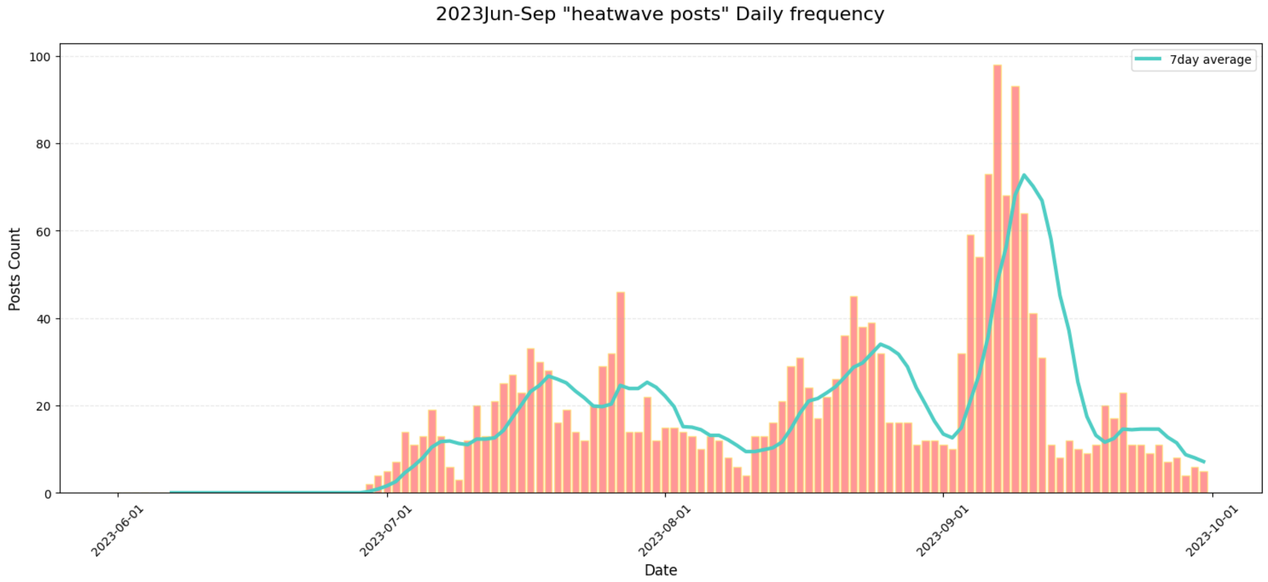
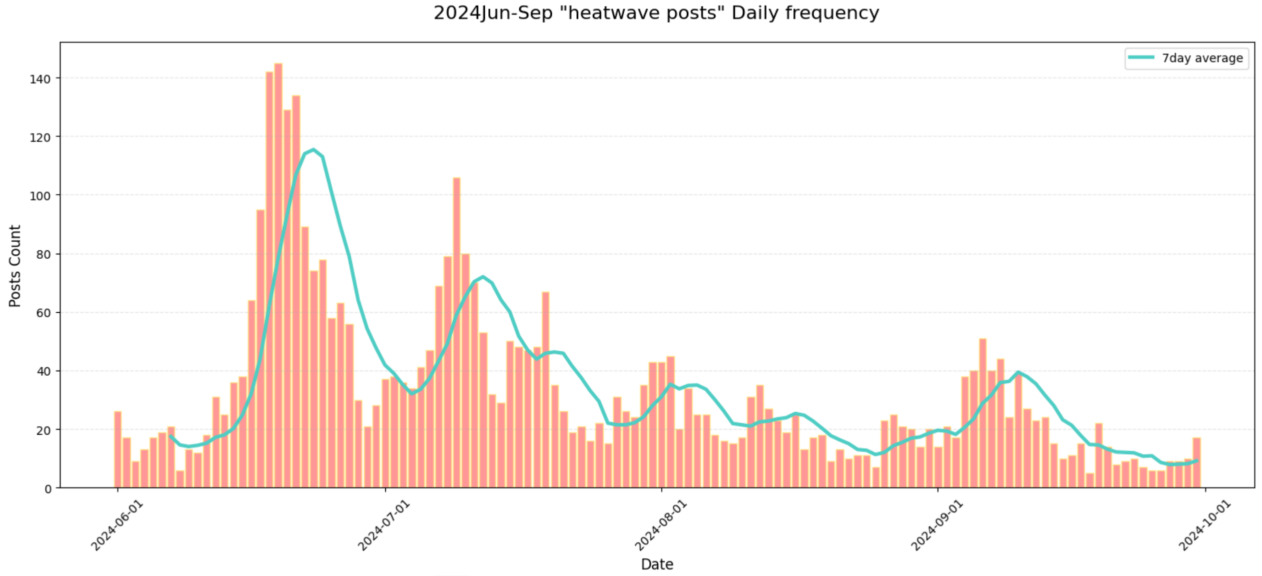
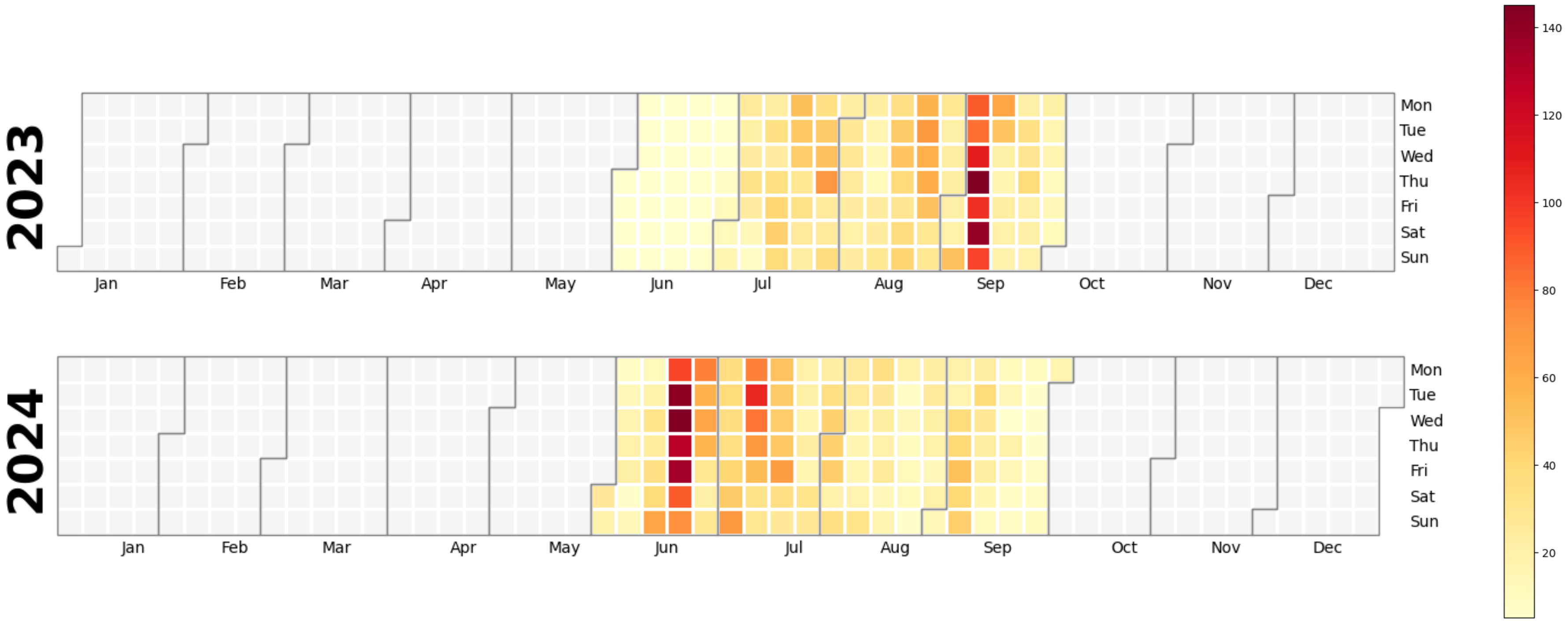
File 01 Bluesky API heatwave post frequency

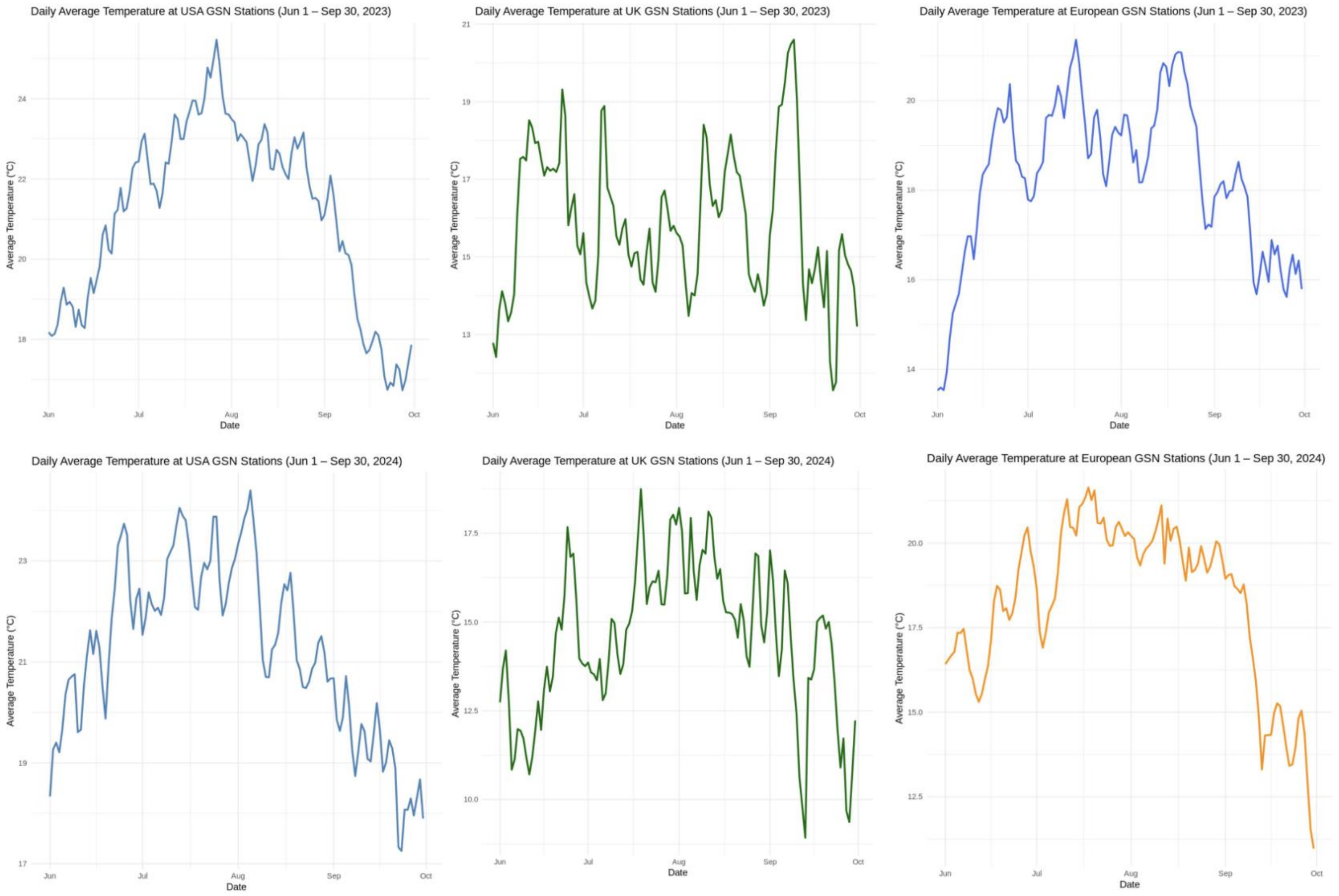






The data files related to post responses are saved in the 'heat post data' folder as 'heat_posts2023.jsonl' and 'heat_posts2024_standardized .jsonl'.

File 02 UK, USA and European temperature data

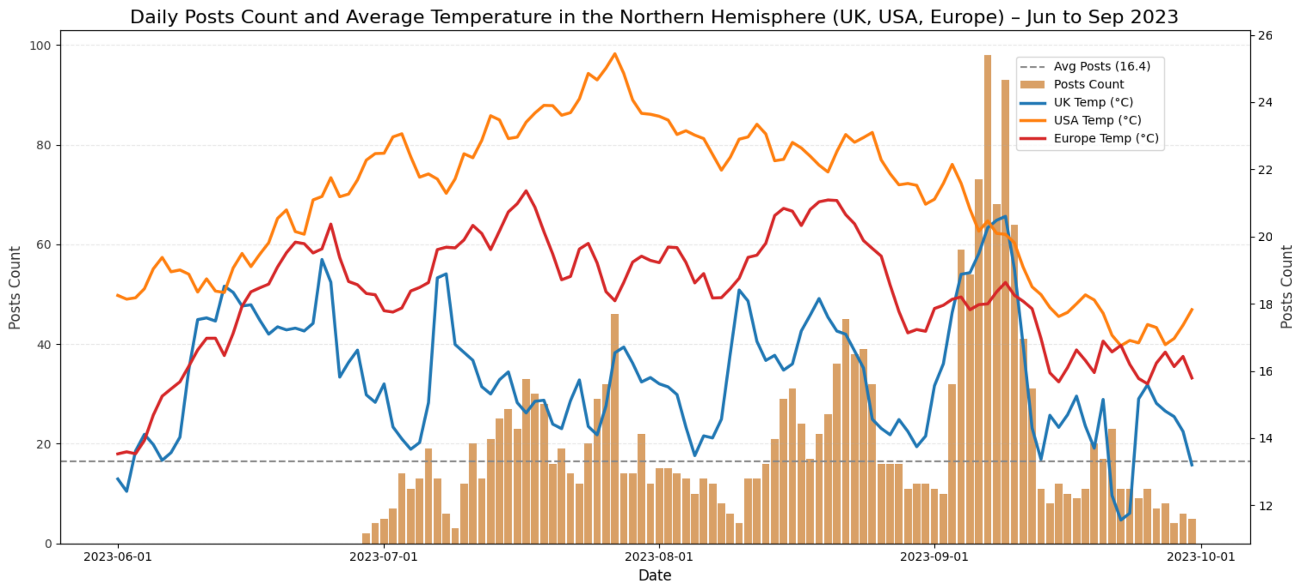


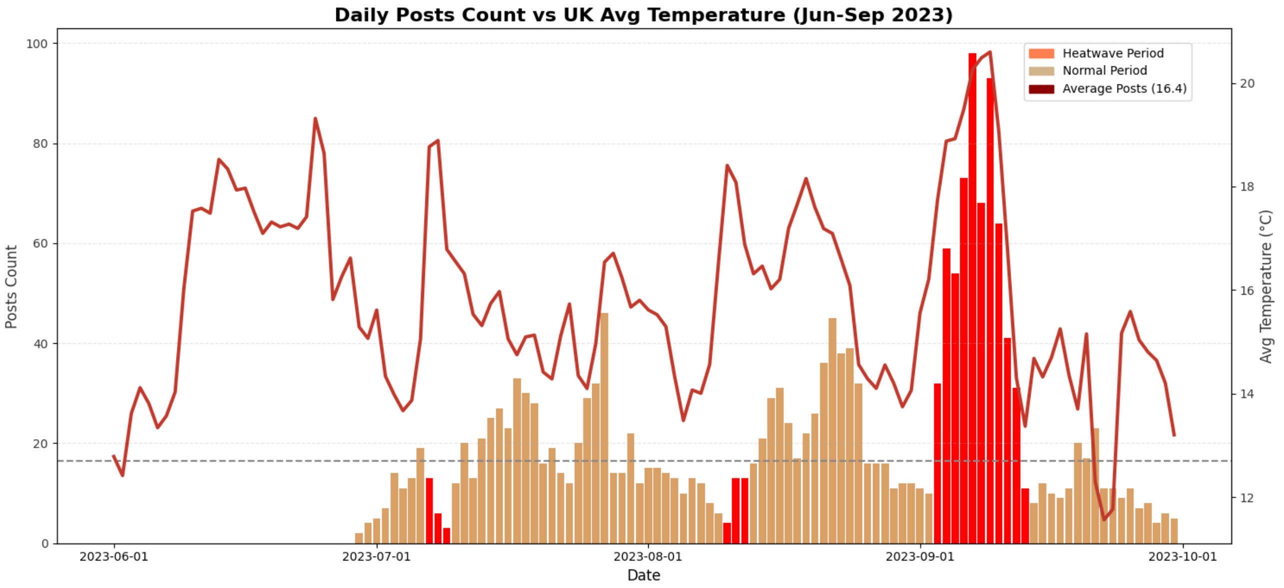
Save under the "UK, European and USA temp data" folder with the following files:

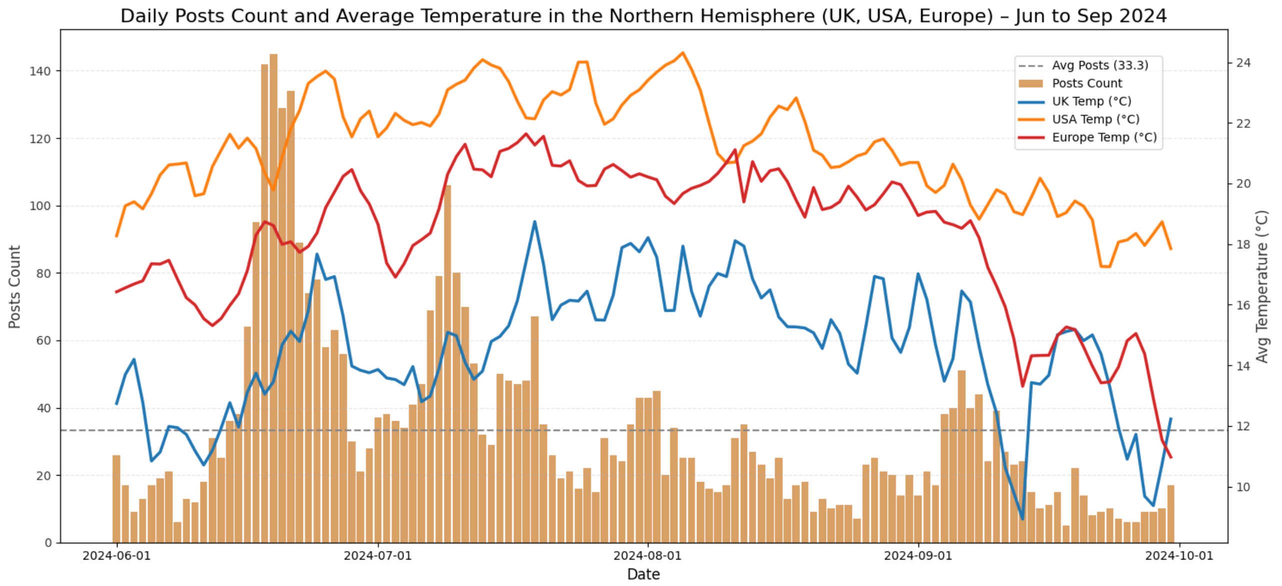
Files with "gsn" in the name contain data downloaded from all stations;
Files with "avg" in the name contain the daily average values across all stations.

