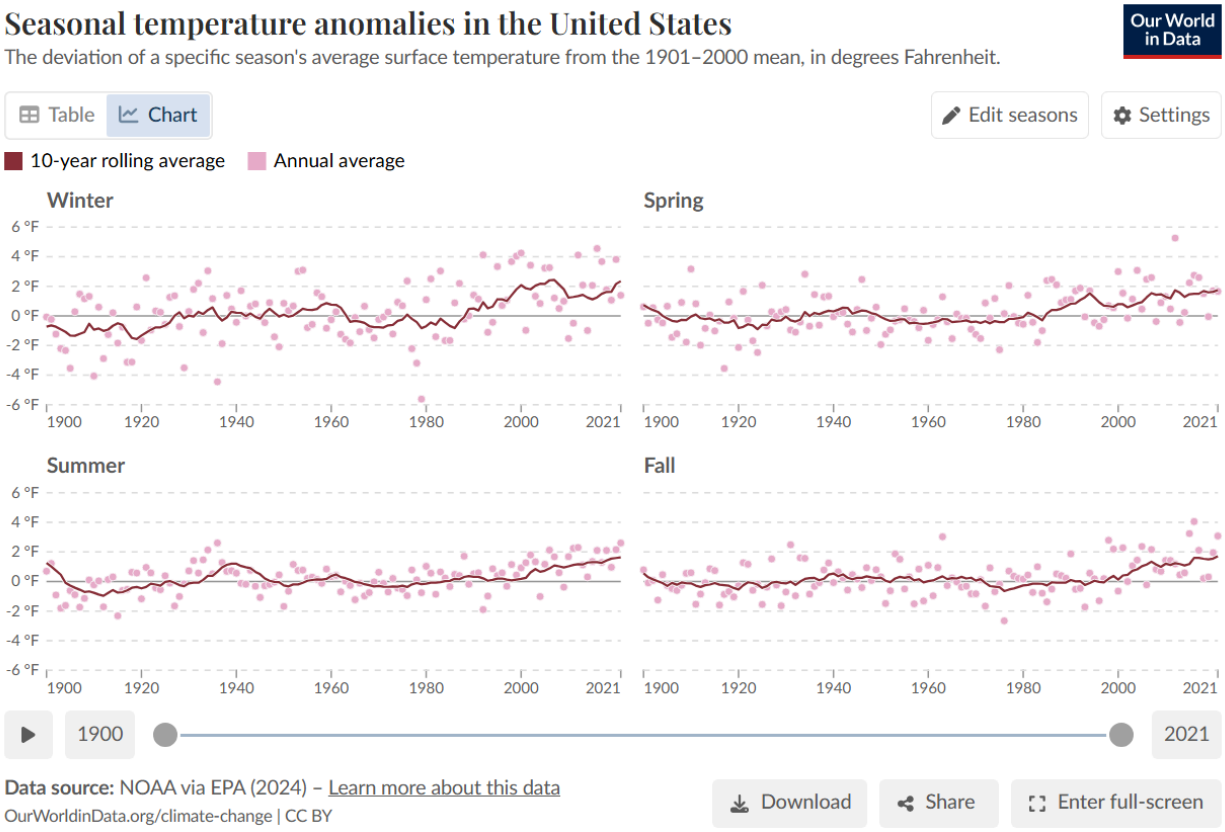
Kyle Gustafson

Overview

Github Username: gusky14

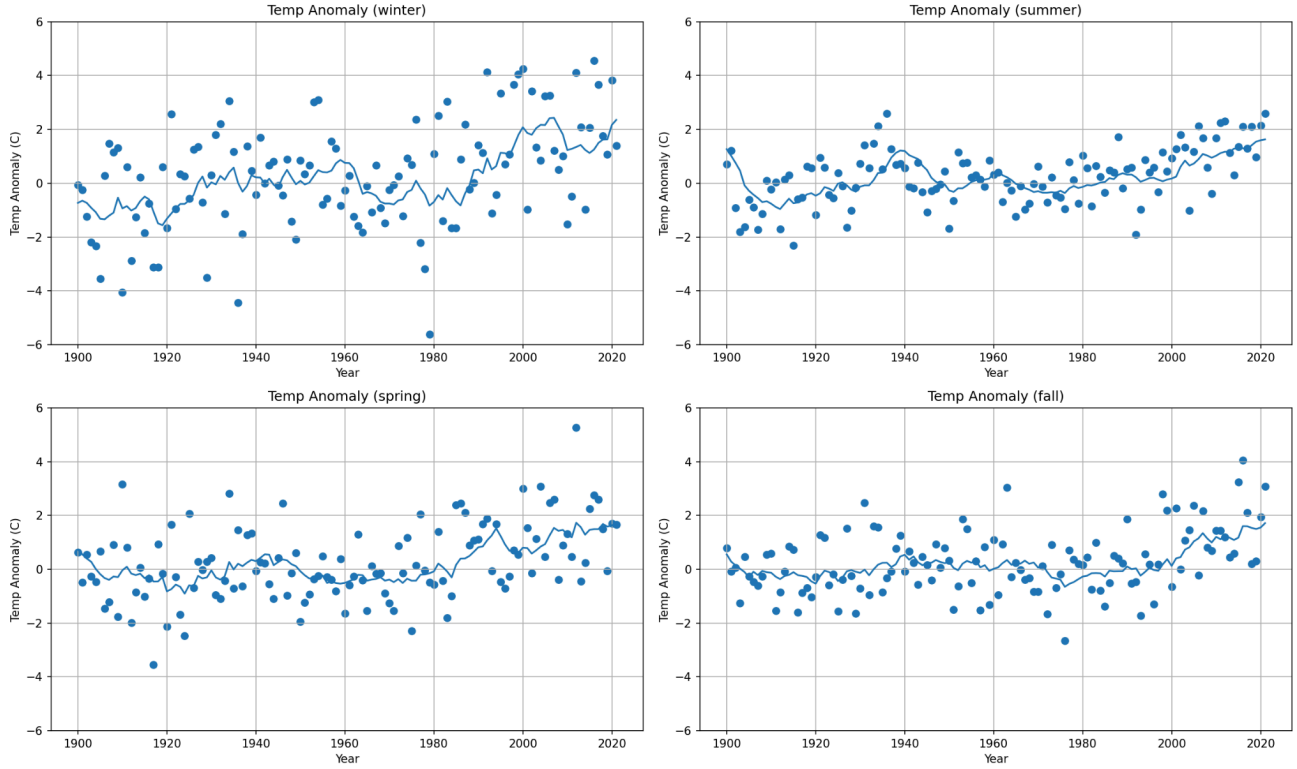
Visualization

When I went to ourworldindata.org, I immediately looked for weather data, as this is often the type of data I am most interested in. I chose this one because it expands upon a fairly common analyzed data topic (temperature rise) by showing per season.



