Global Surface Temperature

Anomalies Analysis

Name: M. Govindu S. Sathruwan

Table of Contents

Introduction	IV
Research Question	5
Relevance:	5
Data Source:	5
Data:	5
Data Preparation	6
Data Tidying:	6
Missing Values:	8
Outliers:	9
Data Type:	10
Season observations:	11
Exploratory Data Analysis (EDA)	12
Univariate Analysis	12
Multivariate Analysis	16
Data Storytelling	24
Introduction	24
Trends Over Time	24
Seasonal Patterns	25
Geographical Patterns	26
Conclusion	28
References	29

Introduction

Climate change has become one of the most pressing global challenges, with rising temperatures posing significant environmental, economic, and societal threats. Understanding the extent and patterns of global temperature changes is crucial for effective decision-making and policy implementation. This project aims to analyze historical surface temperature anomalies across different geographical zones and time periods, focusing on long-term trends, seasonal variations, and regional disparities.

By leveraging data from NASA's Goddard Institute for Space Studies (GISS), this study explores how Earth's surface temperature anomalies have evolved over time. Using advanced data visualization techniques, we uncover critical insights into the warming patterns affecting different parts of the world. Through this analysis, we seek to provide a data-driven perspective on the progression of global warming, highlighting the urgency for sustainable actions to mitigate its impact.

Research Question

How have Earth's surface temperature anomalies varied across different geographical zones, and what patterns can be observed in the monthly and seasonal variations of the surface temperature anomalies?

Relevance:

Global warming is a growing problem that has risen in popularity recently. Therefore, it is crucial to examine historical data to understand the true extent of this issue and its recent prominence.

Furthermore, uncovering the long-term trends and patterns in global temperature anomalies is crucial for understanding climate change. By analysing data across different geographical zones and temporal scales, we can gain insights into regional and seasonal variations in temperature anomalies.

Data Source:

The data for this research will be estimates of global surface temperature change from NASA's Goddard Institute for Space Studies (GISS).

- GISTEMP Team, 2024: GISS Surface Temperature Analysis (GISTEMP), version 4. NASA Goddard Institute for Space Studies. Dataset accessed 2024-12-14 at https://data.giss.nasa.gov/gistemp/.
- Lenssen, N., G.A. Schmidt, M. Hendrickson, P. Jacobs, M. Menne, and R. Ruedy, 2024: A GISTEMPv4 observational uncertainty ensemble. J. Geophys. Res. Atmos., 129, no. 17, e2023JD040179, doi:10.1029/2023JD040179.

Data:

The following are files in tabular format of temperature anomalies, i.e. deviations from the corresponding 1951-1980 means.

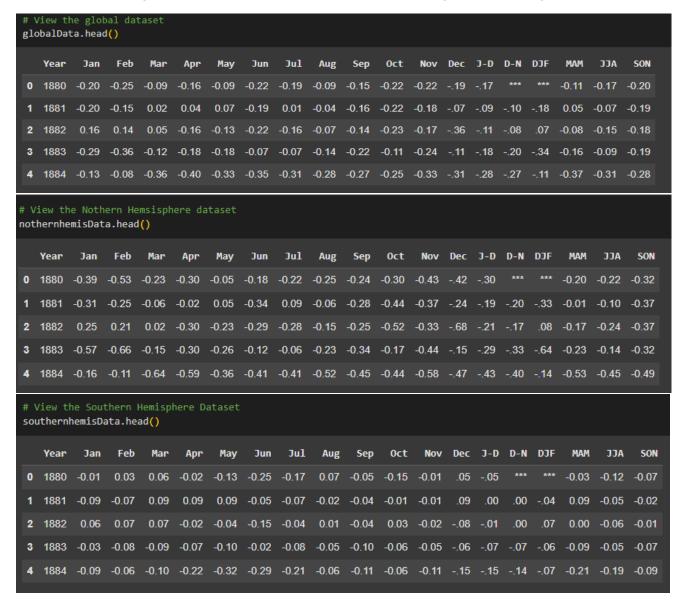
- Global-mean monthly, seasonal, and annual means, 1880-present, updated through most recent month: https://data.giss.nasa.gov/gistemp/tabledata_v4/GLB.Ts+dSST.csv
- Southern Hemisphere-mean monthly, seasonal, and annual means, 1880-present, updated through most recent month:
 - https://data.giss.nasa.gov/gistemp/tabledata_v4/SH.Ts+dSST.csv
- Northern Hemisphere-mean monthly, seasonal, and annual means, 1880-present, updated through most recent month:
 - https://data.giss.nasa.gov/gistemp/tabledata_v4/NH.Ts+dSST.csv
- **Zonal annual means**, 1880-present, updated through most recent complete year: https://data.giss.nasa.gov/gistemp/tabledata_v4/ZonAnn.Ts+dSST.csv

Data Preparation

To ensure that the data obtained from the above sources are ready for analysis, several cleaning and transformation techniques will be applied to the data to meet the requirements for the analysis. Python (Colab) was the tool used for data preprocessing in this project.

Data Tidying:

These four datasets were not tidy; thus, the melt function was used to transform the data from a wide format to a long, tidy format. They were then sorted to be arranged in chronological order.