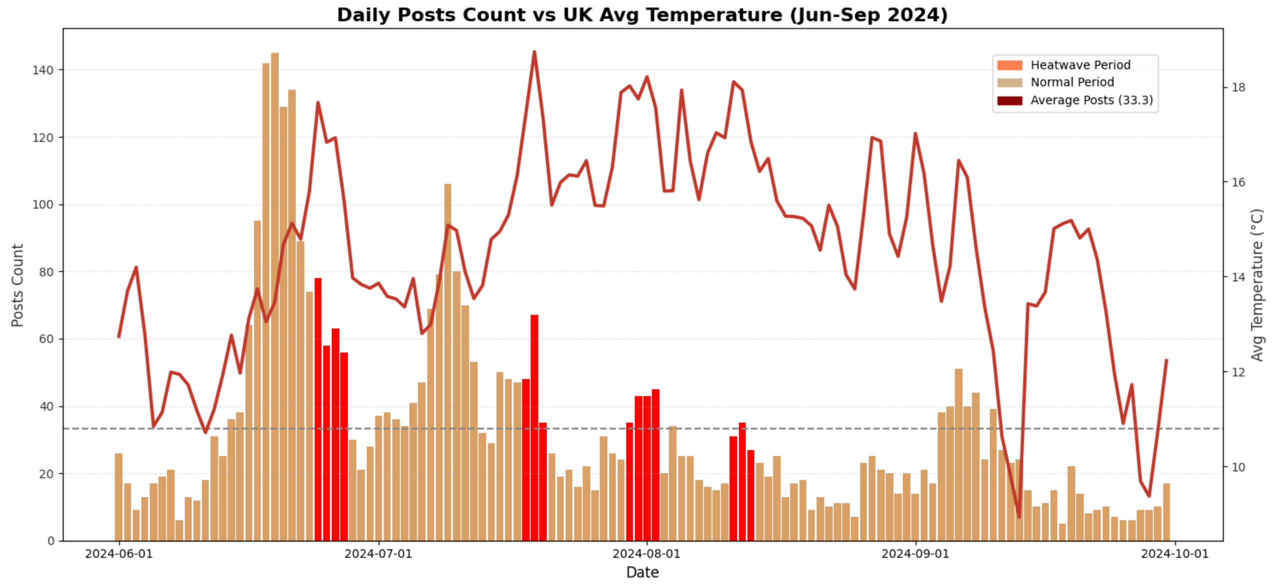
- europe_avg_temperature_2023_Jun-Oct.csv
- europe_avg_temperature_2024_Jun-Oct.csv
- europen_temperature_gsn_2023.csv
- europen_temperature_gsn_2024.csv
- uk_avg_temperature_2023_Jun-Oct.csv
- uk_avg_temperature_2024_Jun-Oct.csv
- uk_temperature_gsn_2023.csv
- uk_temperature_gsn_2024.csv
- usa_avg_temperature_2023_Jun-Oct.csv
- usa_avg_temperature_2024_Jun-Oct.csv
- usa_temperature_gsn_2023.csv
- usa_temperature_gsn_2024.csv

File 03A Temporal Trends of Post Frequency and Temperature

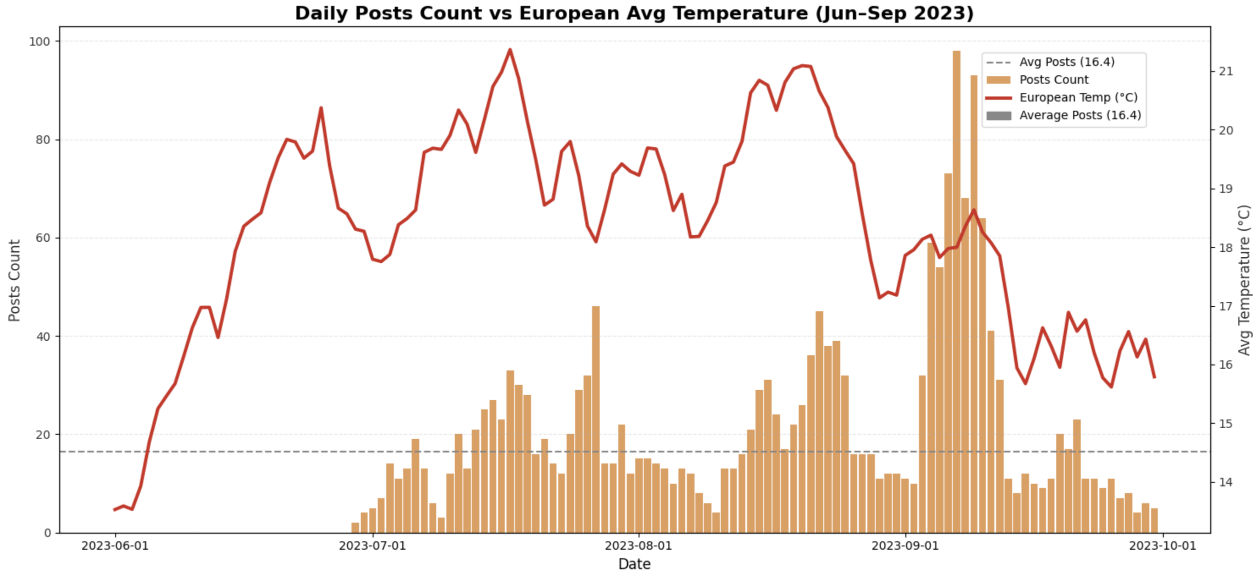


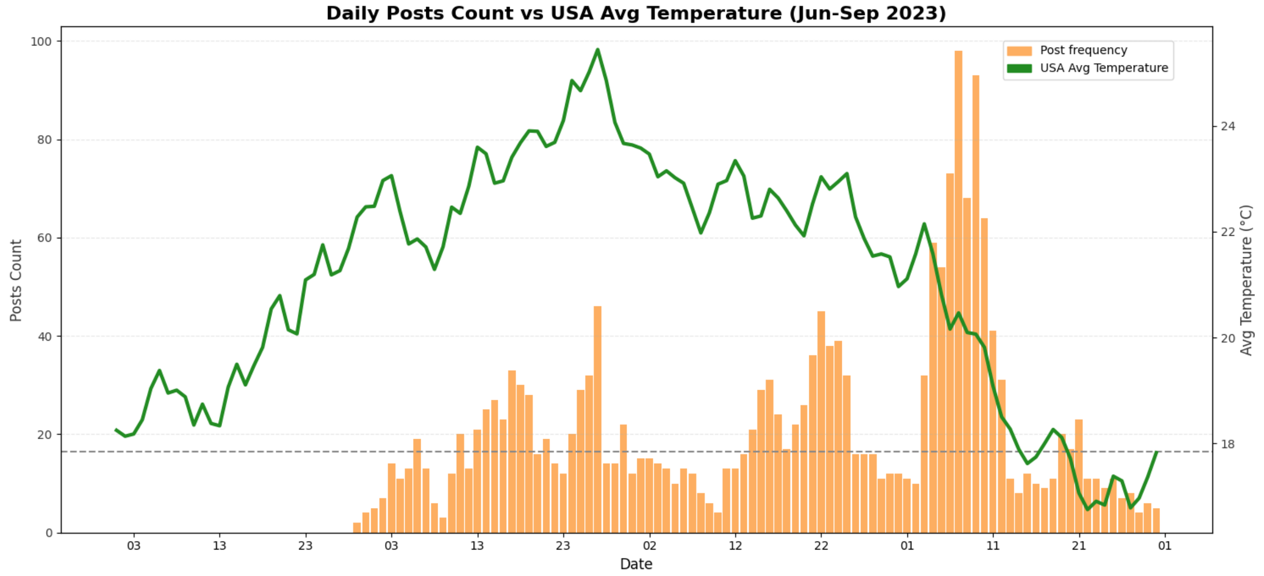


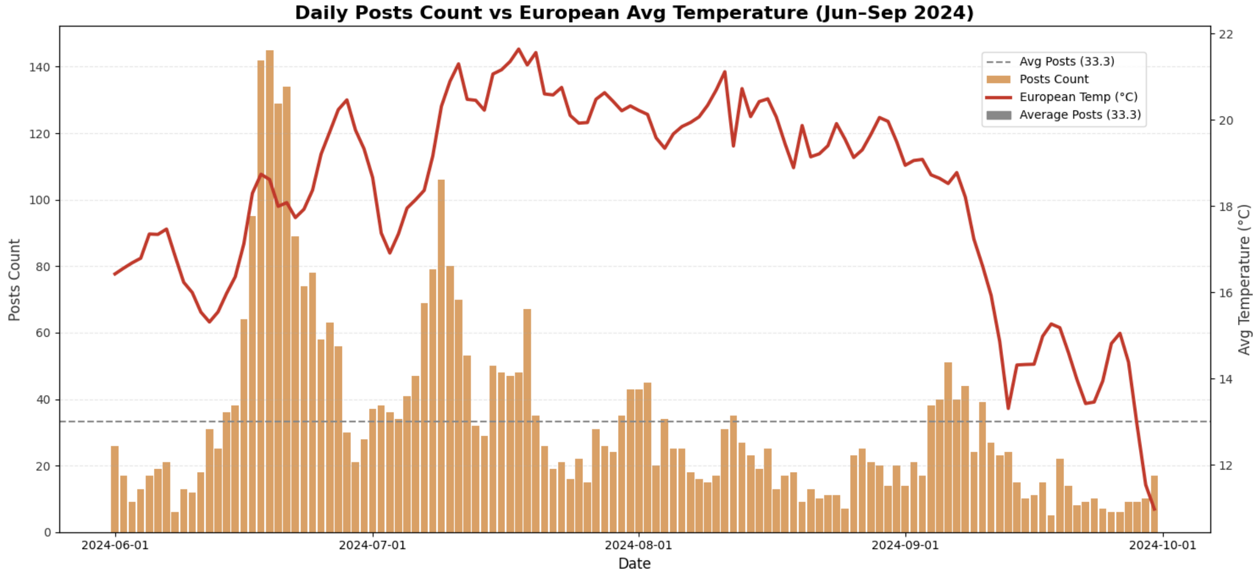


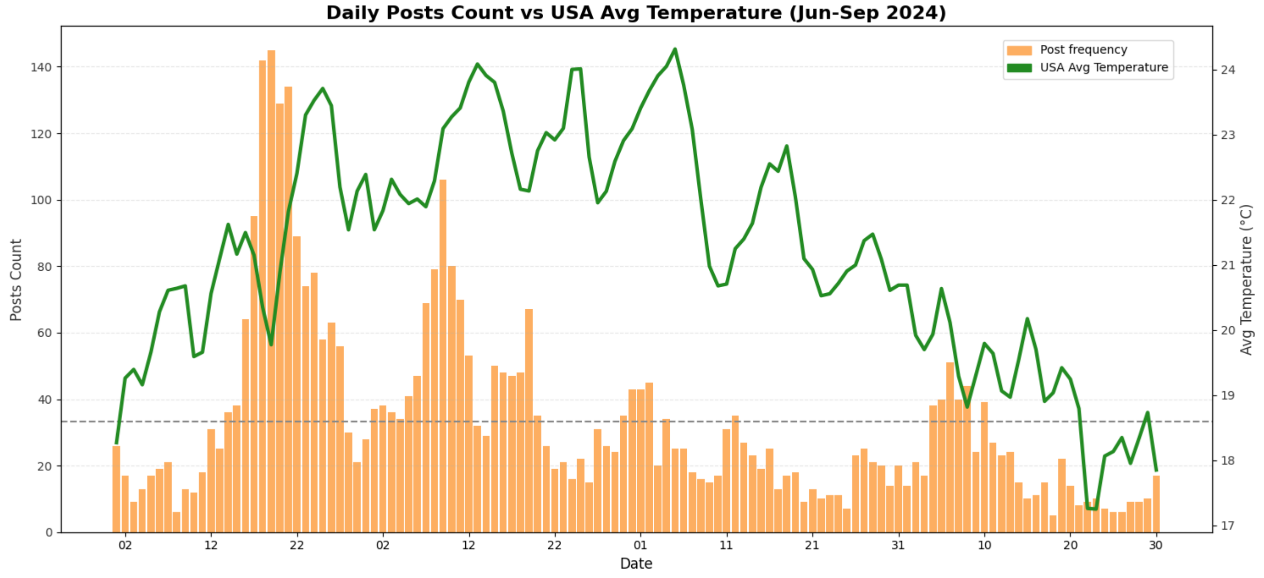


File 03A Temporal Trends of Post Frequency and Temperature

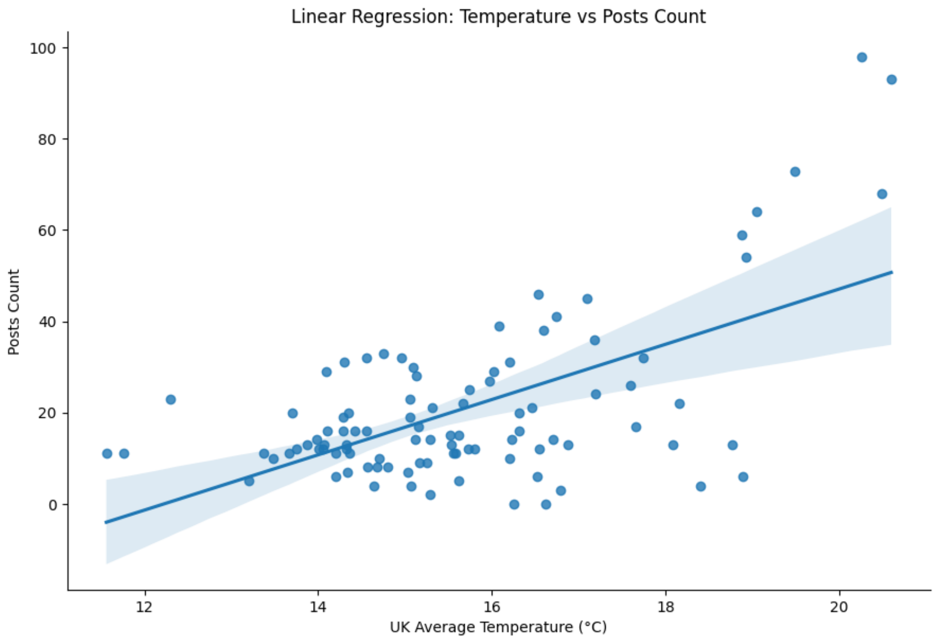








File 03B UK tem&post with lag and forecast effect 2023



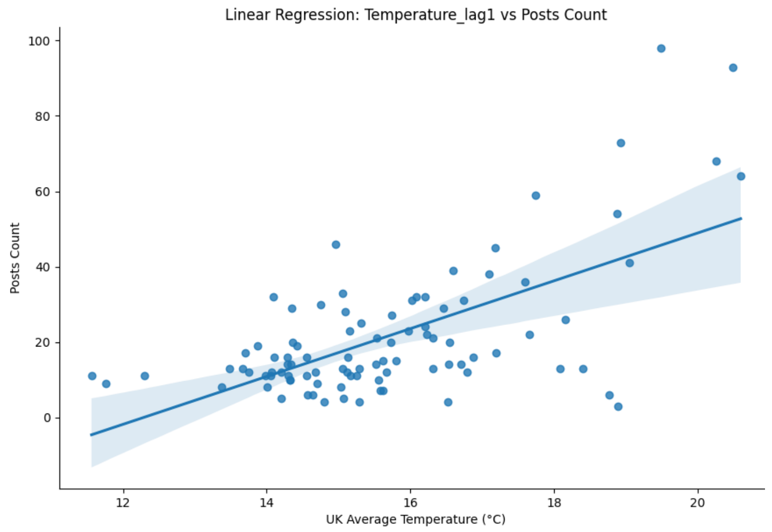
- Dependent Variable: post_count (daily post count)
- R-squared = 0.361: Model explains 36.1% of the variance in post counts.

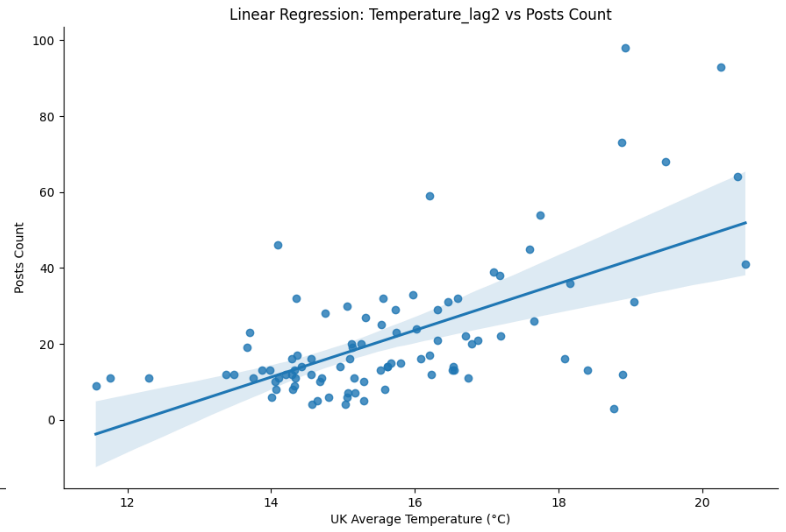
Significance (P>|t|):

avg_temp (p = 0.001) is significant — today's temperature affects post counts.

