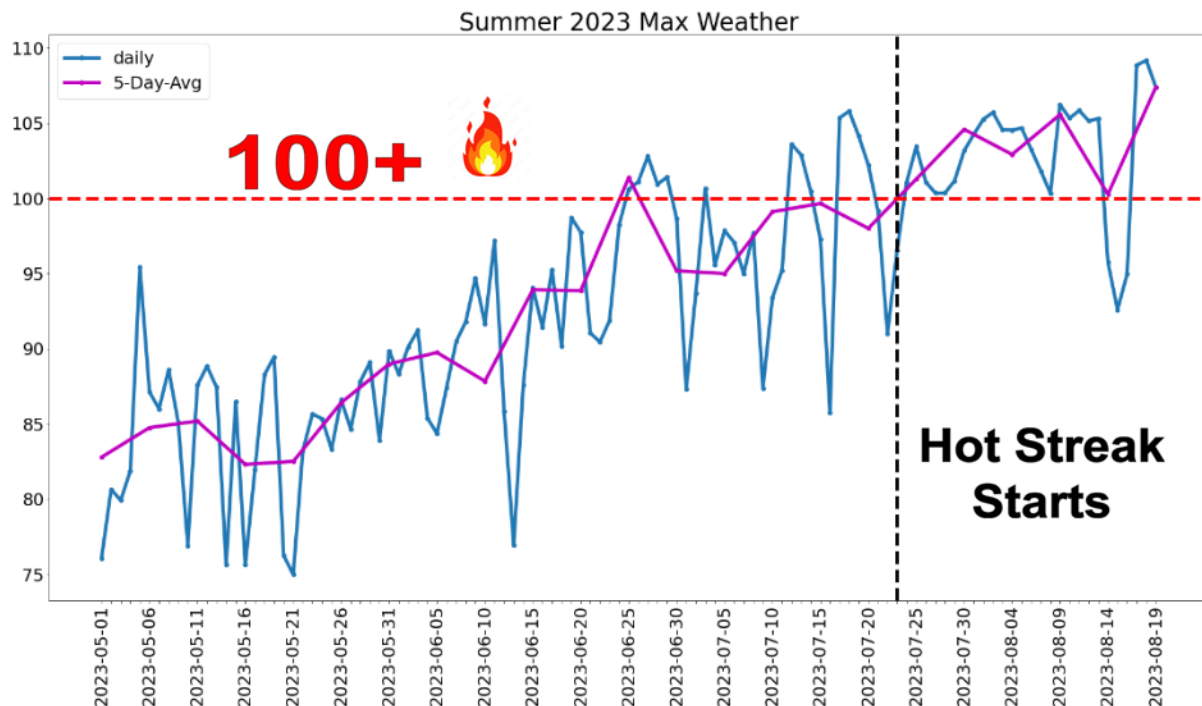


TEXAS HEAT EXPLAINED

My dad and I were having a silly debate on which Texas month was hotter, July or August. He believed August was the hottest month, and the weather got progressively hotter from June - August; I thought July was the hottest month. I determined this once and for all by doing some data analysis on Dallas, Texas weather. I used Python's matplotlib library for this endeavor specifically. In this post I will show the highest temperature and average temperature in each summer month (May - August) as well as the overall temperature distribution.



Narrowing data population scope to Summer 2023 (May - August) and calculating five day average weather

```
In [9]: 1 in_scope = weather[weather['date'] >= '2023-05-01']
        2 max_weather = in_scope.groupby('date')['temp'].max()
```

Local N day average calculation

```
In [10]: 1 three_day_average = max_weather.groupby(np.arange(len(max_weather))//3).mean()
        2 three_day_average.index = [max_weather.index[i] for i in range(0, max_weather.shape[0], 3)]
        3
        4 five_day_average = max_weather.groupby(np.arange(len(max_weather))//5).mean()
        5 five_day_average.index = [max_weather.index[i] for i in range(0, max_weather.shape[0], 5)]
```

The blue line is the original weather signal, the purple line is the consecutive 5 day average