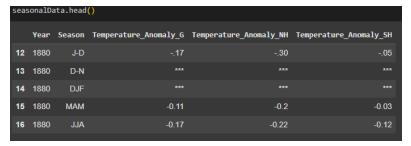
Year Month Temperature_Anomaly_G Year Month Temperature_Anomaly_NH Year Month Temperature_Anomaly_SH 1880 Jan -0.2 1880 Feb -0.53 1880 Jan -0.01 1880 Mar -0.09 1880 Mar -0.23 1880 Mar 0.06 1880 May -0.016 1880 May -0.05 1880 May Apr -0.02 1880 Jun -0.22 1880 Jun -0.05 1880 May -0.01 1880 Jul -0.19 1880 Jul -0.02 1880 Jul -0.17 1880 Aug -0.09 1880 Aug -0.25 1880 Jul -0.17 1880 Sep -0.15 1880 Aug -0.25 1880 Aug 0.07 1880 Sep -0.15 1880 Oct -0.3 1880 Nov -0.05 1880 Nov -0.22 1880 Nov -0.43 1880 Nov -0.05 1880 Dec -1.17 1880 Dec -0.43 1880 Nov -0.01 1880 DF *** 1880 Dec -0.02 1880 Dec -0.05 1880 DF <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>								
1880 Jan -0.2 1880 Jan -0.61 1880 Feb -0.25 1880 Feb 0.03 1880 Mar -0.09 1880 Mar -0.23 1880 Mar 0.06 1880 Apr -0.16 1880 Apr -0.3 1880 Apr -0.02 1880 May -0.09 1880 May -0.05 1880 May -0.13 1880 Jun -0.22 1880 Jun -0.25 1880 Jun -0.25 1880 Jul -0.19 1880 Jul -0.22 1880 Jul -0.25 1880 Aug -0.09 1880 Sep -0.24 1880 Sep -0.07 1880 Sep -0.15 1880 Sep -0.24 1880 Sep -0.05 1880 Oct -0.15 1880 Nov -0.43 1880 Nov -0.05	Year	Month	Temperature Anomaly G				Month	Temperature_Anomaly_SH
1880 Feb -0.25 1880 Feb -0.53 1880 Feb 0.03 1880 Mar -0.09 1880 Mar -0.23 1880 Mar 0.06 1880 Apr -0.16 1880 Apr -0.3 1880 Apr -0.02 1880 May -0.05 1880 May -0.03 1880 Jun -0.21 1880 Jun -0.22 1880 Jun -0.25 1880 Jul -0.19 1880 Jul -0.22 1880 Jul -0.17 1880 Aug -0.09 1880 Aug -0.25 1880 Aug 0.07 1880 Sep -0.15 1880 Sep -0.24 1880 Sep -0.05 1880 Oct -0.22 1880 Oct -0.3 1880 Oct -0.15 1880 Dec -1.17 1880 Dec -4.2 1880			. – 1880					
1880 Mar -0.09 1880 Mar -0.23 1880 Mar 0.66 1880 Apr -0.16 1880 Apr -0.3 1880 Apr -0.02 1880 May -0.05 1880 May -0.13 1880 Jun -0.18 1880 Jun -0.25 1880 Jul -0.19 1880 Jul -0.22 1880 Jul -0.17 1880 Aug -0.09 1880 Aug -0.25 1880 Aug 0.07 1880 Sep -0.15 1880 Sep -0.24 1880 Sep -0.05 1880 Oct -0.22 1880 Nov -0.43 1880 Oct -0.05 1880 Dec 19 1880 Dec 42 1880 Dec .05 1880 Dec 17 1880 Dec 42 1880 Dec .05			-0.25				Feb	
1880 Apr			-0.09			1880	Mar	0.06
1880 May -0.09 1880 May -0.13 1880 May -0.13 1880 Jun -0.22 1880 Jul -0.18 1880 Jul -0.25 1880 Aug -0.09 1880 Aug -0.22 1880 Aug 0.07 1880 Sep -0.15 1880 Sep -0.24 1880 Sep -0.05 1880 Oct -0.22 1880 Oct -0.3 1880 Oct -0.15 1880 Nov -0.43 1880 Nov -0.15 1880 Dec 42 1880 Nov -0.01 1880 Dec 42 1880 Nov -0.01 1880 J-D 17 1880 Dec 42 1880 Dec .05 1880 J-D 17 1880 D-N **** 1880 D-N **** 1880 DF **** <t< td=""><td>1880</td><td>Apr</td><td>-0.16</td><td></td><td></td><td></td><td>Apr</td><td></td></t<>	1880	Apr	-0.16				Apr	
1880 Jun -0.22 1880 Jun -0.18 1880 Jun -0.25 1880 Jul -0.19 1880 Jul -0.17 1880 Aug -0.25 1880 Aug 0.67 1880 Sep -0.15 1880 Sep -0.24 1880 Sep -0.05 1880 Oct -0.22 1880 Oct -0.3 1880 Oct -0.15 1880 Dec -0.22 1880 Nov -0.43 1880 Oct -0.15 1880 Dec -1.19 1880 Dec 42 1880 Dec .05 1880 J-D 30 1880 J-D .05 1880 J-D .05 1880 D-N **** 1880 D-N **** 1880 D-N **** 1880 DJF **** 1880 DJF **** 1880 DJF **** 1880 <t< td=""><td></td><td></td><td>1990</td><td>May</td><td>-0.05</td><td>1880</td><td>May</td><td>-0.13</td></t<>			1990	May	-0.05	1880	May	-0.13
1880 Jul -0.19 1880 Jul -0.22 1880 Jul -0.17 1880 Aug -0.09 1880 Aug -0.25 1880 Aug 0.67 1880 Sep -0.15 1880 Sep -0.24 1880 Sep -0.05 1880 Oct -0.22 1880 Oct -0.3 1880 Oct -0.15 1880 Nov -0.43 1880 Nov -0.01 1880 Dec 19 1880 Dec 42 1880 Dec .05 1880 J-D 17 1880 J-D 30 1880 J-D 05 1880 D-N **** 1880 D-N **** 1880 D-N **** 1880 DJF **** 1880 DJF *** 1880 DJF *** 1880 JJA -0.11 1880 JJA -0.22 1880 <td< td=""><td></td><td></td><td>1000</td><td>Jun</td><td>-0.18</td><td>1880</td><td>Jun</td><td>-0.25</td></td<>			1000	Jun	-0.18	1880	Jun	-0.25
1880 Aug -0.09 1880 Aug -0.25 1880 Aug 0.07 1880 Sep -0.15 1880 Sep -0.24 1880 Sep -0.05 1880 Oct -0.3 1880 Oct -0.15 1880 Nov -0.43 1880 Nov -0.61 1880 Dec 42 1880 Dec .05 1880 J-D 17 1880 J-D 30 1880 J-D 05 1880 D-N **** 1880 D-N **** 1880 DJF **** 1880 D-N **** 1880 DJF **** 1880 DJF **** 1880 DJF **** 1880 DJF **** 1880 MAM -0.11 1880 MAM -0.2 1880 MAM -0.03 1880 SON -0.2 1880 SON -0.32 </td <td></td> <td></td> <td>1990</td> <td>Jul</td> <td>-0.22</td> <td>1880</td> <td>Jul</td> <td>-0.17</td>			1990	Jul	-0.22	1880	Jul	-0.17