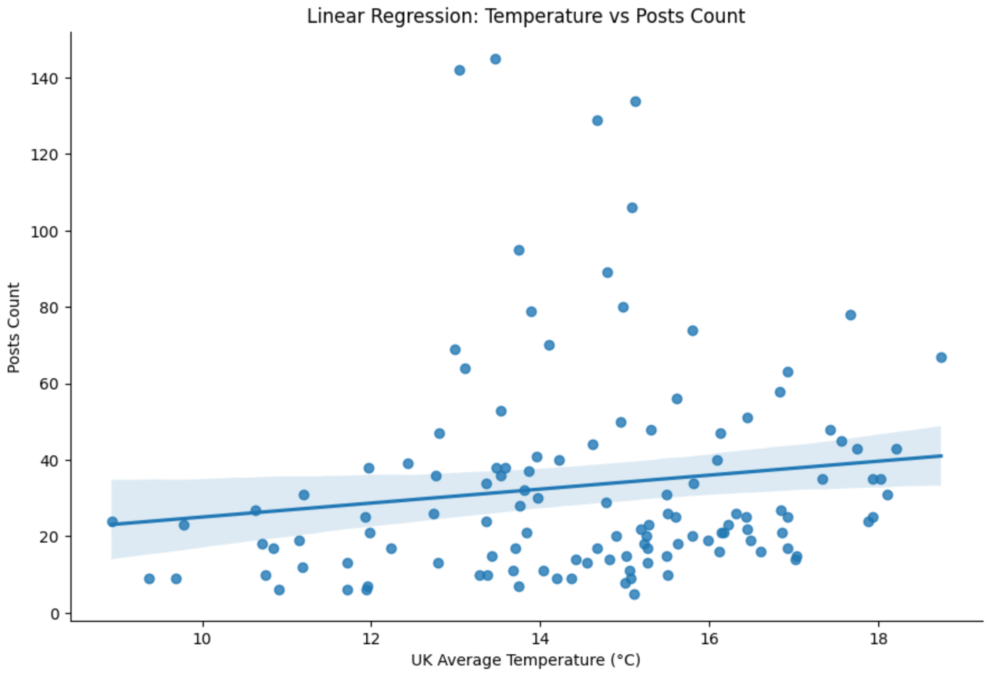
avg_temp_lag1 and avg_temp_lag2 (p = 0.888 and 0.526) are not significant — future temperatures have no effect.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:            posts_count   R-squared:                       0.361
Model:                            OLS   Adj. R-squared:                  0.354
Method:                 Least Squares   F-statistic:                     53.15
Date:                Sat, 26 Apr 2025   Prob (F-statistic):           9.55e-11
Time:                        19:44:15   Log-Likelihood:                -392.60
No. Observations:                  96   AIC:                             789.2
Df Residuals:                      94   BIC:                             794.3
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const        -74.0988     13.108     -5.653      0.000    -100.125     -48.073
UK             6.0585      0.831      7.290      0.000       4.408       7.709
==============================================================================
Omnibus:                        6.314   Durbin-Watson:                   0.536
Prob(Omnibus):                  0.043   Jarque-Bera (JB):                6.687
Skew:                           0.390   Prob(JB):                       0.0353
Kurtosis:                       4.032   Cond. No.                         139.
==============================================================================
```





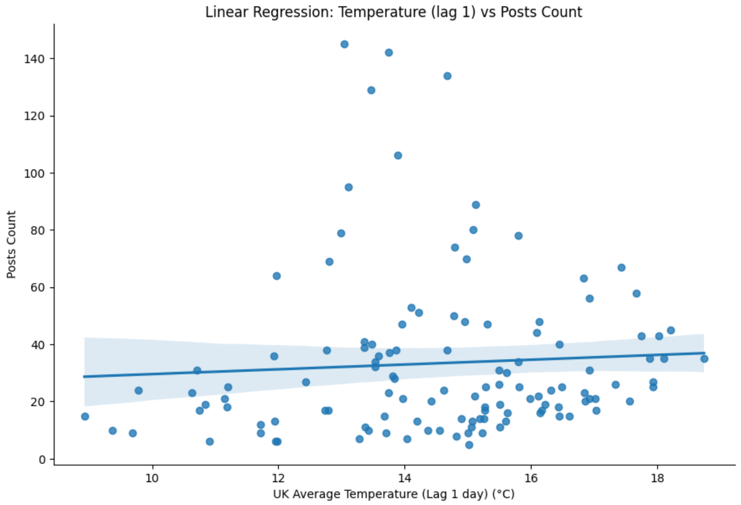
File 03C UK tem&post with lag and forecast effect 2024

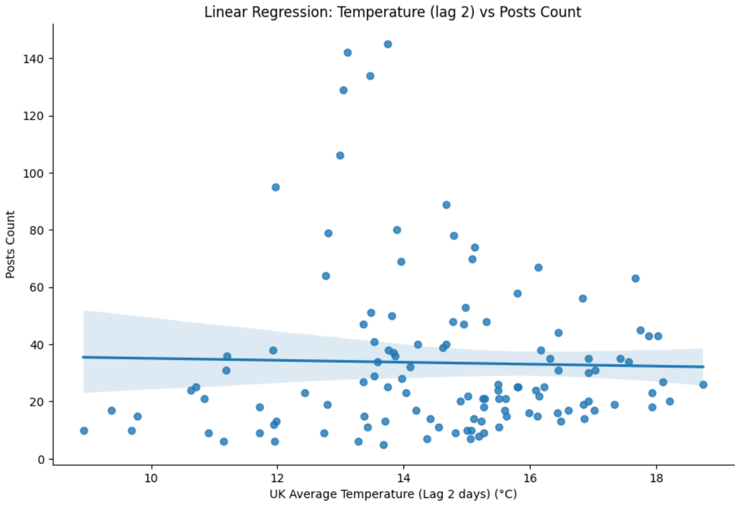


Dependent Variable: posts_count (daily post count)
R-squared = 0.019: Model explains 1.9% of the variance in post counts.

Significance (P>|t|):
UK average temperature (p = 0.129) is **not significant** — temperature has no significant effect on post counts in 2024.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:            posts_count   R-squared:                       0.019
Model:                            OLS   Adj. R-squared:                  0.011
Method:                 Least Squares   F-statistic:                     2.331
Date:                Sat, 26 Apr 2025   Prob (F-statistic):              0.129
Time:                        20:05:32   Log-Likelihood:                -578.34
No. Observations:                 122   AIC:                             1161.
Df Residuals:                     120   BIC:                             1166.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          6.8018     17.557      0.387      0.699     -27.960      41.564
UK             1.8256      1.196      1.527      0.129      -0.542       4.193
==============================================================================
Omnibus:                       68.626   Durbin-Watson:                   0.188
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              228.008
Skew:                           2.163   Prob(JB):                     3.08e-50
Kurtosis:                       8.113   Cond. No.                         102.
==============================================================================
```





File 04 global temp and humidity & time series

In the GitHub "global temperature data" folder, due to the large number of files retrieved from the ERA5 API, a Google Colab link is provided, as shown in the image. It includes all maximum temperature and maximum dew point temperature data from 2018 to 2024.

My Drive › 01 ERA5 post-process... › 2018

Type | People | Modified | Source

Name ↑	Owner	Last mo...
.DS_Store	me	Apr 24, 2025
2m_dewpoint_temperature_stream-oper_daily-max.nc	me	Apr 8, 2025
2m_temperature_0_daily-max.nc	me	Apr 8, 2025
daily_mean_2m_temperature2018.csv	me	Apr 8, 2025

The source files after applying the land-sea mask can be found under the links starting with "02" and "03".

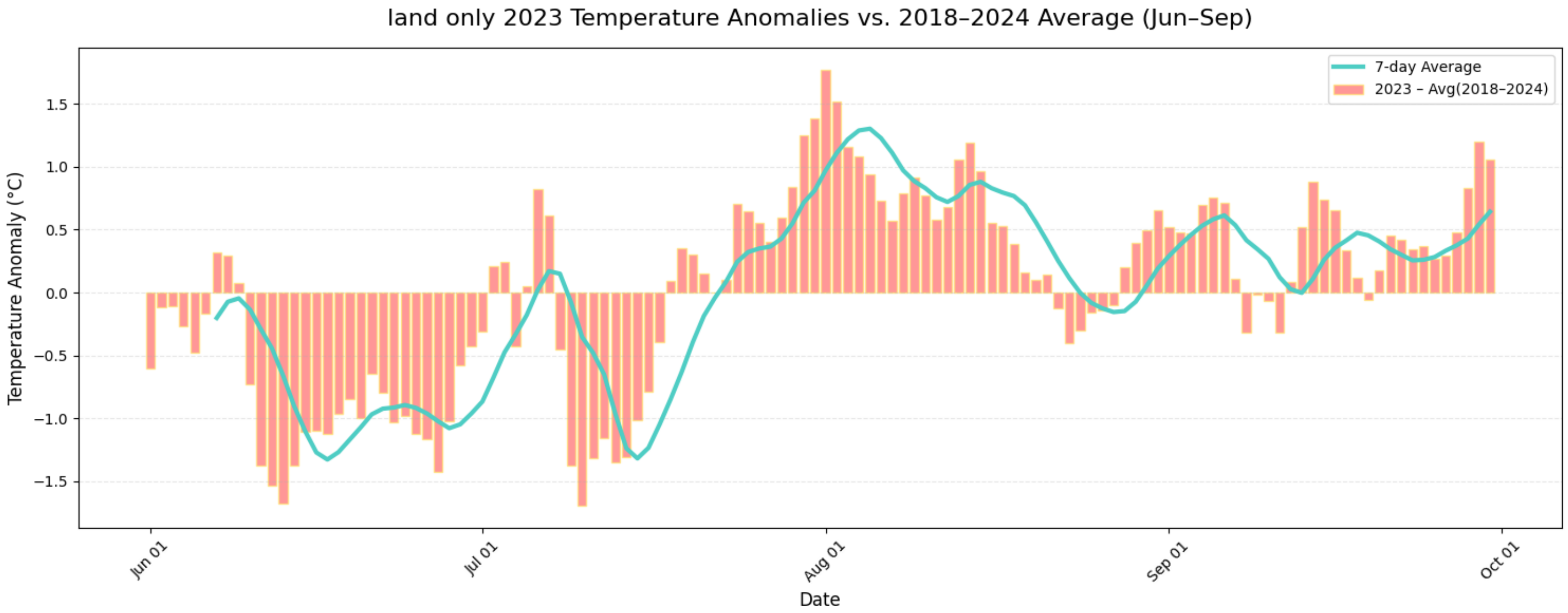
The "cdsapirc" file contains the official ERA5 land-sea mask.

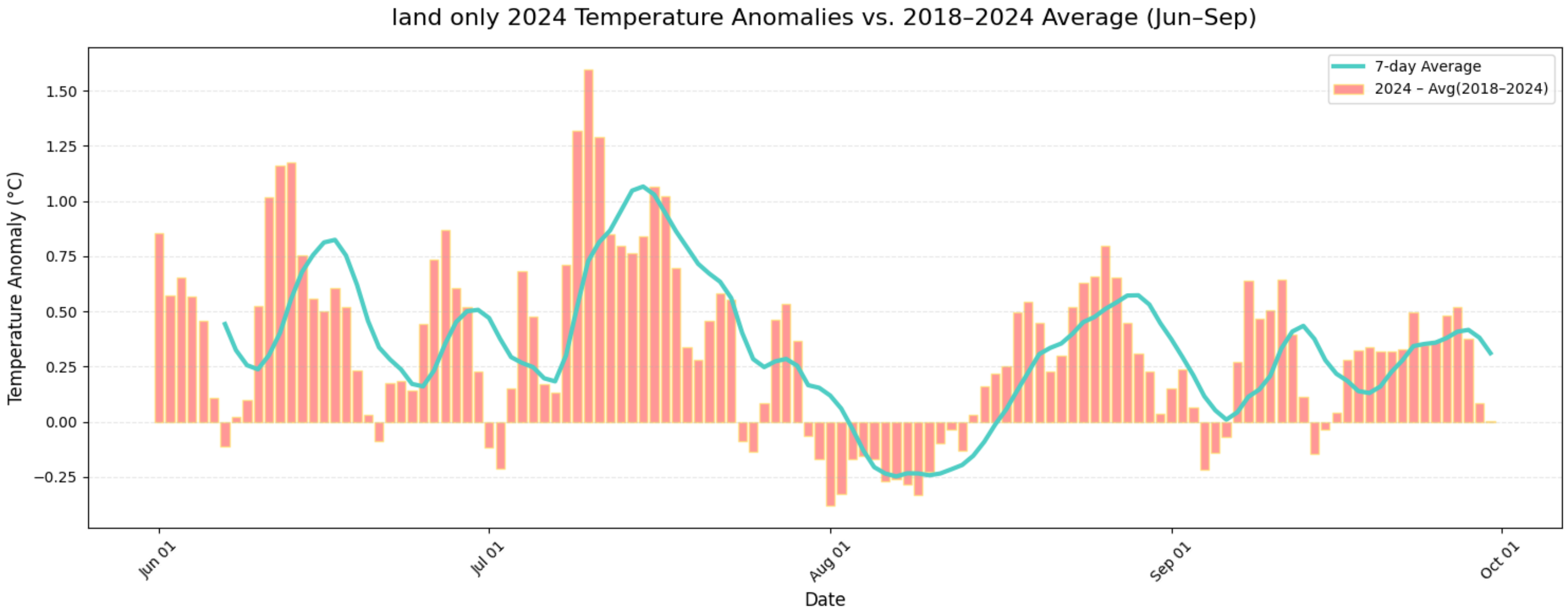
The two remaining Excel files document all calculation steps for identifying land surface anomalies in the Northern Hemisphere during this period.

- 01 ERA5 post-processed daily statistics raw data
- 02 After Land-sea mask tem 2m data
- 03 After Land-sea mask dewpoint tem data
- 2018-2024 land only 2m temperature daily average.xlsx
- 2018-2024 land only dewpoint temperature daily average dataset.xlsx
- cdsapirc

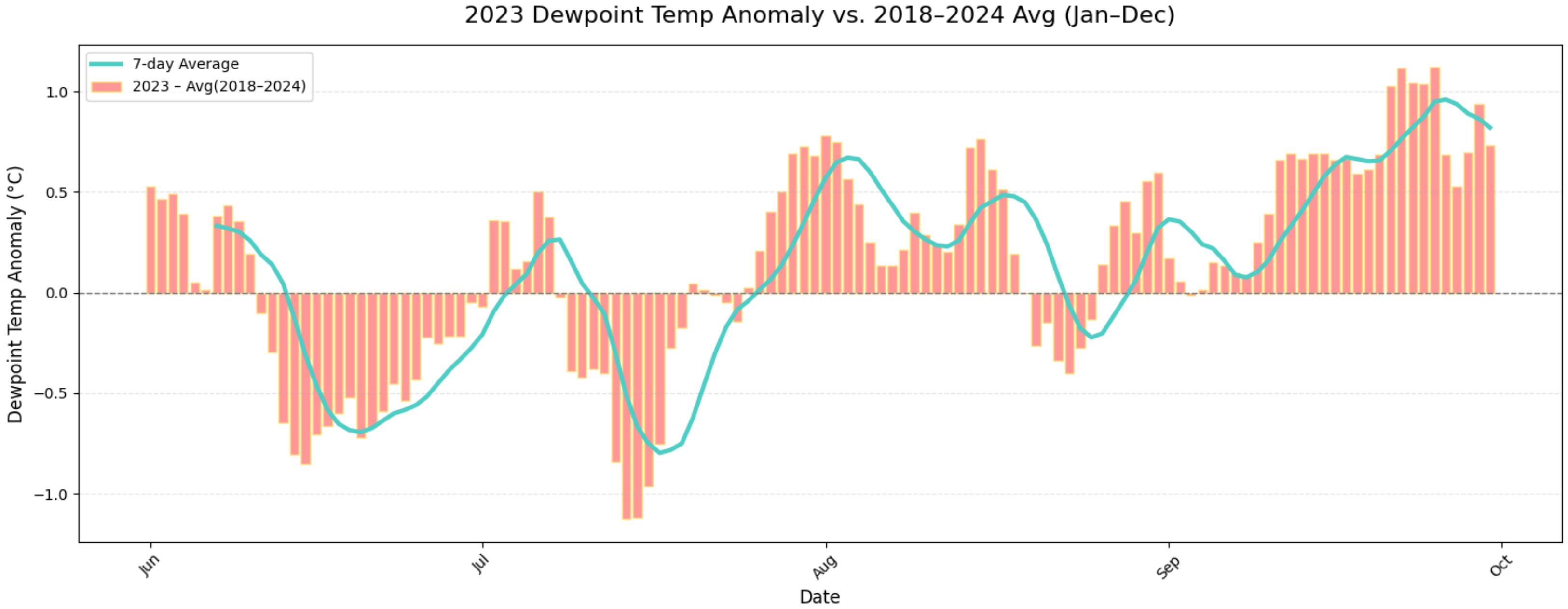
Month-Day	2018	2019	2020	2021	2022	2023	2024	2018-2024 average	2023 temperature minus average temperat	2024 temperature minus average temperature
06-01	21.666	22.324	22.2442	23.2504	23.1146	22.9866	23.59311	22.73984657	-0.606538	0.853263429
06-02	21.7471	22.7991	22.5906	23.6556	23.3698	23.4987	23.611969	23.03897086	-0.113312	0.572998143
06-03	21.8281	23.098	22.933	24.0361	23.0705	23.7749	23.882599	23.23187686	-0.107697	0.650722143
06-04	22.3637	23.2166	23.4249	24.0858	22.8625	23.6667	23.933746	23.364855	-0.266998	0.568891
06-05	23.2045	23.6539	23.7793	23.8343	23.0897	23.5818	24.056244	23.59996686	-0.474396	0.456277143
06-06	23.7364	24.019	23.9253	23.3421	23.5347	23.6611	23.829285	23.72113914	-0.168183	0.108145857
06-07	23.9483	24.3234	23.9671	23.2935	23.8039	24.095	23.772156	23.886187	0.322876	-0.114031
06-08	24.0041	24.4966	24.0165	23.4619	24.0344	24.3886	24.094818	24.070984	0.293732	0.023834
06-09	23.9438	24.5387	24.2203	23.5139	24.3241	24.3365	24.258057	24.16219671	0.07843	0.095860286
06-10	23.9328	24.2739	24.2811	23.5036	24.797	24.0152	24.744171	24.22109986	-0.728943	0.523071143
06-11	23.9845	24.0766	24.0244	23.9426	24.7368	23.9221	25.301117	24.284001	-1.379028	1.017116
06-12	24.3141	23.903	24.1046	24.1956	24.646	24.0151	25.549072	24.389666	-1.533935	1.159406
06-13	24.2299	23.8517	24.3352	24.1983	24.787	23.9099	25.590698	24.41465543	-1.680847	1.176042571
06-14	24.4227	23.9161	24.6488	24.5362	25.0967	23.9302	25.30188	24.55035843	-1.371674	0.751521571
06-15	24.6321	24.0627	24.9057	24.7556	25.2087	24.1626	25.273712	24.71443829	-1.111145	0.559273714
06-16	24.8766	24.6722	24.981	24.692	25.1067	24.248	25.343262	24.84569886	-1.095215	0.497563143
06-17	24.9573	25.0041	24.8898	24.9073	24.8841	24.4283	25.55075	24.94595343	-1.122466	0.604796571
06-18	24.8657	25.1213	24.9763	25.3795	24.6704	24.5709	25.536804	25.01726429	-0.965912	0.519539714
06-19	24.9683	25.4118	25.2009	25.336	24.8671	24.4601	25.30957	25.07909271	-0.849487	0.230477286
06-20	25.1187	25.6726	25.279	25.5812	24.8647	24.1522	25.147156	25.116516	-0.994965	0.03064