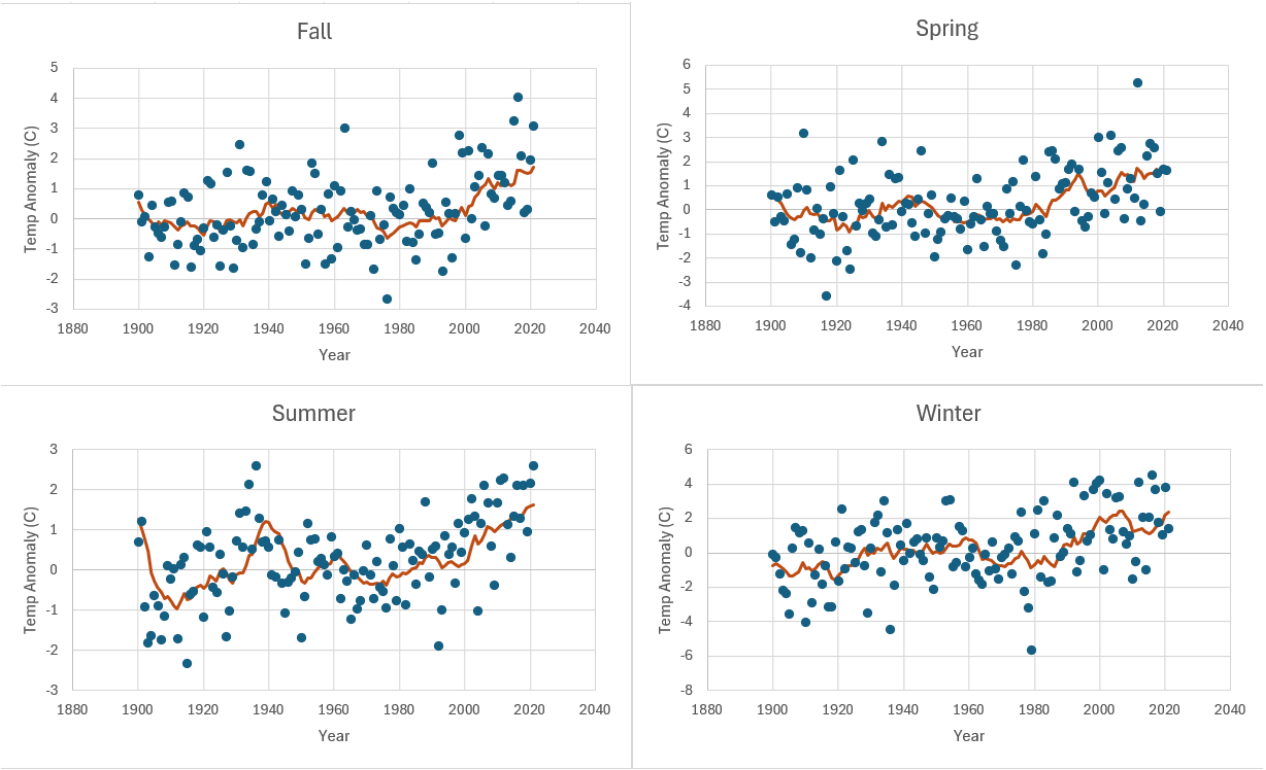
Recreations

#1: Python



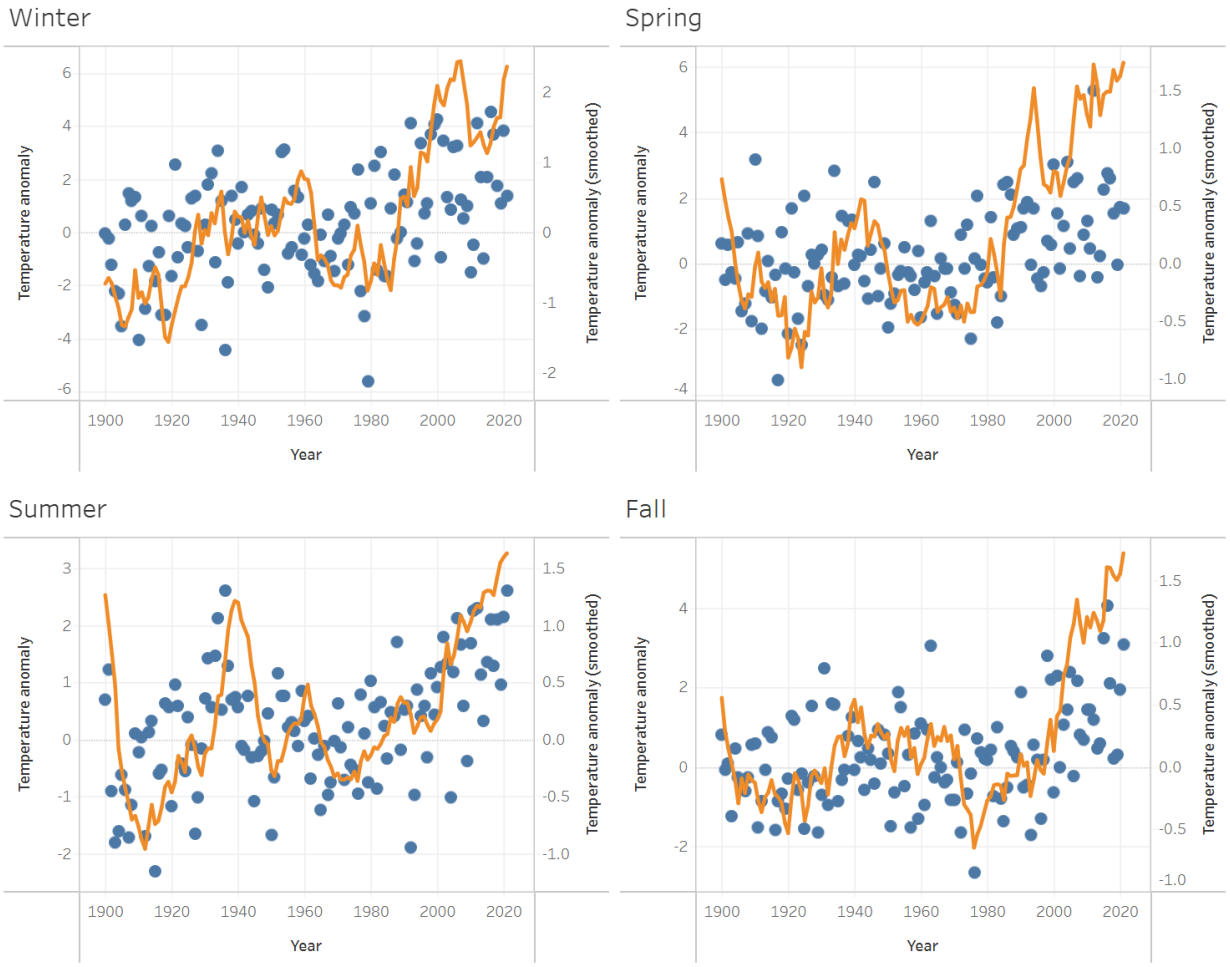
Using python ended up taking the longest, mostly because I kept deciding to go down unnecessary rabbit holes, including trying to create a data visualization package that I could install with pip and optimizing my code with libraries and functions. Aside from this, I created sub data frames for each season and used subplots to put all of the charts in the same figure. I spend the majority of my time figuring out subplots and figuring out my functions and library; plotting the data did not take long.

#2: Excel



Using excel also ended up taking me a long time, as I had to figure out how to use PivotTables to separate out the seasons in the data. I also went through a lengthy process of learning how to correctly format data series. This whole process probably took me a little less time than using python, but still took way longer than I had intended.

#3: Tableau (web)



This method was probably the quickest of the three. That said, I found I had to tackle a steep learning curve, so this method still took me a very long time. I used the seasons column as a filter, creating separate charts for each season and plotting them in a dashboard. I used a dual axis to plot the line and scatter plots in the same figure. Most of my time was spent learning how to use Tableau rather than actually analyzing the data.

Reflection

In reality, all three methods ended up taking me a similar amount of time. However, most of my time working in python was spent trying to figure out how to make an importable package out of my very small visualization helper library, which I ended up completely abandoning around two hours in after getting more confused than it was worth. Ignoring time spent learning how to properly use each method, I think that Tableau is by far the easiest for the scope of this analysis. It utilizes simple drag and drop features and has many simple methods to accomplish different goals. Python is by far the most customizable, with essentially infinite different libraries and methods to accomplish different goals, as long as you know how to use them for what you want to do. Tableau is probably the best for communication (at least within the scope that I used all the tools), with good visual formatting all around.