1880 Sep -0.15 1880 Sep -0.24 1880 Sep -0.65 1880 Oct -0.22 1880 Oct -0.3 1880 Oct -0.15 1880 Nov -0.43 1880 Nov -0.61 1880 Dec 42 1880 Dec .05 1880 J-D 17 1880 J-D 30 1880 J-D 05 1880 D-N **** 1880 D-N **** 1880 D-N **** 1880 DJF **** 1880 DJF **** 1880 DJF **** 1880 MAM -0.11 1880 MAM -0.2 1880 MAM -0.03 1880 JJA -0.03 1880 JJA -0.12 1881 Jan -0.07			1000	Aug	-0.25	1880	Aug	0.07
1880 Oct -0.22 1880 Oct -0.15 1880 Nov -0.22 1880 Nov -0.43 1880 Nov -0.91 1880 Dec -1.19 1880 Dec -4.2 1880 Dec .05 1880 J-D -30 1880 J-D -0.5 1880 D-N *** 1880 D-N *** 1880 DJF *** 1880 DJF *** 1880 MAM -0.11 1880 MAM -0.2 1880 MAM -0.03 1880 JJA -0.17 1880 JJA -0.22 1880 JJA -0.12 1880 SON -0.2 1880 SON -0.32 1880 SON -0.07 1881 Jan -0.2 1881 Jan -0.31 1881 Jan -0.09 1881 Mar 0.06 1881 Mar 0.06 1881			1000	Sep	-0.24	1880	Sep	-0.05
1880 Nov -0.22 1880 Nov -0.43 1880 Nov -0.01 1880 Dec 19 1880 Dec 42 1880 Dec .05 1880 J-D 17 1880 J-D 85 1880 J-D 05 1880 D-N **** 1880 D-N **** 1880 D-N **** 1880 DJF **** 1880 DJF **** 1880 DJF **** 1880 MAM -0.11 1880 MAM -0.2 1880 MAM -0.03 1880 JJA -0.17 1880 JJA -0.22 1880 JJA -0.12 1880 SON -0.22 1880 SON -0.32 1880 SON -0.07 1881 Jan -0.15 1881 Jan -0.31 1881 Jan -0.07 1881 Mar 0.02 1881 <			1000	0ct	-0.3	1880	0ct	-0.15
1880 Dec 19 1880 Dec 42 1880 Dec .05 1880 J-D 17 1880 J-D 30 1880 J-D 05 1880 D-N **** 1880 D-N **** 1880 DJF *** 1880 DJF **** 1880 DJF *** 1880 DJF **** 1880 MAM -0.11 1880 MAM -0.2 1880 MAM -0.03 1880 JJA -0.17 1880 JJA -0.22 1880 JJA -0.12 1880 SON -0.2 1880 SON -0.32 1880 SON -0.07 1881 Jan -0.2 1881 Jan -0.07 1881 Jan -0.09 1881 Feb -0.15 1881 Feb -0.25 1881 Jan -0.07 1881 Mar 0.02 1881			4000	Nov	-0.43	1880	Nov	-0.01
1880 J-D 17 1880 J-D 30 1880 J-D 05 1880 D-N **** 1880 D-N **** 1880 DJF **** 1880 DJF **** 1880 DJF **** 1880 DJF **** 1880 MAM -0.2 1880 DJF **** 1880 JJA -0.2 1880 MAM -0.63 1880 JJA -0.12 1880 JJA -0.12 1881 Jan -0.2 1881 Jan -0.07 1881 Jan -0.31 1881 Jan -0.09 1881 Feb -0.25 1881 Feb -0.07 1881 Mar -0.06 1881 Mar 0.09 1881 Apr 0.09 1881 Apr 0.09			1000	Dec	42	1880	Dec	.05
1880 D-N **** 1880 D-N **** 1880 DJF **** 1880 DJF **** 1880 MAM -0.11 1880 DJF **** 1880 MAM -0.2 1880 MAM -0.31 1880 SON -0.12 1880 SON -0.32 1880 SON -0.07 1881 Jan -0.2 1881 Jan -0.31 1881 Jan -0.09 1881 Feb -0.15 1881 Feb -0.25 1881 Feb -0.07 1881 Mar 0.02 1881 Mar -0.06 1881 Mar 0.09 1881 Apr 0.04 1881 Apr -0.02 1881 Apr 0.09			1000	J-D	30	1880	J-D	
1880 DJF *** 1880 DJF **** 1880 DJF **** 1880 MAM -0.11 1880 MAM -0.2 1880 MAM -0.03 1880 JJA -0.12 1880 JJA -0.12 1880 SON -0.32 1880 SON -0.07 1881 Jan -0.21 1881 Jan -0.31 1881 Jan -0.09 1881 Feb -0.15 1881 Feb -0.25 1881 Feb -0.07 1881 Mar 0.02 1881 Mar -0.06 1881 Mar 0.09 1881 Apr 0.04 1881 Apr -0.02 1881 Apr 0.09				D-N	***	1880	D-N	***
1880 MAM -0.11 1880 MAM -0.2 1880 MAM -0.03 1880 JJA -0.17 1880 JJA -0.22 1880 JJA -0.12 1880 SON -0.32 1880 SON -0.07 1881 Jan -0.2 1881 Jan -0.07 1881 Feb -0.25 1881 Feb -0.07 1881 Mar 0.02 1881 Mar -0.06 1881 Mar 0.09 1881 Apr 0.04 1881 Apr -0.02 1881 Apr 0.09			*** 1880	DJF	***	1880	DJF	***
1880 JJA -0.17 1880 JJA -0.22 1880 JJA -0.12 1880 SON -0.2 1880 SON -0.32 1880 SON -0.67 1881 Jan -0.2 1881 Jan -0.09 1881 Feb -0.25 1881 Feb -0.67 1881 Mar -0.06 1881 Mar 0.09 1881 Apr 0.04 1881 Apr -0.02 1881 Apr 0.09			-9 11 1880	MAM	-0.2	1880	MAM	-0.03
1880 SON -0.2 1880 SON -0.32 1880 SON -0.07 1881 Jan -0.31 1881 Jan -0.09 1881 Feb -0.25 1881 Feb -0.07 1881 Mar -0.06 1881 Mar 0.09 1881 Apr 0.04 1881 Apr -0.02 1881 Apr 0.09 1881 Apr -0.02 1881 Apr 0.09			4000	JJA	-0.22	1880	JJA	-0.12
1881 Jan -0.2 1881 Jan -0.31 1881 Jan -0.09 1881 Feb -0.25 1881 Feb -0.07 1881 Mar 0.02 1881 Mar -0.06 1881 Mar 0.09 1881 Apr 0.09 1881 Apr 0.09 1881 Apr 0.09			4000	SON	-0.32	1880	SON	-0.07
1881 Feb -0.25 1881 Feb -0.07 1881 Mar 0.02 1881 Mar -0.06 1881 Mar 0.09 1881 Apr 0.04 1881 Apr -0.02 1881 Apr 0.09 1881 Apr -0.02 1881 Apr 0.09			4004	Jan	-0.31	1881	Jan	-0.09
1881 Mar 0.02 1881 Mar -0.06 1881 Mar 0.09 1881 Apr 0.02 1881 Apr 0.09 1881 Apr -0.02 1881 Apr 0.09			4004	Feb	-0.25	1881	Feb	-0.07
1881 Apr 9.94 1881 Apr -0.02 1881 Apr 9.09			4004	Mar	-0.06	1881	Mar	0.09
1991 May 9 00				Apr	-0.02	1881	Apr	0.09
1881 May 0.07 1881 May 0.05 1881 May 0.09	1881	May	0.07 1881	May	0.05	1881	May	0.09