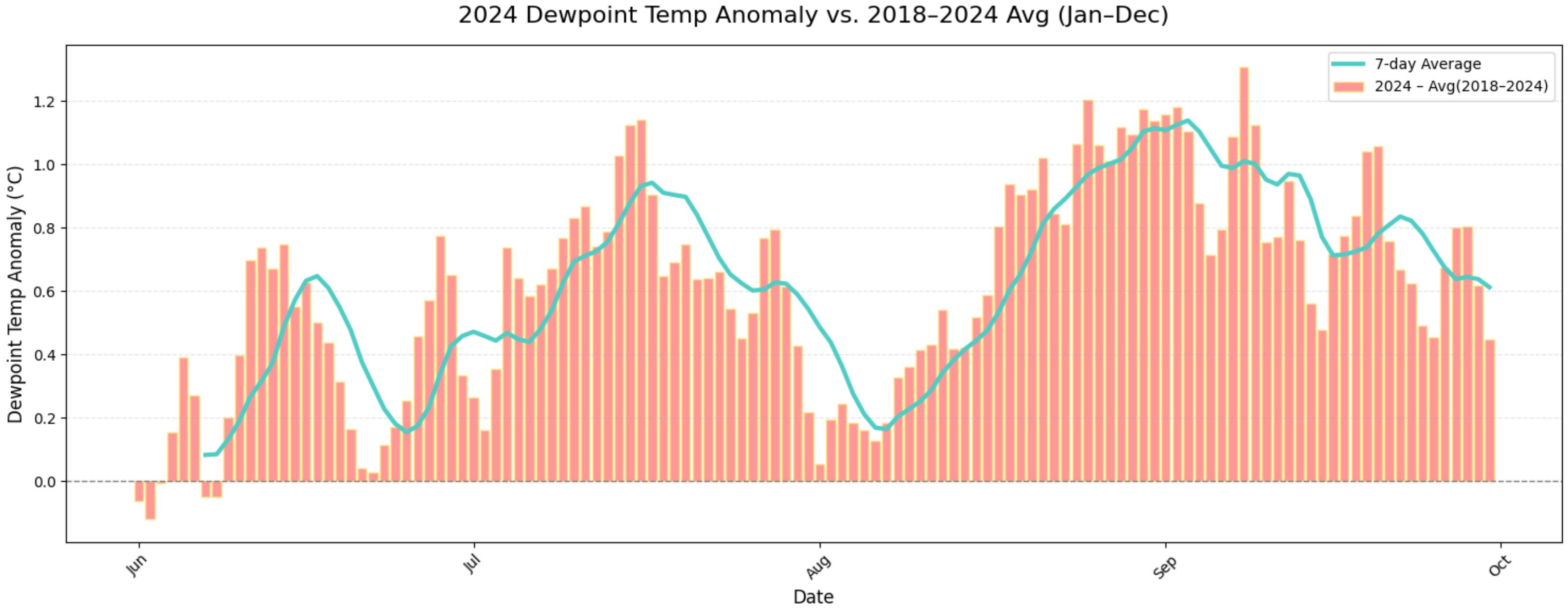
File 04 global temp and humidity & time series





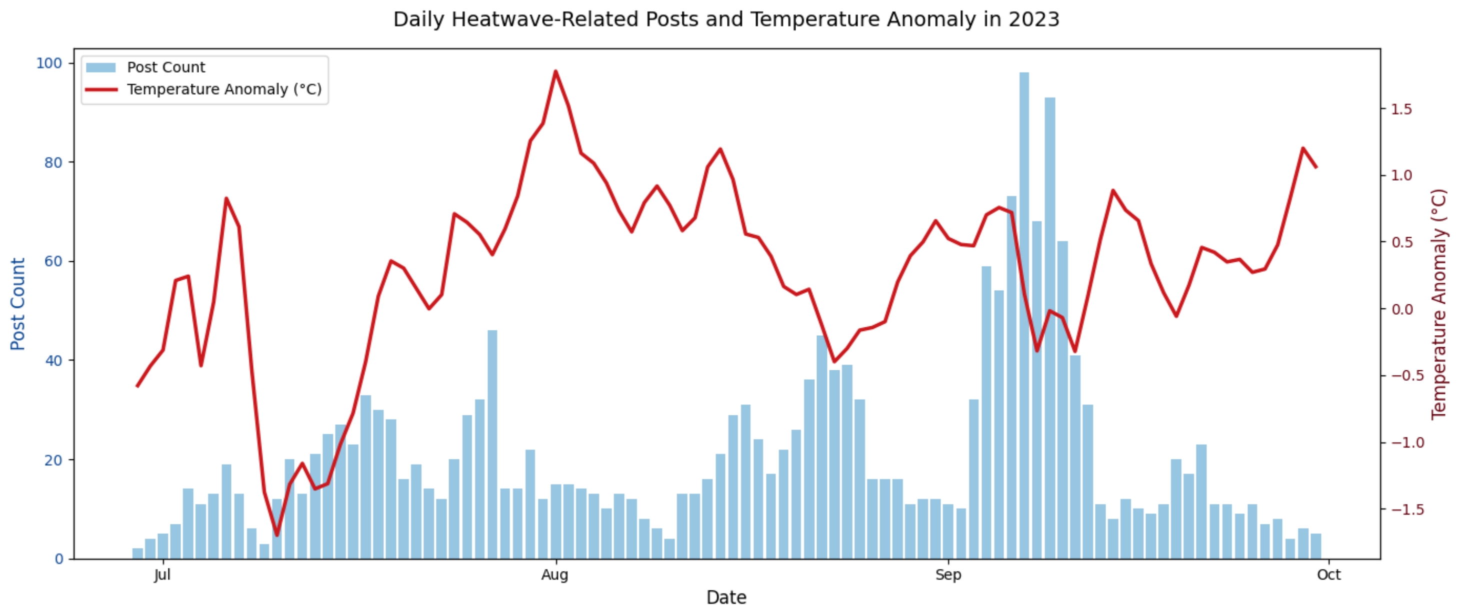
File 04 global temp and humidity & time series



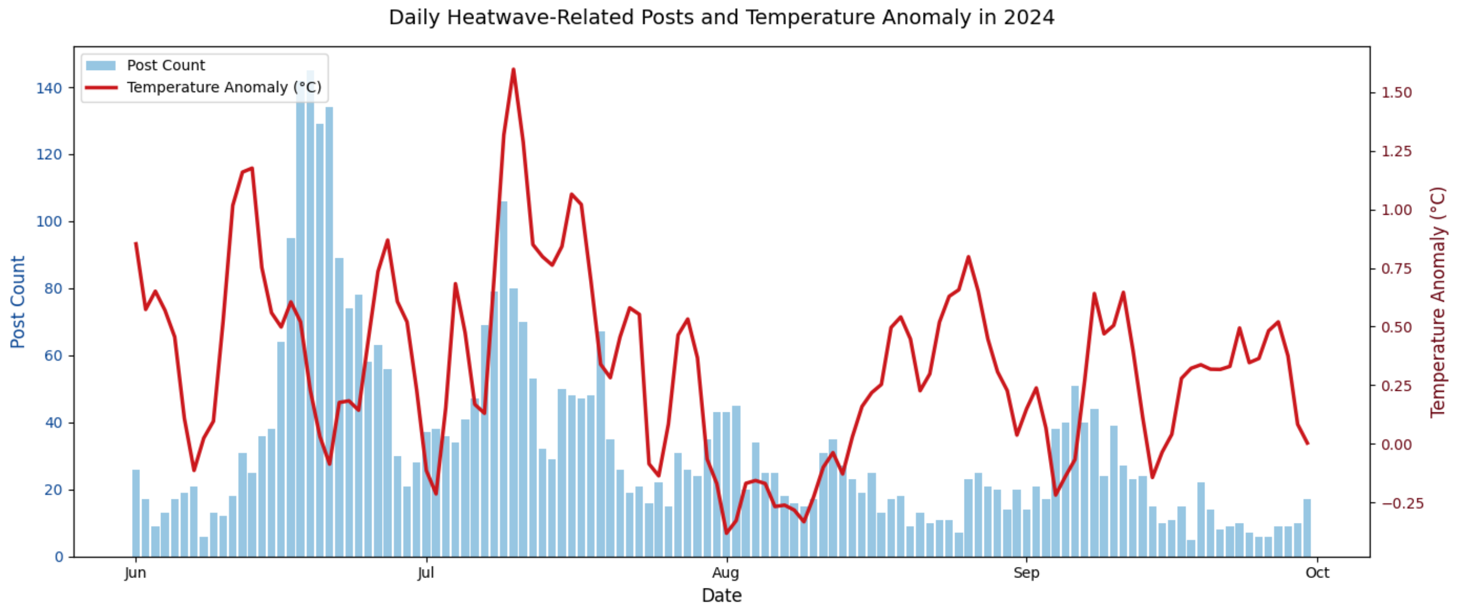


File 04 global temp and humidity & time series

For the time series plots, statistical analysis, and linear relationships, please refer to the document.



In 2023, the relationship between temperature anomalies and post count had an R^2 of 0.011 — temperature anomalies explained only 1.1% of the variation in post count. The temp_anomaly coefficient was –2.89, meaning each 1°C increase was associated with an average decrease of about 2.89 posts, but the result was not statistically significant ($p = 0.307$).

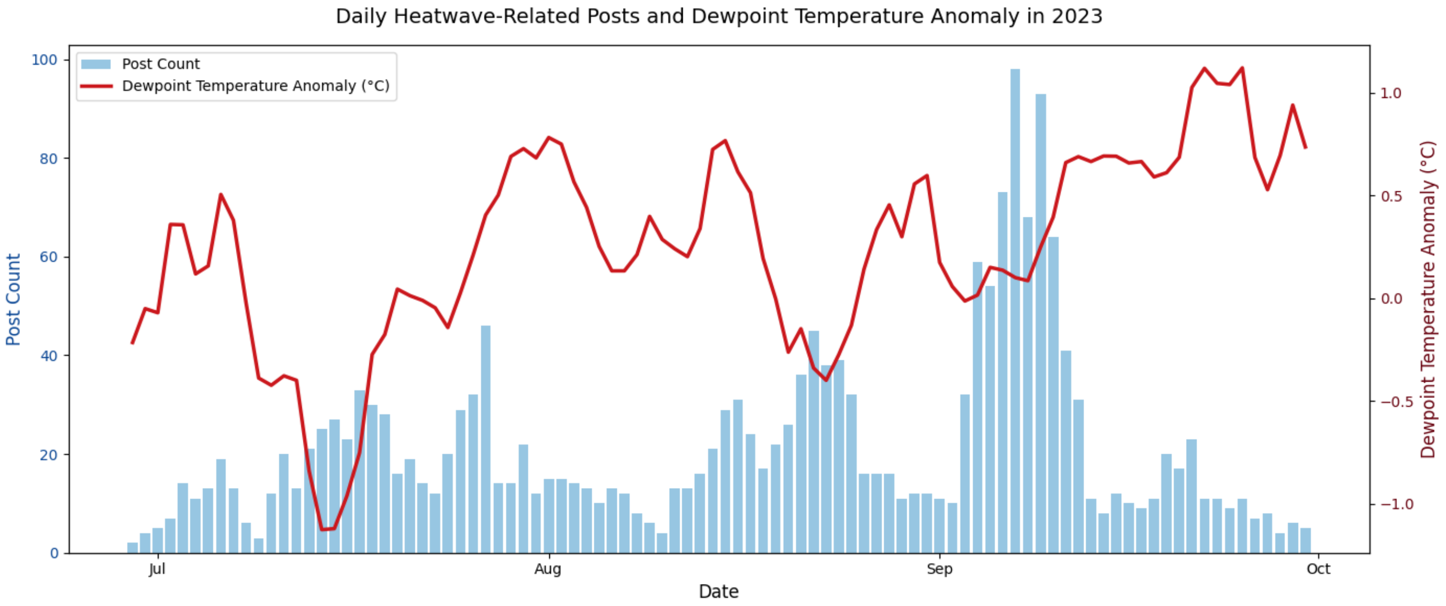


In 2024, the R^2 was 0.023 - the model explained 2.3% of the variation. The temp_anomaly coefficient was +10.88, indicating an increase of about 10 posts per 1°C rise, but again, the result was not statistically significant ($p = 0.097$).

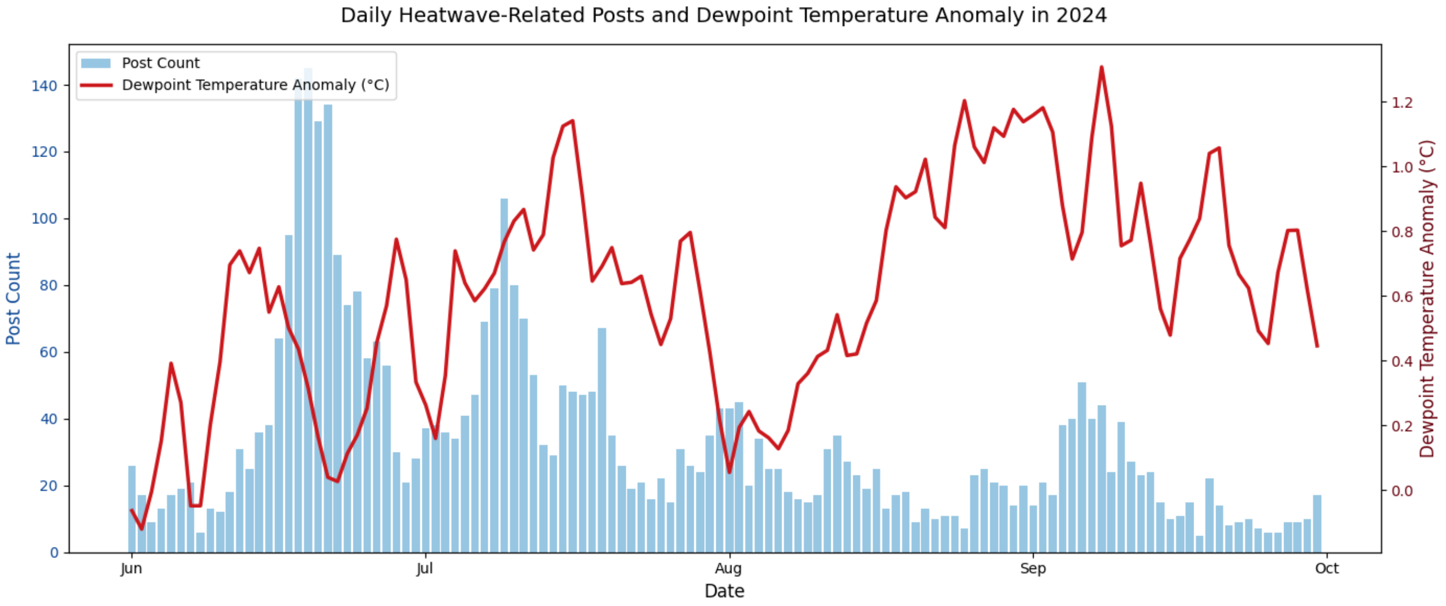
Although a positive trend appeared in 2024, the relationship between temperature anomalies and post count was not statistically significant in either year.

File 04 global temp and humidity & time series

For the time series plots, statistical analysis, and linear relationships, please refer to the document.