Plotting max weather data and 5 day average with matplotlib

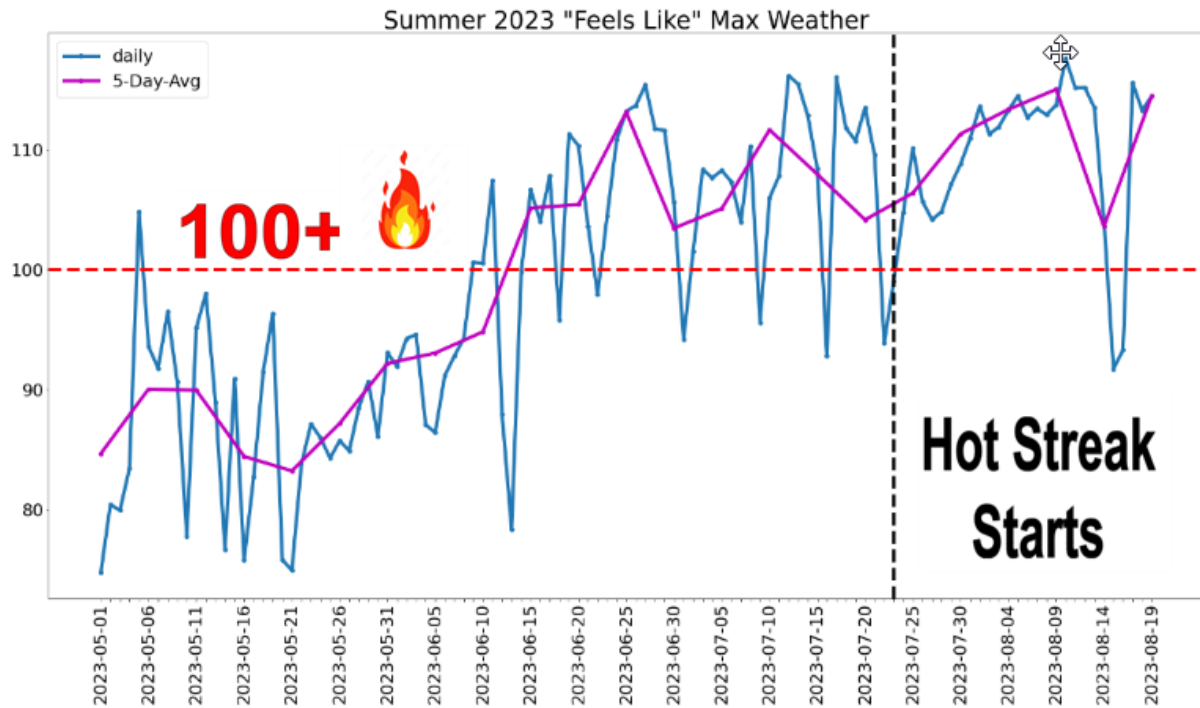
```
In [42]: 1 plt.rcParams["figure.figsize"] = (30, 15)
2 fig, ax = plt.subplots()
3
4 ax.plot_date(max_weather.index, max_weather, linestyle='solid', linewidth=5, label='daily')
5 ax.plot_date(five_day_average.index, five_day_average, linestyle='solid', linewidth=5, color='m', label='5-Day-Avg')
6 original_labels = [str(label) for label in ax.get_xticks()]
7 labels_of_interest = [str(i) for i in np.arange(0, len(original_labels), 5)]
8 new_labels = [str(max_weather.index[int(label)]) if label in labels_of_interest else "" for label in original_labels]
9 ax.set_xticklabels(new_labels)
10 plt.xticks(rotation=90)
11 plt.yticks(fontsize=25)
12 plt.xticks(fontsize=25)
13 plt.title('Summer 2023 Max Weather', fontsize=35)
14 plt.axhline(y = 100, color = 'r', linestyle = 'dashed', linewidth=5)
15 plt.axvline(x = 83, color = 'k', linestyle = 'dashed', linewidth=5)
16 plt.legend(fontsize=25)
17 print()
18 print()
19 plt.show()
20 print()
21 print()
22 fig.savefig('Summer_2023_Max_Weather.png')
```

The hot streak really starts in that last week of July, I know there have been a few dips, but that has been the largest hot streak this summer season.

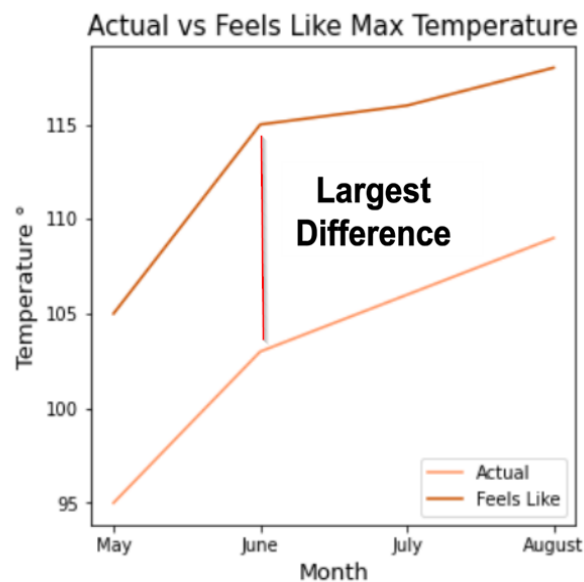
I wanted to clarify the 5 day average is not a moving average, I did that because there are 111 days that are in this dataset, so I wanted to condense the data down.

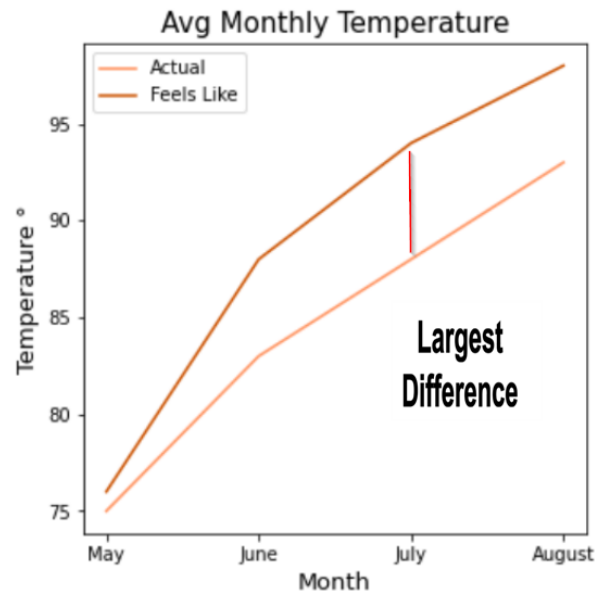
For example

Day	Weather	5 Day Avg
1	0	10
2	5	35
3	10	
4	15	
5	20	
6	25	
7	30	
8	35	
9	40	
10	45	



I will show it more clearly below, but the “feels like” temperature, definitely exceeds the actual temperature, the difference between the two temperatures is only apparent in extremely cold or hot weather, as shown in the graph below.





Conclusion

From this analysis we can see that August is the hottest month, since August has the highest average and max temperature. The analysis I showed above only scratches the surface of what is needed to determine the hottest month. In my later posts I will also show the number of consecutive hot days in a month, as well as the heat duration in a particular month, i.e. how late into the day the weather stays above 90 degrees. I believe this year is a weather anomaly for hot weather as well, so I will also be looking at historical summer weather for the past 10 years to decide once and for all which summer month is the hottest.