<pre># View the Zonal Dataset zonalData.head()</pre>															
	Year	Glob	NHem	SHem	24N-90N	24S-24N	905-245	64N-90N	44N-64N	24N-44N	EQU-24N	24S-EQU	445-245	64S-44S	905-645
0	1880	-0.17	-0.30	-0.05	-0.40	-0.12	-0.01	-0.81	-0.52	-0.31	-0.13	-0.11	-0.04	0.05	0.68
1	1881	-0.09	-0.19	0.00	-0.37	0.11	-0.07	-0.93	-0.49	-0.21	0.11	0.10	-0.05	-0.07	0.60
2	1882	-0.11	-0.21	-0.01	-0.32	-0.04	0.01	-1.41	-0.31	-0.14	-0.03	-0.05	0.01	0.04	0.64
3	1883	-0.18	-0.29	-0.07	-0.36	-0.16	-0.01	-0.17	-0.59	-0.27	-0.17	-0.16	-0.04	0.07	0.51
4	1884	-0.28	-0.43	-0.15	-0.61	-0.14	-0.14	-1.30	-0.67	-0.47	-0.12	-0.17	-0.19	-0.02	0.66
1 1 1	/ear 1880 1881 1882 1883	Zor Glo Glo Glo Glo	ob ob ob	Temp	eratur	 -(-(naly 3.17 3.09 3.11 3.18 3.28								