In 2023, the relationship between dewpoint temperature anomalies and post count had an R^2 of 0.041 — humidity anomalies explained only 4.1% of the variation in post count.
The temp_anomaly coefficient was –7.72, meaning each 1°C increase in dew point was associated with an average decrease of about 7.72 posts, the result was statistically significant ($p = 0.049$).



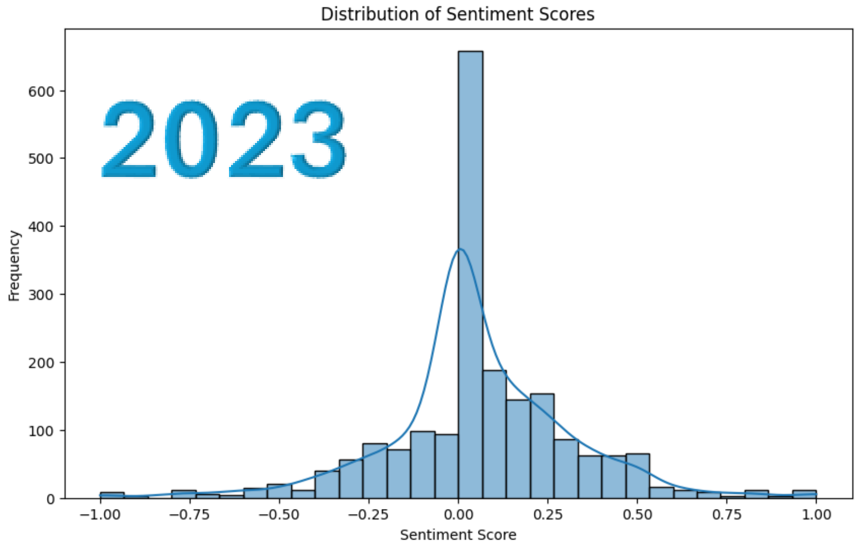
In 2024, the R^2 was 0.042 — the model explained 4.2% of the variation in post count. The temp_anomaly coefficient was –17.26, indicating a decrease of about 17 posts per 1 °C rise in dewpoint, and this result was statistically significant ($p = 0.024$).

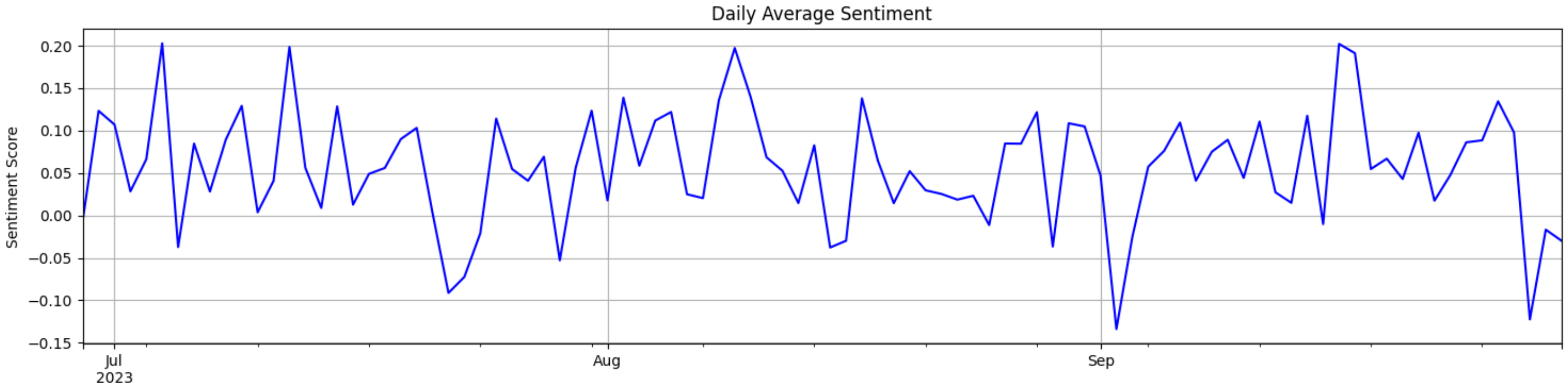
File 05 Sentimental analysis

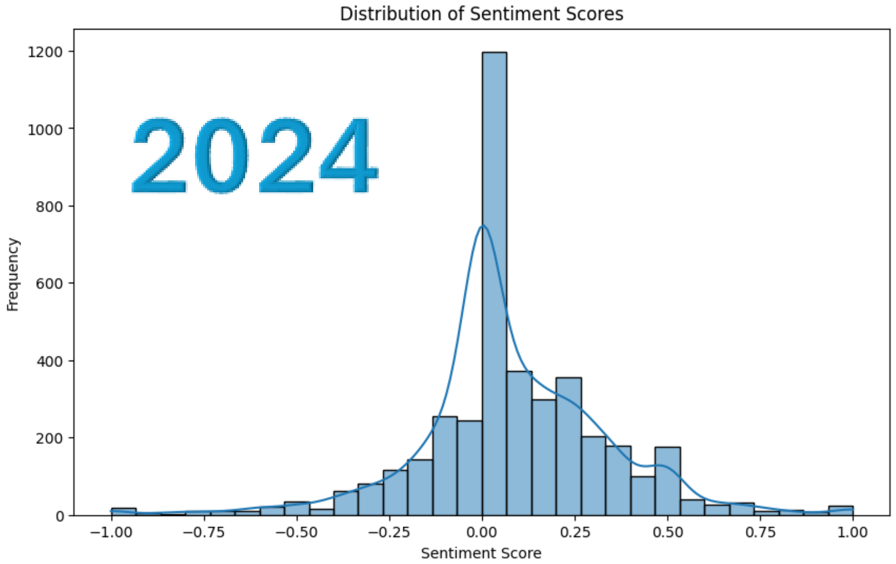
Save under the " sentimental statistics " folder

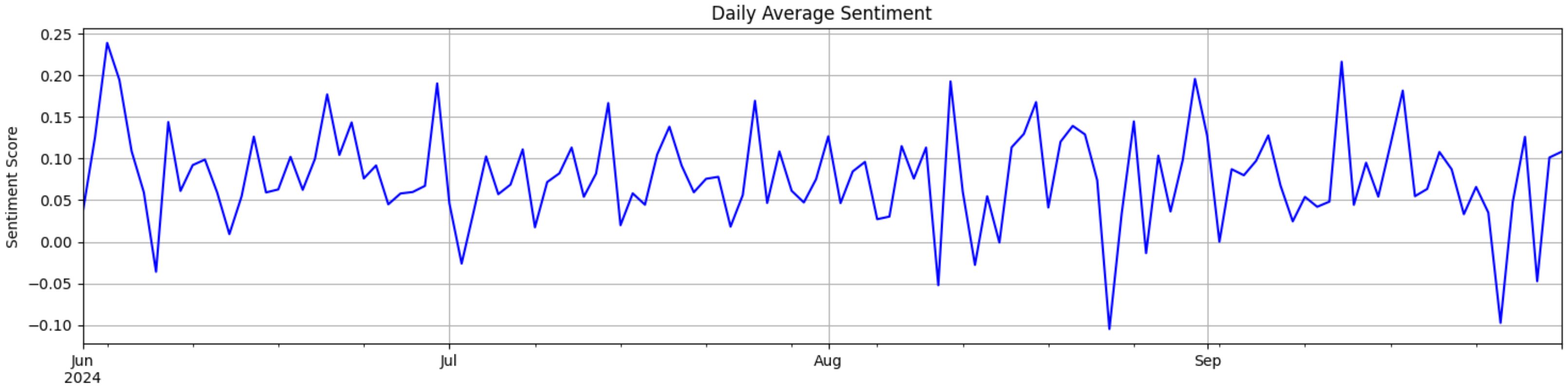
2023daily_sentiment.csv

2024daily_sentiment.csv



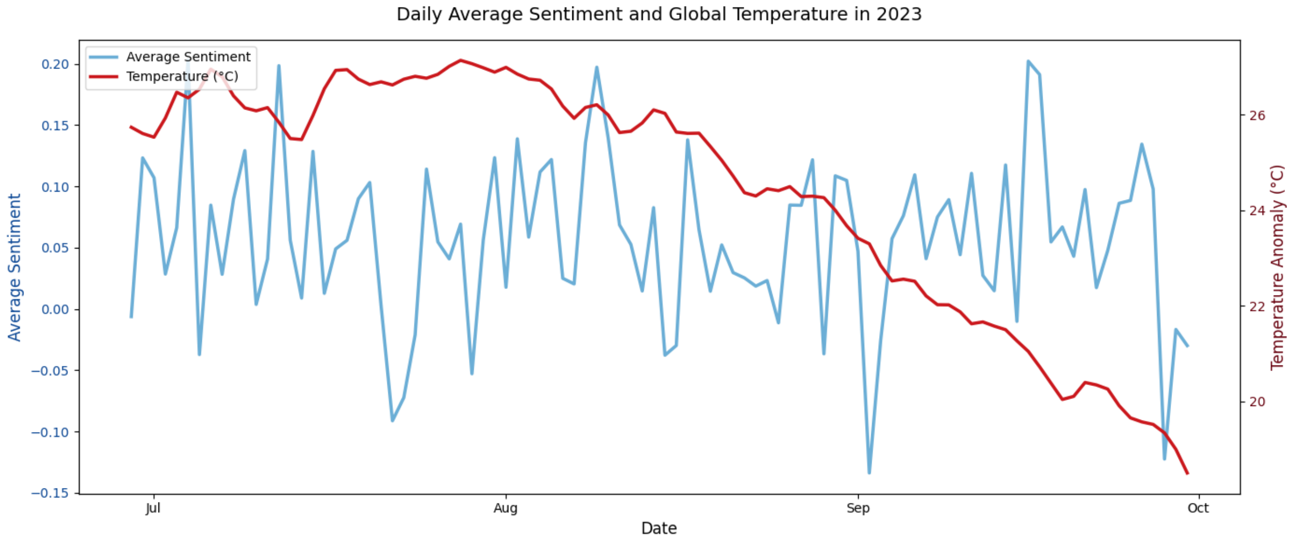




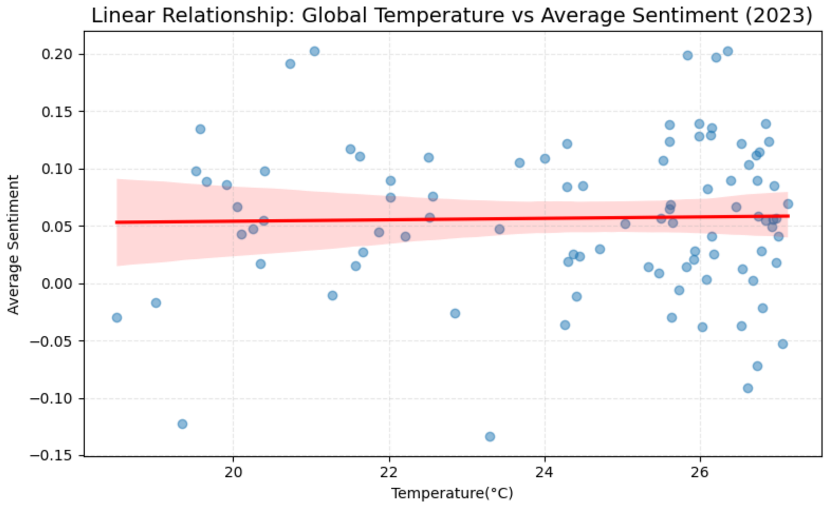


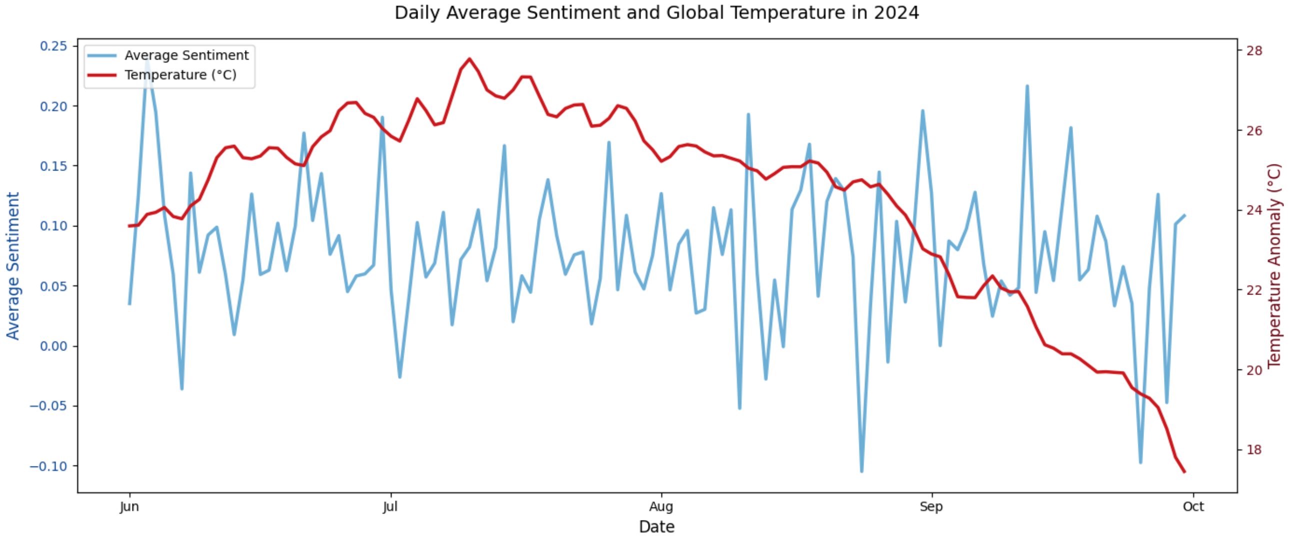
File 06A Temperature, Post Volume, and Sentiment Changes: A Statistical Investigation

6.1.Exploring the Relationship Between Global Temperature Variations and Human Sentiment: Evidence From Bluesky Data

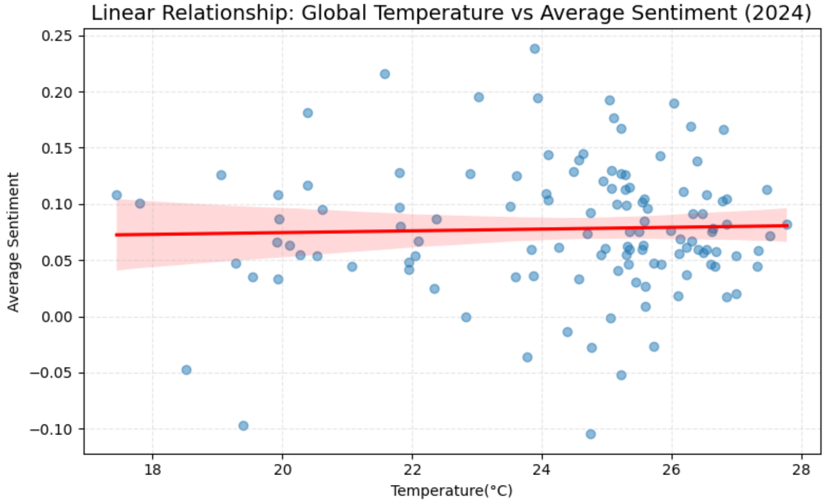


2023
R^2 = 0.001
Pearson Correlation (same day)=0.02
Pearson Correlation (lag +1 day)= 0.03



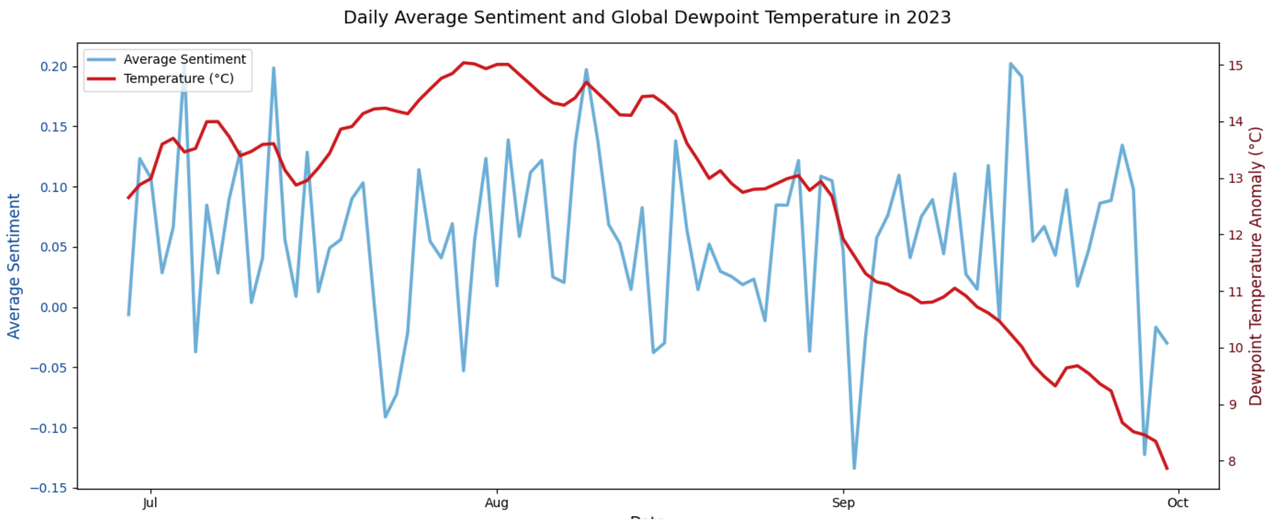


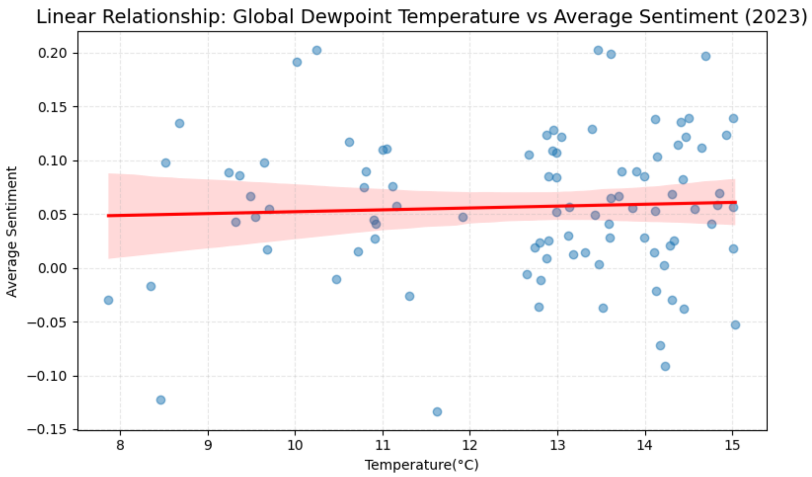
2024
R^2 = 0.001
Pearson Correlation (same day)=0.03
Pearson Correlation (lag +1 day)= 0.03