The first three datasets reveal similar variables: Year, Month and Average temperature anomaly globally, northern hemisphere and southern hemisphere, respectively. Therefore, to improve efficiency, these three datasets have been merged, leaving the last columns separate. Following that, the combined dataset was then split into monthly data and seasonal data based on the Month column to ensure that the monthly averages are stored separately from the seasonal averages, which would improve the ease of analysing this data. Three finalised datasets have now been obtained for the analysis.

comb	oined_c	lata.he	ad (45)		
	Year	Month	Temperature_Anomaly_G	Temperature_Anomaly_NH	Temperature_Anomaly_SH
0	1880	Jan	-0.2	-0.39	-0.01
1	1880	Feb	-0.25	-0.53	0.03
2	1880	Mar	-0.09	-0.23	0.06
3	1880	Apr	-0.16	-0.3	-0.02
4	1880	Мау	-0.09	-0.05	-0.13
5	1880	Jun	-0.22	-0.18	-0.25
6	1880	Jul	-0.19	-0.22	-0.17
7	1880	Aug	-0.09	-0.25	0.07
8	1880	Sep	-0.15	-0.24	-0.05
9	1880	Oct	-0.22	-0.3	-0.15
10	1880	Nov	-0.22	-0.43	-0.01
11	1880	Dec	19	42	.05
12	1880	J-D	17	30	05
13	1880	D-N	***	***	***
14	1880	DJF	***	***	***
15	1880	MAM	-0.11	-0.2	-0.03
16	1880	JJA	-0.17	-0.22	-0.12
17	1880	SON	-0.2	-0.32	-0.07
18	1881	Jan	-0.2	-0.31	-0.09





Moving on, these three finalised data sets (monthly, seasonal, and zonal) can now be cleaned by checking and resolving missing values, outliers, and datatype errors.

Missing Values:

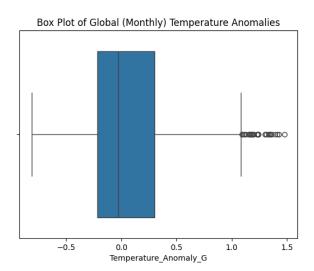
• It was discovered that the missing values had "***" assigned as placeholders in these datasets. Thus, they were removed and replaced with NA so that they would not interfere with any plots or calculations. Moreover, these few missing values are due to a lack of data beyond the time frame of the analysis, i.e. before 1880 and for 2024 December.

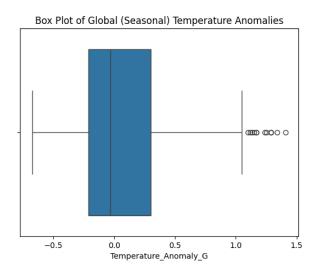


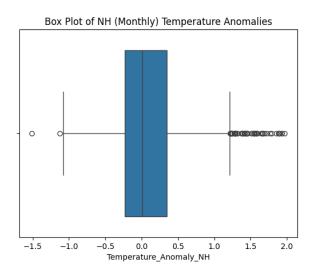
Replace missing value placeholders with NaN
monthlyData.replace(['***'], pd.NA, inplace=True)
seasonalData.replace(['***'], pd.NA, inplace=True)
zonalData.replace(['***'], pd.NA, inplace=True)

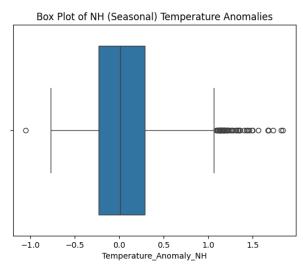
Outliers:

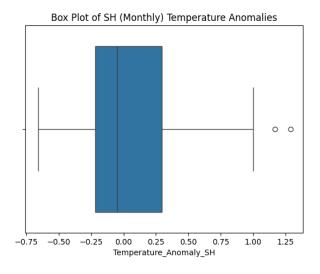
• Box plots were used to observe for any outliers in these datasets. While there were figures that went beyond the upper and lower bounds, none of these values were such extremes that they had to be erroneous data. Thus, no action was taken on those data points.

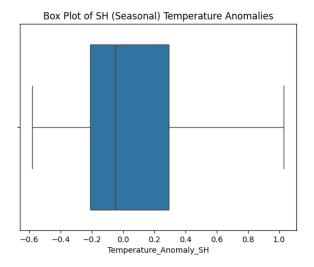


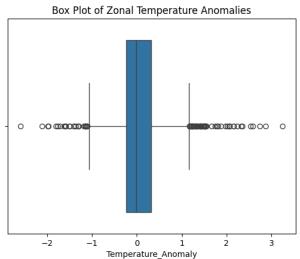






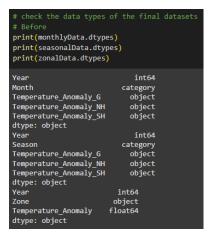






Data Type:

• To ensure the consistency and reliability of these datasets for the analysis, the data type of some of the temperature anomaly variables was changed so that these variables had uniform and appropriate data types.



```
check the data types of the final datasets after conversion
print(monthlyData.dtypes)
print(seasonalData.dtypes)
print(zonalData.dtypes)
                                      int64
                                  category
Temperature_Anomaly_G
Temperature_Anomaly_NH
Temperature_Anomaly_SH
                                   float64
                                   float64
dtype: object
Year
                                      int64
Season
                                  category
Temperature_Anomaly_G
Temperature_Anomaly_NH
Temperature_Anomaly_SH
                                   float64
                                   float64
dtype: object
Year
                              object
float64
Zone
Temperature_Anomaly
dtype: object
```

Season observations:

- J-D January to December
- D-N December of the previous year to November of the current year
- DJF December, January, February; Winter in NH and Summer in SH
- MAM March, April, May; Spring in NH and Autumn in SH
- JJA June, July, August; Summer in NH and Winter in SH
- SON September, October, November; Autumn in NH and Spring in SH

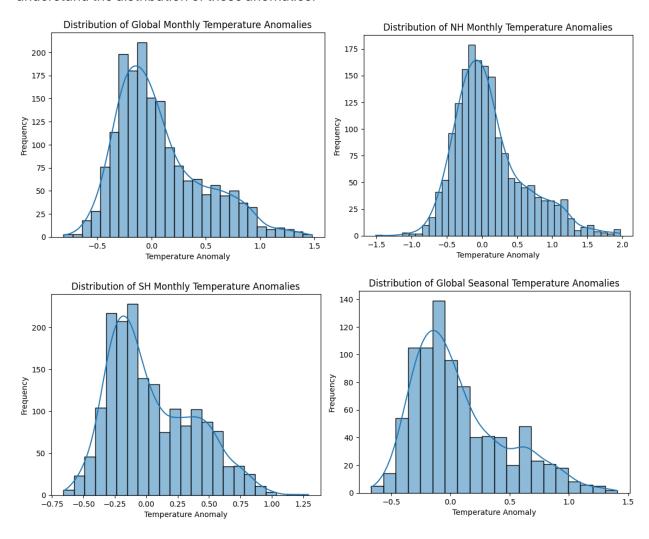
Exploratory Data Analysis (EDA)

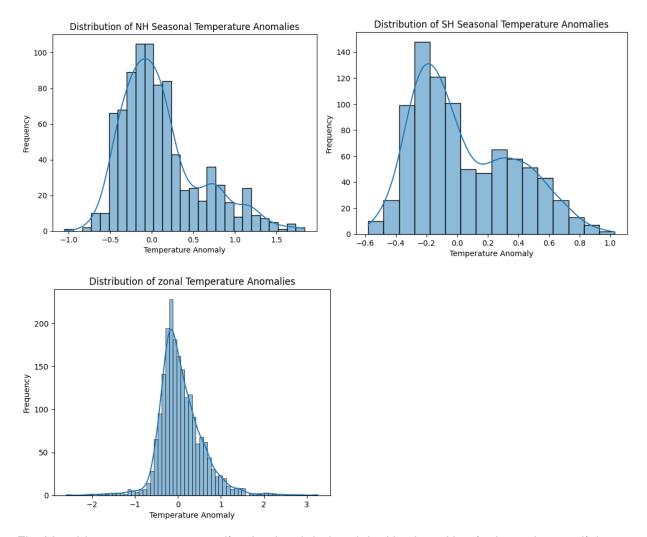
The purpose of the EDA is to explore the data to understand its structure and discover potential trends and patterns that might aid in answering the research question. It enables us to gain insights into the data.

Univariate Analysis

This is an analysis of individual variables, and several techniques have been used to carry out this analysis, including box plots, histograms and descriptive statistics.

First and foremost, histograms will be plotted for all the temperature anomaly variables to understand the distribution of these anomalies.





The Monthly temperature anomalies for the global and the Northern Hemisphere show a slight positive skewness as the tail extends to the positive side, meaning that there are more anomalies slightly lower than or equal to the base year averages. Furthermore, the zonal temperature anomalies could be said to follow a normal distribution as the graph is symmetrical around 0, which would be the mean. However, the rest of the temperature anomalies follow a bimodal distribution with two distinct peaks, one below 0 and the other above 0. The zonal anomalies could also be said to have the most temperature variance as their values range over 5 degrees compared to the other anomalies, which range over 2 – 3 degrees.

Following this, observing these variables' descriptive statistics could help further understand them.

	Year	Temperature_Anomaly_G	Temperature_Anomaly_NH	Temperature_Anomaly_SH
count	1740.000000	1739.000000	1739.000000	1739.000000
mean	1952.000000	0.075509	0.108712	0.043122
std	41.868932	0.402125	0.503651	0.338282
min	1880.000000	-0.810000	-1.510000	-0.660000
25%	1916.000000	-0.220000	-0.230000	-0.220000
50%	1952.000000	-0.030000	0.010000	-0.050000
75%	1988.000000	0.300000	0.350000	0.295000
max	2024.000000	1.480000	1.970000	1.290000

Figure 1: Descriptive Stats - Monthly Data

	Year	Temperature_Anomaly_G	Temperature_Anomaly_NH	Temperature_Anomaly_SH
count	870.000000	867.000000	867.000000	867.000000
mean	1952.000000	0.074729	0.107935	0.042295
std	41.880976	0.391821	0.479988	0.327609
min	1880.000000	-0.670000	-1.050000	-0.580000
25%	1916.000000	-0.210000	-0.230000	-0.210000
50%	1952.000000	-0.030000	0.010000	-0.050000
75%	1988.000000	0.300000	0.290000	0.295000
max	2024.000000	1.410000	1.840000	1.030000