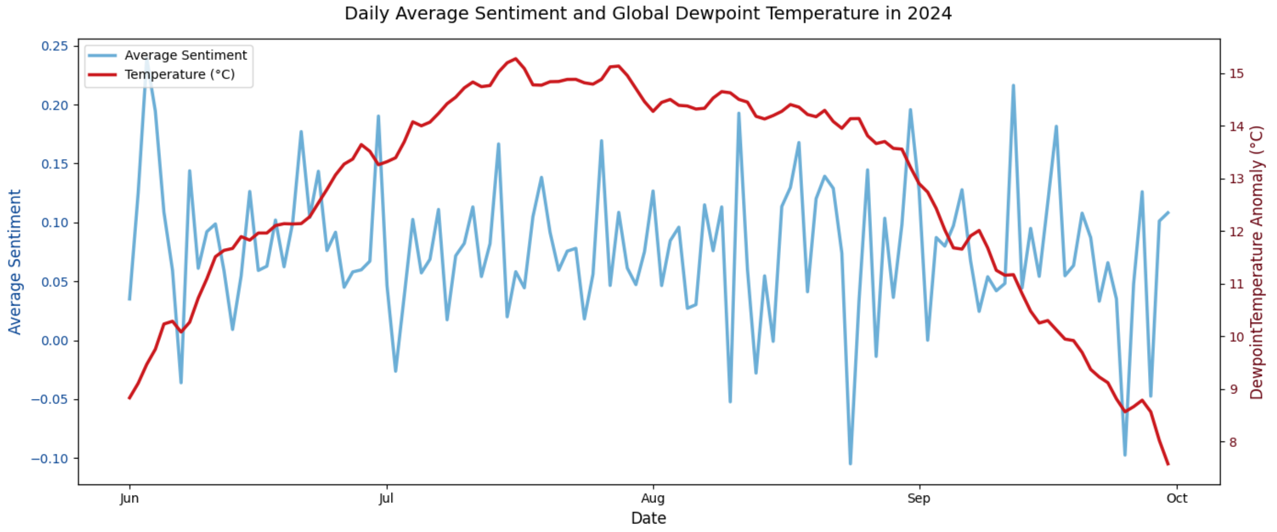


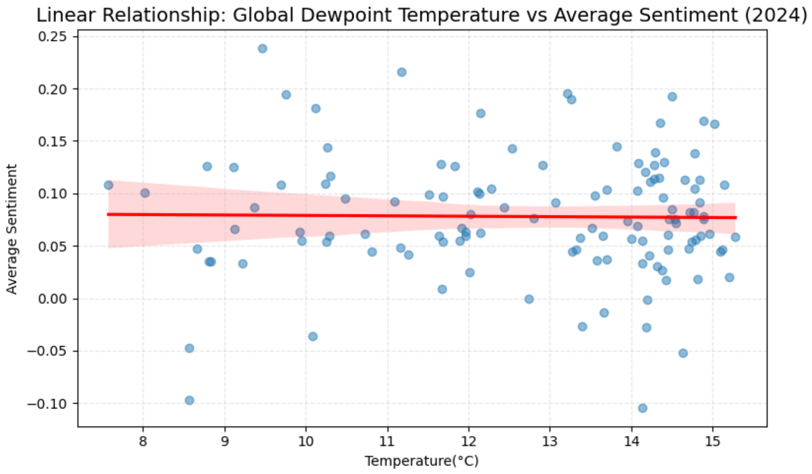
Datasets save under the ‘sentimental statistics’ folder

File 06A 6.2.The Impact of Global Dew Point Temperature (Humidity) Variations on Human Sentiment









2023

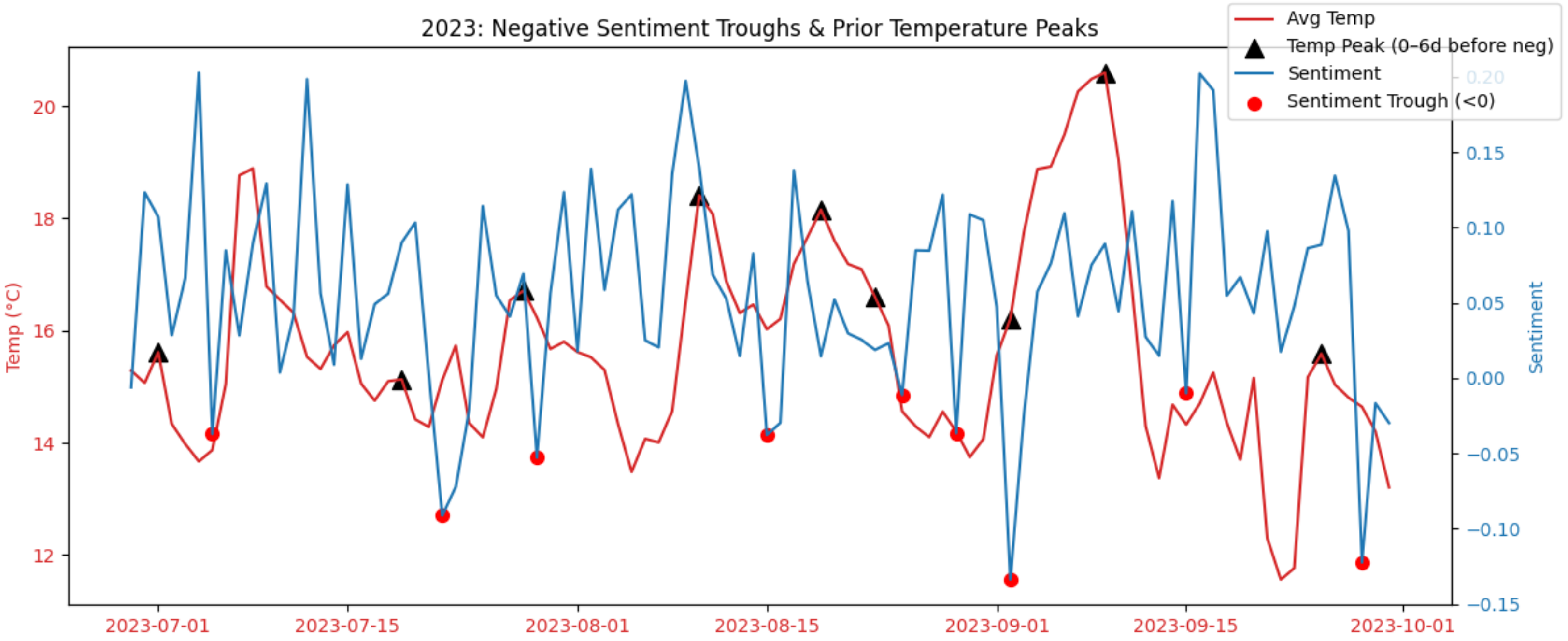
Pearson Correlation (same day): 0.05
Pearson Correlation (lag +1 day): 0.05
$R^2 = 0.002$

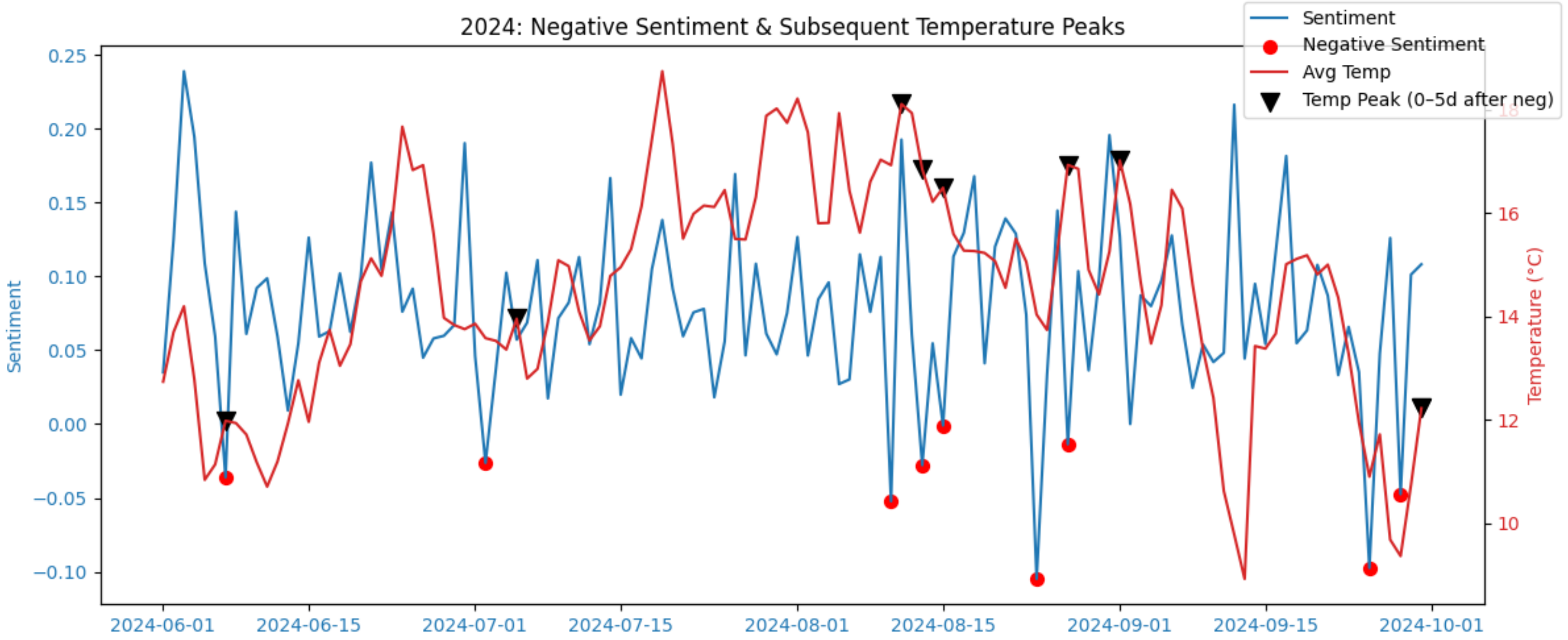
2024

Pearson Correlation (same day): -0.01
Pearson Correlation (lag +1 day): -0.04
$R^2 = 0.000$

Datasets save under the ‘sentimental statistics’ folder

File 06A 6.3. Pearson & OLS: UK Temp Around Sentiment Lows





Does peak temperature predict the lowest point of sentiment on social media?

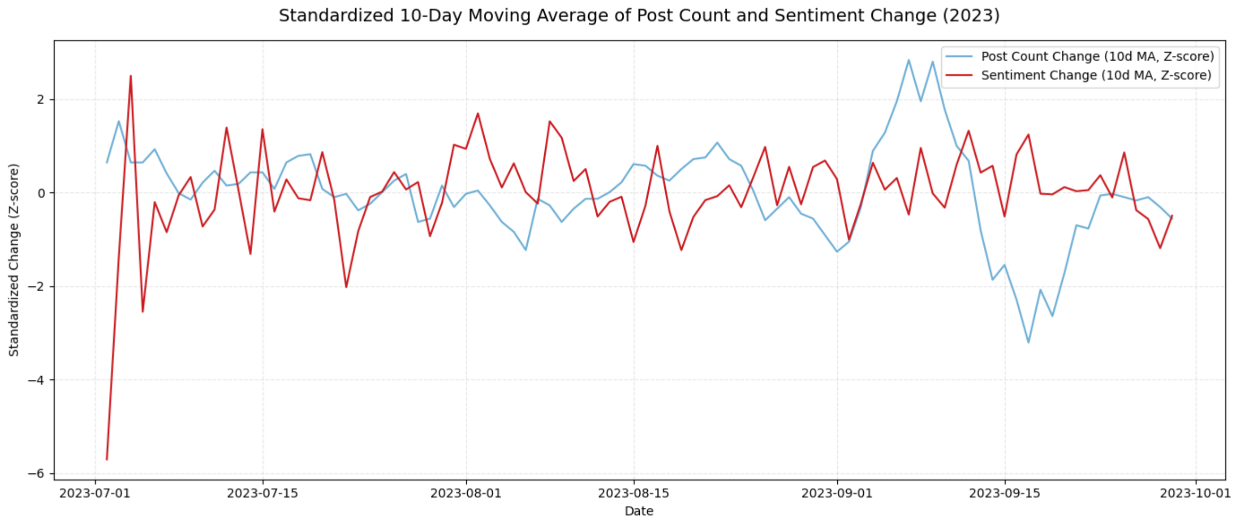
2023
$R^2 = 0.446$
Peak temperature is positively correlated with subsequent sentiment troughs on social media, and the relationship is statistically significant ($p = 0.049$). Approximately 45% of the variance in sentiment troughs can be explained by the single variable peak_temp, indicating a moderate effect of temperature on public sentiment.

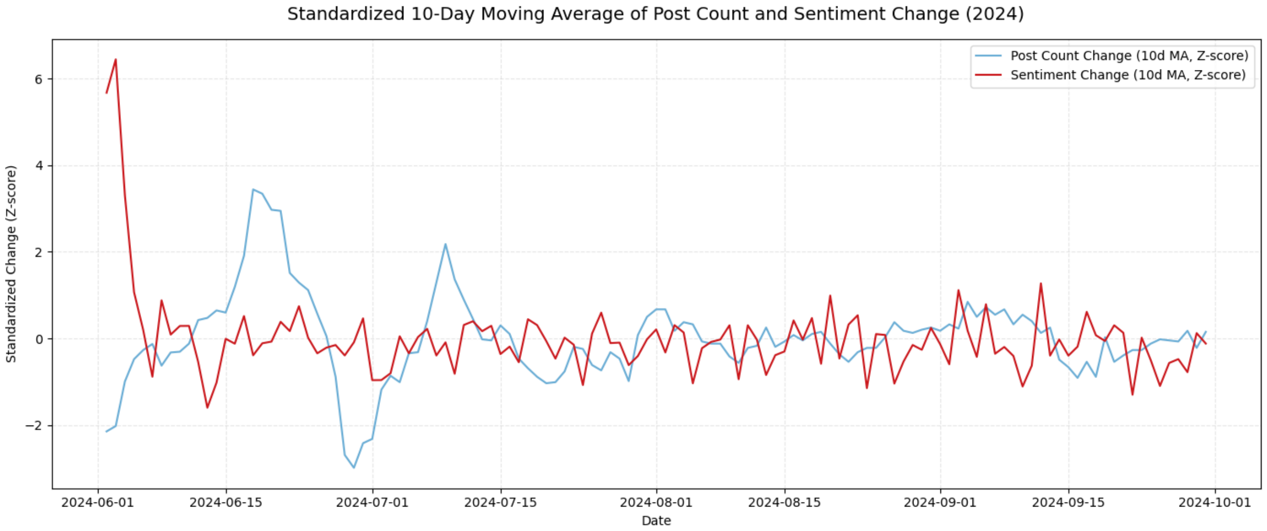
The 2024 regression analysis indicates that negative sentiment on social media does not effectively predict subsequent temperature peaks ($R^2 = 0.034$, $p = 0.636$), and the correlation is not statistically significant ($r = 0.18$).