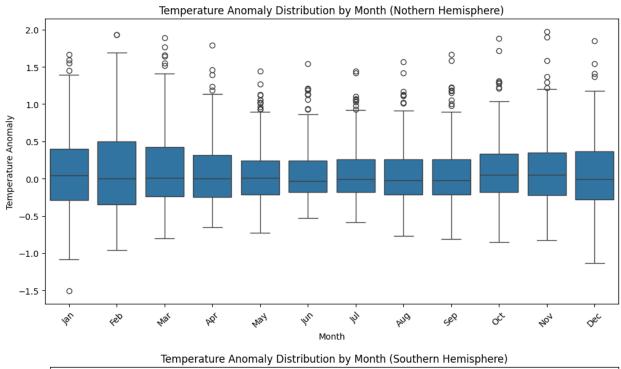
Figure 2: Descriptive Stats - Seasonal Data

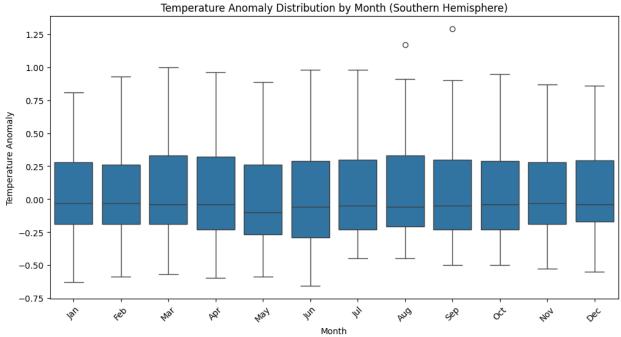
	Year	Temperature_Anomaly
count	2016.00000	2016.000000
mean	1951.50000	0.067594
std	41.57853	0.511211
min	1880.00000	-2.590000
25%	1915.75000	-0.240000
50%	1951.50000	-0.020000
75%	1987.25000	0.320000
max	2023.00000	3.250000

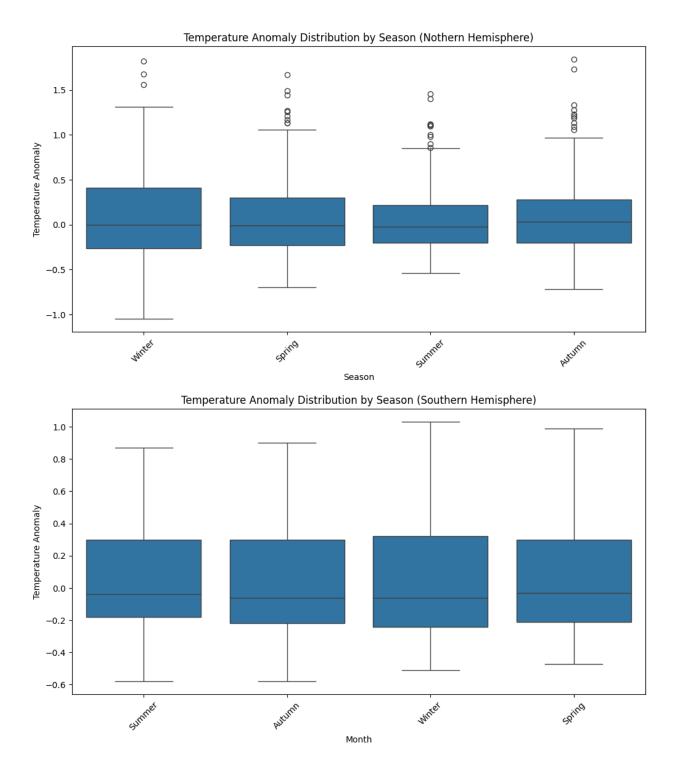
Figure 3: Descriptive Stats - Zonal Data

Using these statistics, we can interpret the time frame of these datasets using the min and max years as well as the variation of the anomalies based on their range, mean and standard deviation, allowing a better understanding of the variables and the datasets.

Moreover, box plots can also be used to gain insight into the distribution of variables, especially by splitting these variables into their categories and observing the distribution.