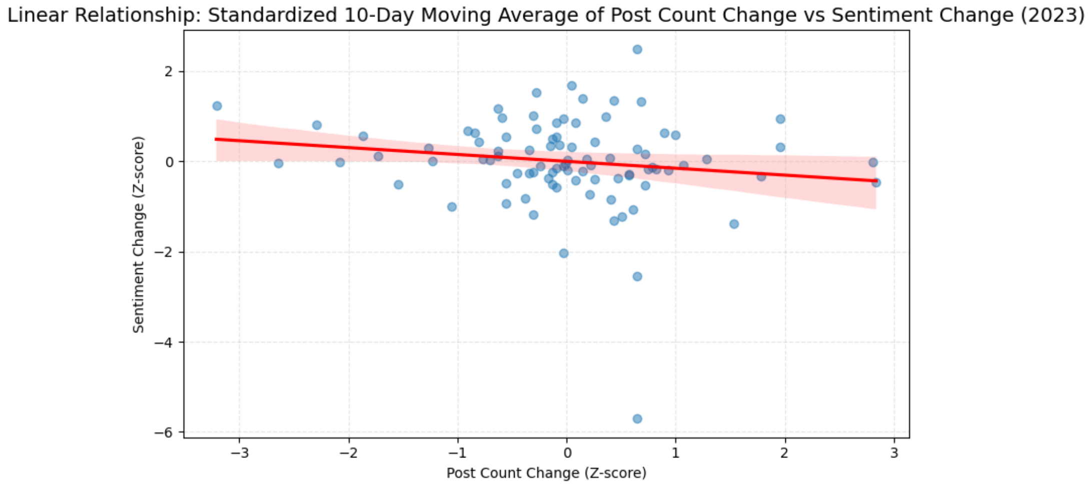
Datasets save under the ‘sentimental statistics’ folder

File 06A 6.4. Bluesky Post Count and Sentiment Fluctuations (Using 10 days moving average)

Datasets save under the 'sentimental statistics' folder





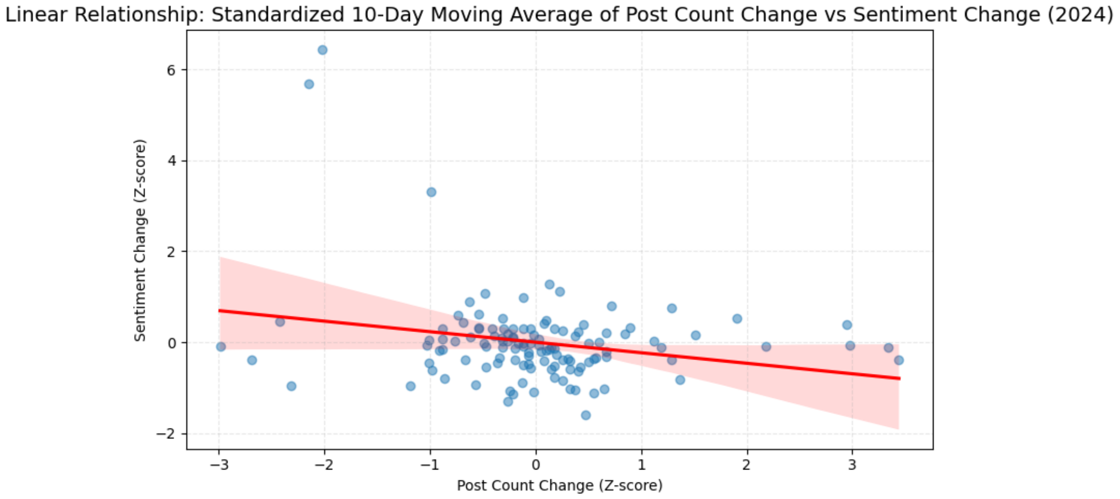


2024

Pearson Correlation (Z-score normalized, 10-day MA): -0.23

Pearson Correlation (lag +1 day, Z-score normalized, 10-day MA): -0.17

$R^2 = 0.053$



2023

Pearson Correlation (Z-score normalized, 10-day moving average): -0.15

Pearson Correlation (lag +1 day, Z-score normalized, 10-day MA): 0.03

$R^2 = 0.023$

File 06B Generalized Additive Model (GAM) 2023

API_heatwave / GAM&Bayesian / GAM2023.csv

Lixuanecho Update and rename GAM2023.csv to GAM2023.csv

Preview | Code | Blame 96 lines (96 loc) · 2.63 KB Code 55

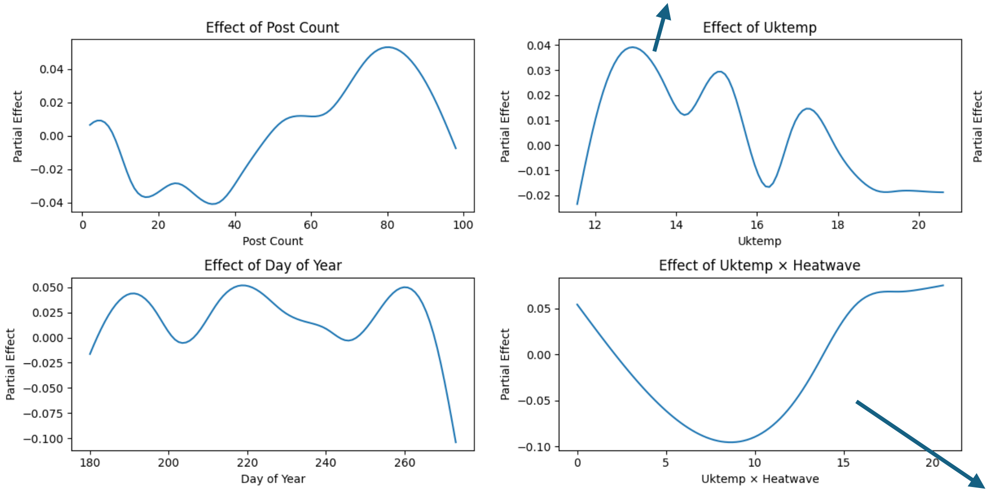
Search this file

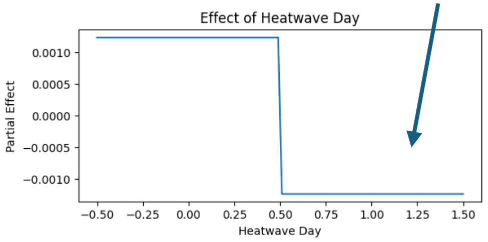
1	date	post_count	Average_Sentiment	Uktemp	heatwaveday
2	2023-06-29	2	-0.01	15.29	0
3	2023-06-30	4	0.12	15.07	0
4	2023-07-01	5	0.11	15.61	0
5	2023-07-02	7	0.03	14.33	0

```
LinearGAM
=============================================== ==========================================================
Distribution:                        NormalDist Effective DoF:                                     26.1619
Link Function:                     IdentityLink Log Likelihood:                                 -7466.1206
Number of Samples:                           94 AIC:                                             14986.565
                                                AICc:                                           15009.8017
                                                GCV:                                                0.0083
                                                Scale:                                              0.0043
                                                Pseudo R-Squared:                                   0.3048
==========================================================================================================
Feature Function                  Lambda               Rank         EDoF         P > x        Sig. Code
================================= ==================== ============ ============ ============ ============
s(0)                              [0.6]                20           10.9         7.95e-01
s(1)                              [0.6]                20           8.2          8.63e-01
f(2)                              [0.6]                2            0.9          1.58e-01
s(3)                              [0.6]                20           5.7          1.74e-01
s(4)                              [0.6]                20           0.4          2.58e-01
intercept                                              1            0.0          8.49e-01
==========================================================================================================
Significance codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Temperatures between 12–14°C were associated with more positive sentiment.

Heatwaves do have an impact on public sentiment.





The overall model fit ($R^2 = 0.305$) was relatively strong, indicating that the inclusion of seasonal trends and heatwave interaction terms substantially improved the model's explanatory power. The heatwave indicator (heatwave day) showed marginal statistical significance.

Initial high temperatures caused discomfort, but sentiment tended to ease once the adaptive threshold was surpassed.

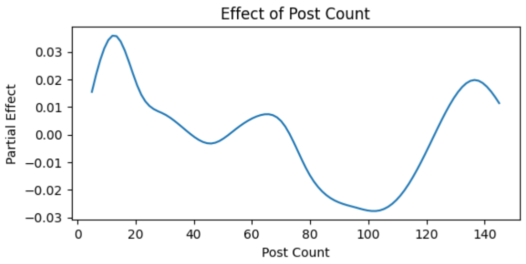
Datasets save under the 'GAM&Bayesian' folder

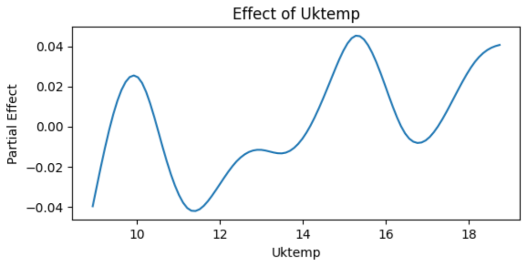
File 06B Generalized Additive Model (GAM) 2024

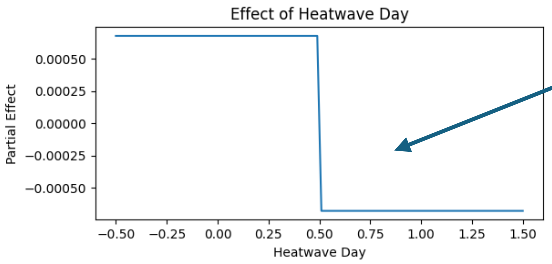
The overall model fit ($R^2 = 0.294$) was relatively strong. Temperature significantly influenced sentiment; whether a day was classified as a heatwave day led to notable differences in emotional expression; and the effect of temperature on sentiment during heatwave days was particularly strong, making it a key interaction variable.

```
LinearGAM
=============================================== ==========================================================
Distribution:                        NormalDist Effective DoF:                                      28.277
Link Function:                     IdentityLink Log Likelihood:                                -14506.8729
Number of Samples:                          122 AIC:                                            29072.2998
                                                AICc:                                            29091.628
                                                GCV:                                                0.0052
                                                Scale:                                              0.0031
                                                Pseudo R-Squared:                                   0.2944
==========================================================================================================
Feature Function                  Lambda               Rank         EDoF         P > x        Sig. Code
================================= ==================== ============ ============ ============ ============
s(0)                              [0.6]                20           10.9         7.68e-01
s(1)                              [0.6]                20           9.5          1.76e-02     *
f(2)                              [0.6]                2            0.8          1.28e-03     **
s(3)                              [0.6]                20           6.3          2.26e-01
s(4)                              [0.6]                20           0.7          4.30e-04     ***
intercept                                              1            0.0          8.69e-01
==========================================================================================================
Significance codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

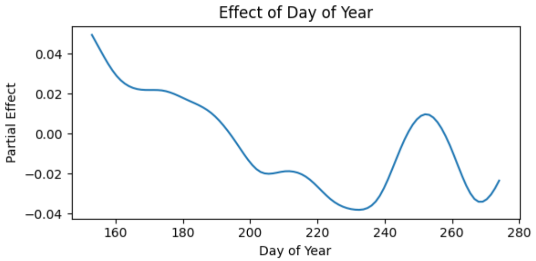
The more negative the sentiment, the more active the posting behavior.

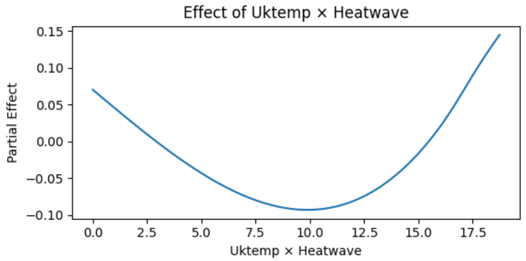






Heatwaves do have an impact on public sentiment.





Compared to 2023, the marginal effect curves in 2024 were more distinct, with stronger explanatory power for sentiment fluctuations. The interaction between heatwave days and temperature exhibited a pronounced nonlinear S/U-shaped pattern, suggesting that public sentiment shows a “phased tolerance” to environmental changes.

Datasets save under the ‘GAM&Bayesian’ folder

File 06B **Bayesian** posterior distributions

Represents the intercept term, indicating the baseline level of sentiment scores.

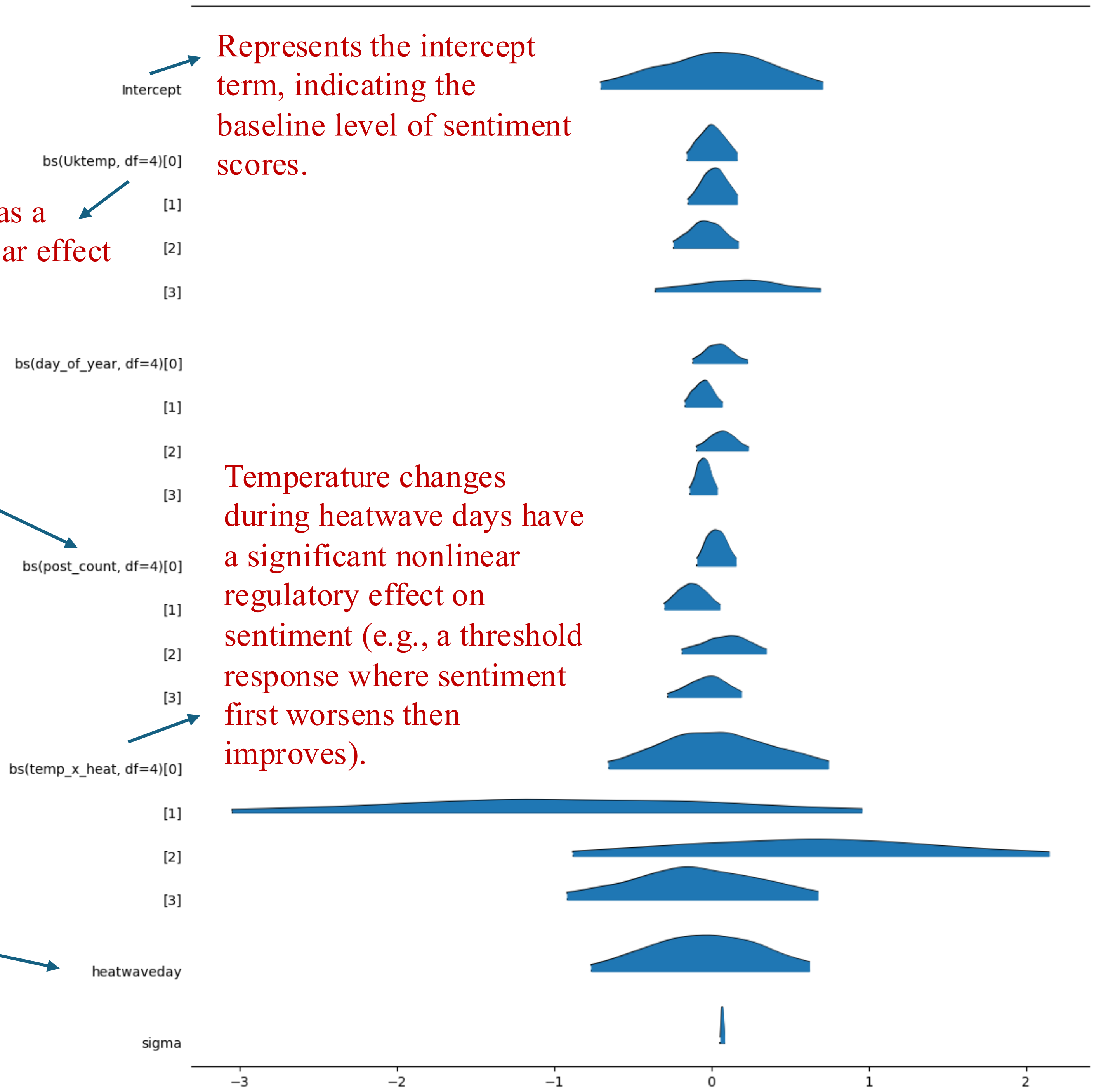
Temperature has a certain nonlinear effect on sentiment.

Seasonal patterns cause fluctuations in emotional expression.

The mid-range of post volume may have a negative effect on sentiment, supporting the hypothesis that negative emotions drive posting behavior.

Temperature changes during heatwave days have a significant nonlinear regulatory effect on sentiment (e.g., a threshold response where sentiment first worsens then improves).

Sentiment tends to be more negative on heatwave days. The overall model residuals are low, indicating good model fit.



2023

Datasets save under the ‘GAM&Bayesian’ folder

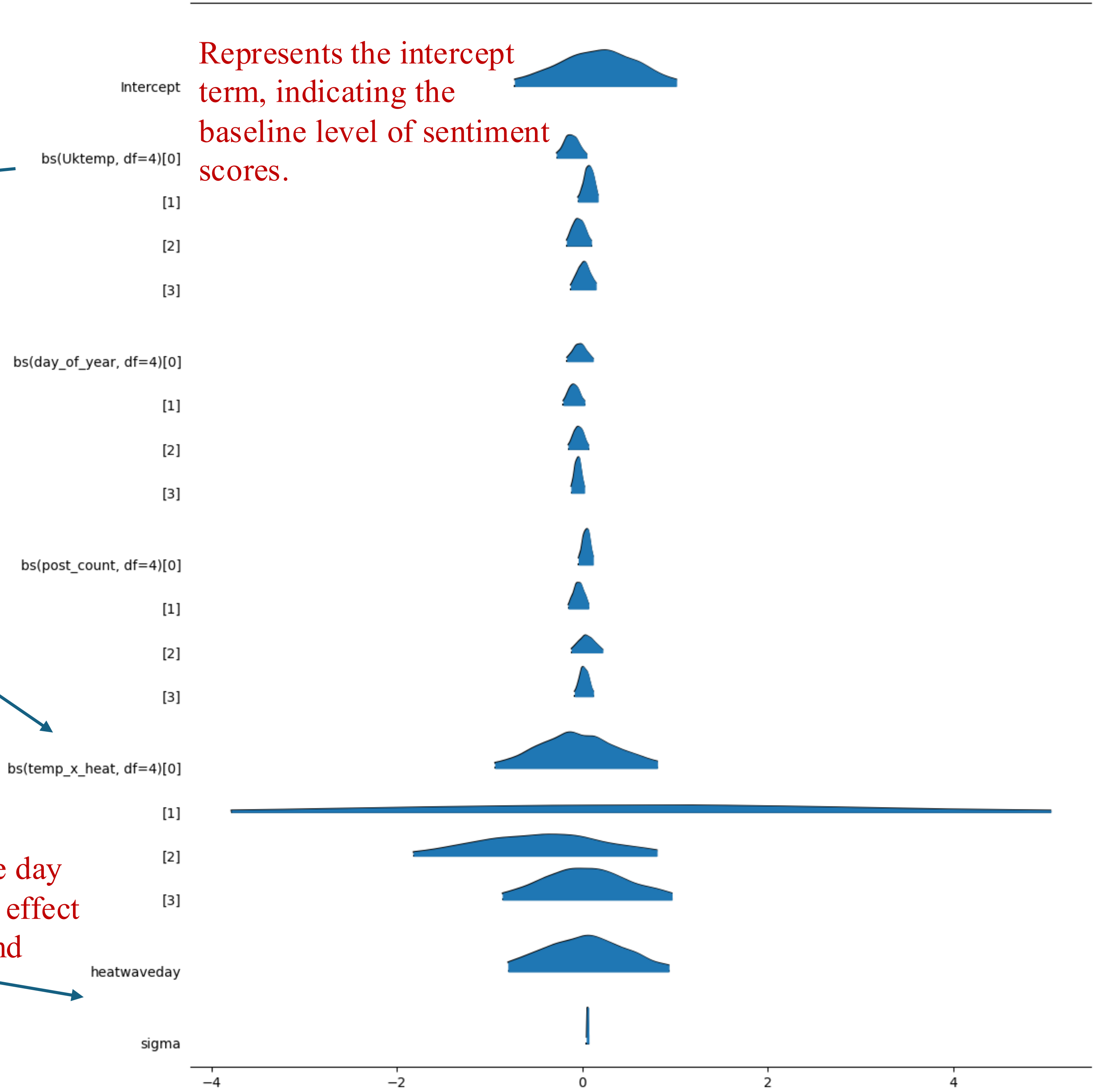
File 06B Bayesian posterior distributions

Represents the intercept term, indicating the baseline level of sentiment scores.

Compared to 2023, the posterior distributions in 2024 are more concentrated, indicating a more stable emotional response to temperature changes.

Some spline segments concentrate in the negative range, suggesting that under heatwave conditions, public sentiment exhibits a “phase-based adaptation” pattern:
Initial temperature rise → more negative emotions
Beyond an adaptation threshold → emotions stabilize or even improve

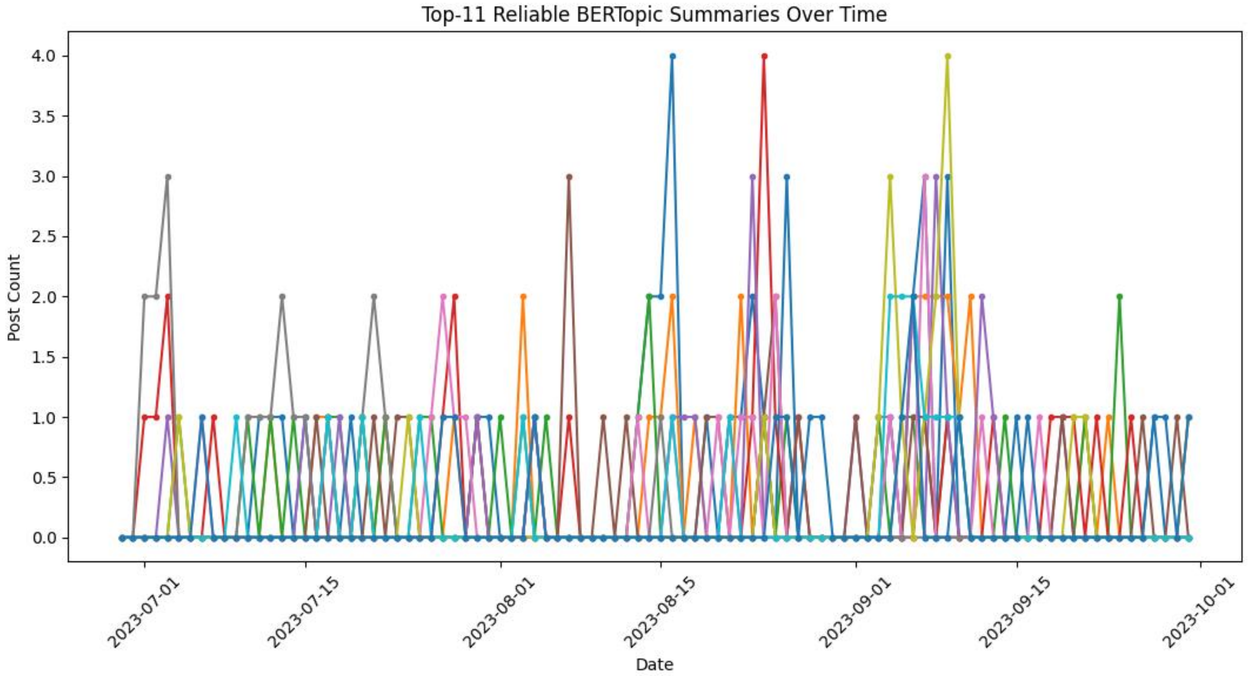
The concentrated posterior distribution for the heatwave day variable also implies that its effect on sentiment is consistent and robust.

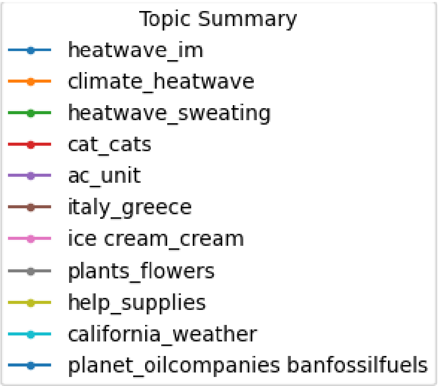


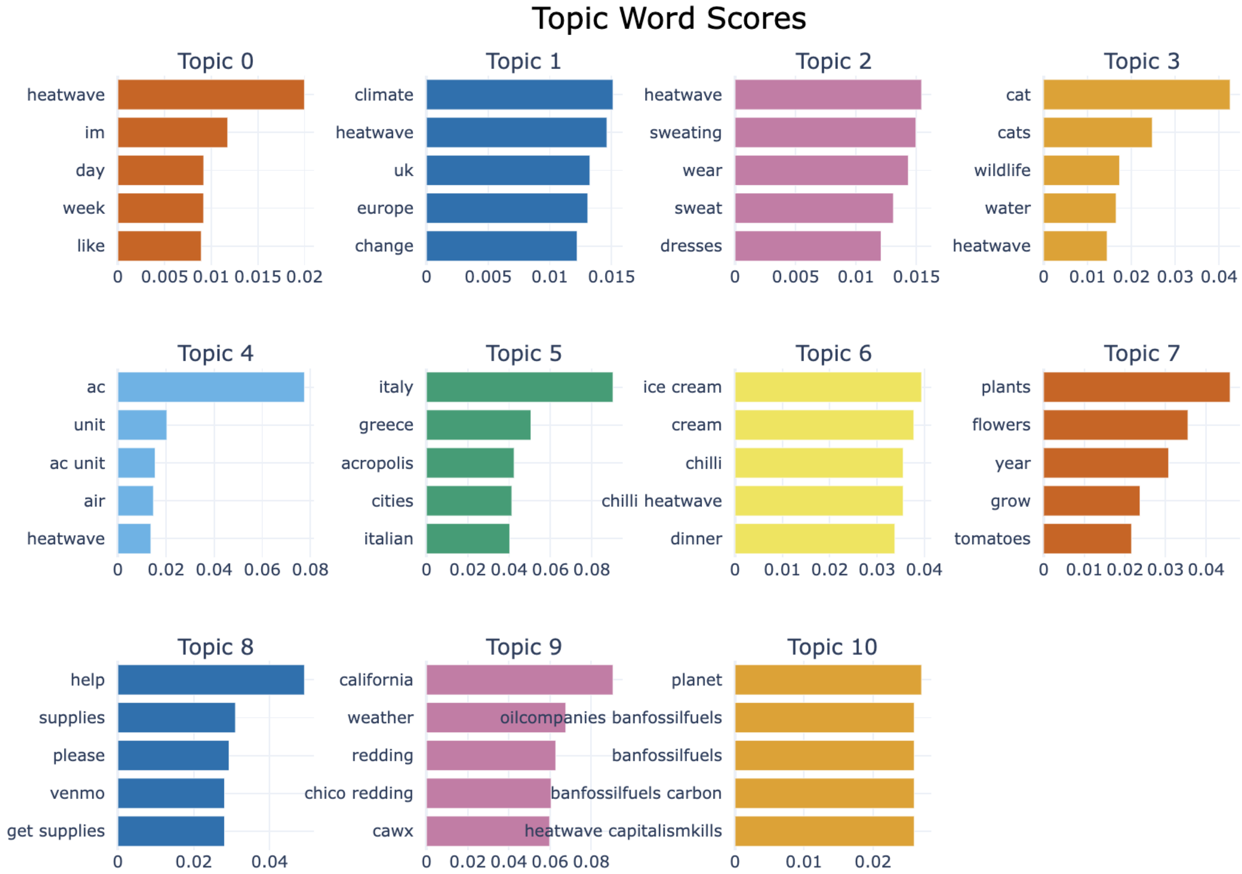
2024

Datasets save under the ‘GAM&Bayesian’ folder

File 07A Topic modelling / 2023 / BERTopic method

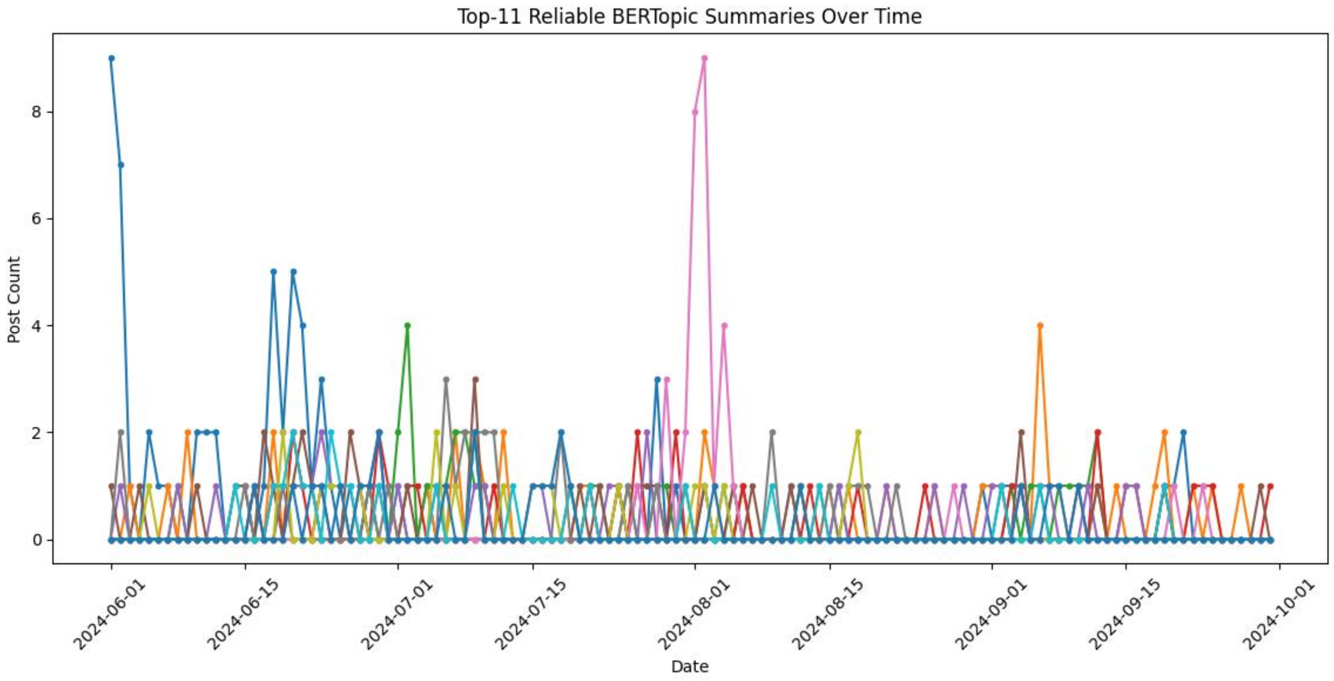


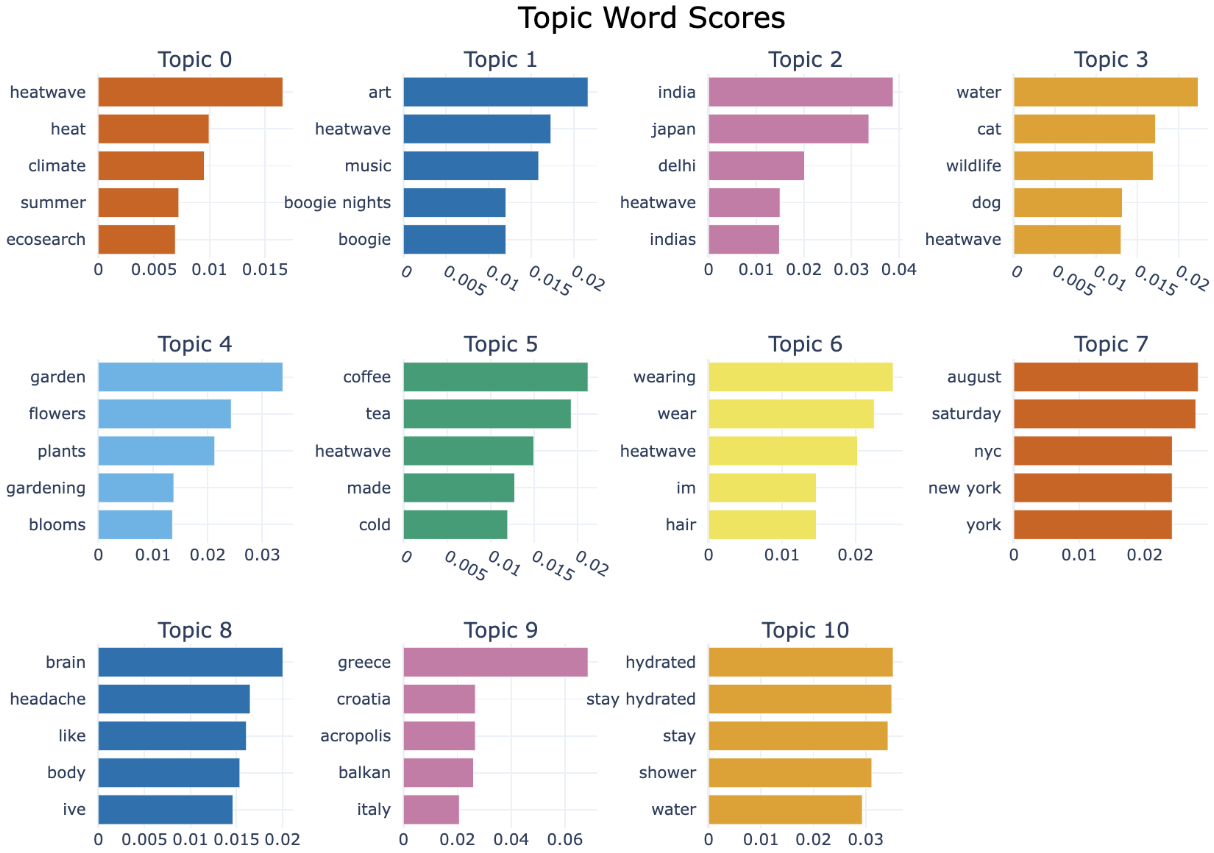




Public concern about climate change is focused on the UK, Italy, and the broader European region. People not only express discomfort caused by heatwaves but also show concern for the impacts on animals and plants, as well as a growing dependence on cooling measures such as air conditioning and ice cream. The discussion also extends to calls for mutual aid, attention to extreme weather events in places like California, and broader reflections on the future of the global climate, fossil fuel use, and the health of our planet’s ecosystems.

File 07A Topic modelling / 2024 / BERTopic method





The discussion focuses on heatwaves in regions such as India and Japan, along with concerns about the survival of animals and plants and the strain on water resources. People also express their emotional responses to extreme heat through culture, music and art. Common coping strategies such as consuming cold drinks, coffee, and adjusting clothing or hairstyles highlight individual-level adaptation. Some topics mention physical discomforts like headaches and fatigue, reflecting the interplay between emotional and physiological responses to heat. Additionally, there is public discussion about high-temperature experiences during travel in Mediterranean countries, involving themes of tourism and cultural heritage.

File 07B Topic modelling / 2023 cluster / Structured topic model method （LDAbate method）

Topic	Summary (Interpretation)	Proper Name
Topic 1	Discussions about the ongoing heatwave into late summer (September), experiences with heat, and hopes for relief.	Late Summer Heatwave Experiences and Hopes for Relief
Topic 2	Strong concerns about fires, death, and the ending of the heatwave, with emotional discussions especially in Texas and Chicago.	Fires, Fatalities, and Emotional Responses to Heatwaves
Topic 3	Immediate reactions to heatwaves, struggles to adapt, mentions of future change (autumn, cooling down), and resilience.	Adaptation Efforts and Hope for Seasonal Change
Topic 4	The impact of heatwaves on daily life — work, home life, feeling trapped, struggling at night with high temperatures.	Everyday Struggles with Heat: Work, Home, and Health
Topic 5	Comparisons of heat this year to previous years, complaints about sweating, rain scarcity, and overall record-breaking conditions.	Record Heat, Lack of Rain, and Yearly Comparisons
Topic 6	Official weather updates and warnings — focused on California (NorCal, Chico), emphasizing heat advisories and forecasts.	Weather Warnings and Heat Advisories in California

File 07B Topic modelling / 2024 cluster / Structured topic model method （LDAbate method）

Topic	Summary (Interpretation)	Proper Name
Topic 1	People sharing personal feelings about the heatwave — describing the sensation of heat inside/outside, dealing with discomfort at home, appreciating small reliefs like opening windows.	Personal Experiences of Heat Discomfort and Relief Efforts
Topic 2	Discussions centered around climate change, ecological crises, wildfires, and scientific news linked to heatwaves — awareness of broader environmental issues.	Climate Change, Wildfires, and Heatwave Awareness
Topic 3	Complaints and emotional reactions to enduring the heat — expressions of frustration, exhaustion, and hope for rain or an end to the heatwave.	Public Frustrations and Emotional Coping During Heatwaves
Topic 4	General discussions about heatwave duration, upcoming weather patterns, seasonal changes, and advice on staying cool and managing expectations.	Heatwave Duration, Cooling Strategies, and Seasonal Outlooks
Topic 5	Coping struggles and expressions of gratitude — mentions of noise, emotional strain, physical conditions, and efforts to stay hydrated and comfortable.	Physical and Emotional Coping Challenges During Heatwaves
Topic 6	Serious concerns about extreme temperatures and health risks — records being broken, calls for hydration, and urgent warnings (even mentioning Japan).	Health Risks, Record Temperatures, and Urgent Public Warnings