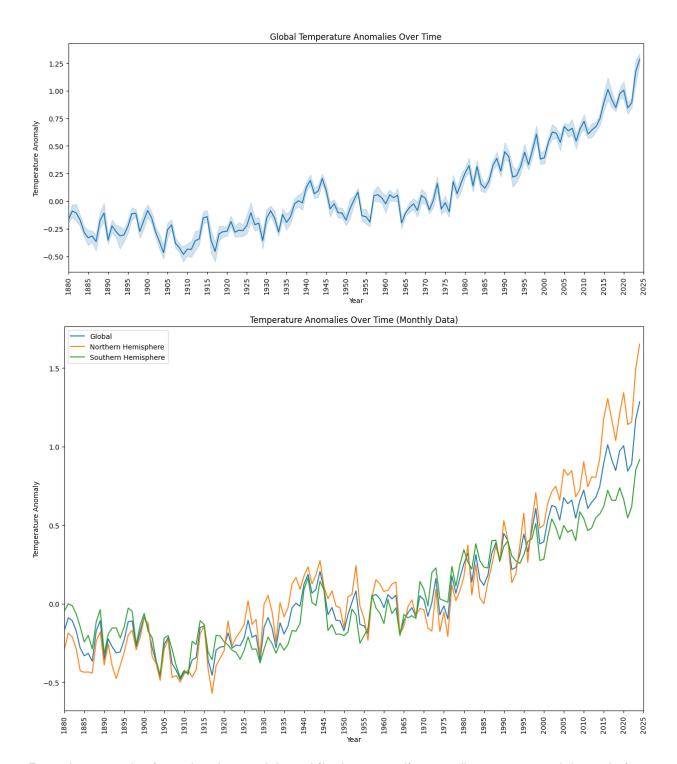




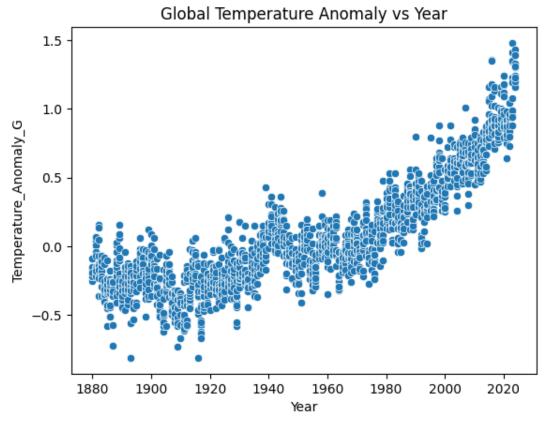
Multivariate Analysis

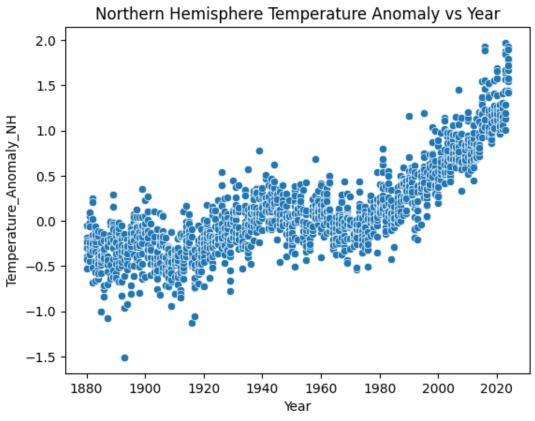
This is an analysis of exploring the relationship between variables, and several techniques have been used for this, including the usage of scatter plots, time series and correlation matrices.

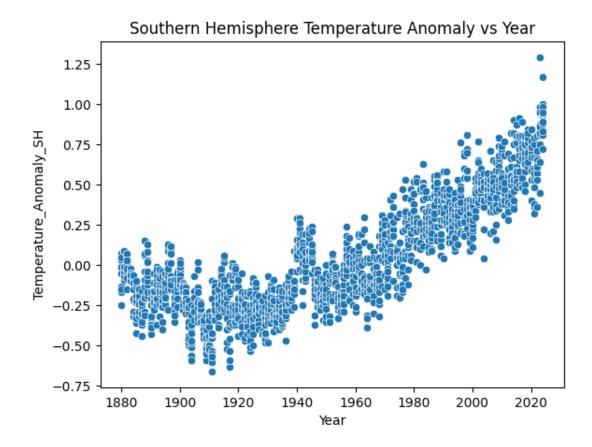
Firstly, a time series analysis of the temperature anomalies would allow us to observe any long-term trends.



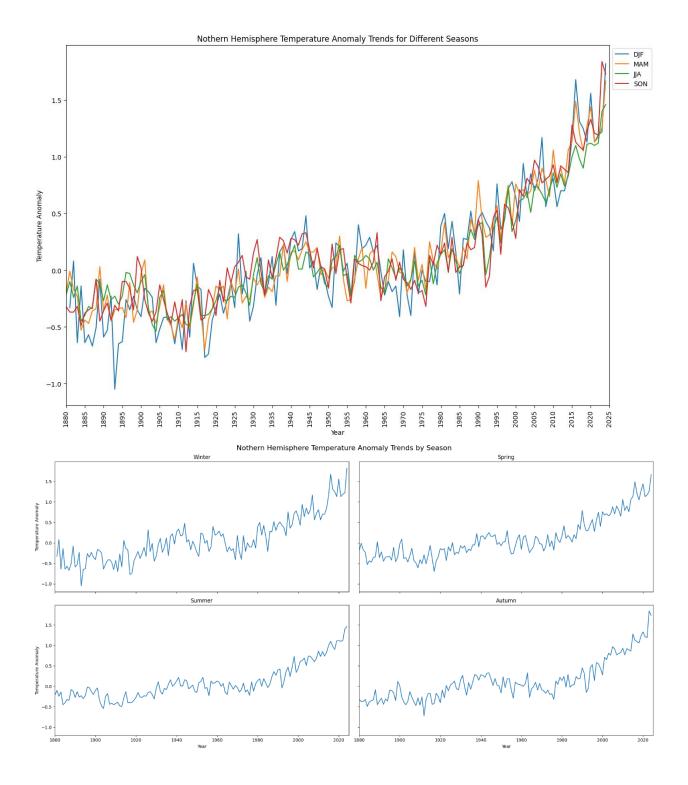
From these graphs, it can be observed that while the anomalies may fluctuate up and down during the short term, over the long term, they are rising at an increasing rate the closer they are to the present day. This clearly shows why global warming has become such a concerning issue lately since the increase in surface temperature has started to increase almost exponentially in recent years. This point is further proven by the following scatter plots.

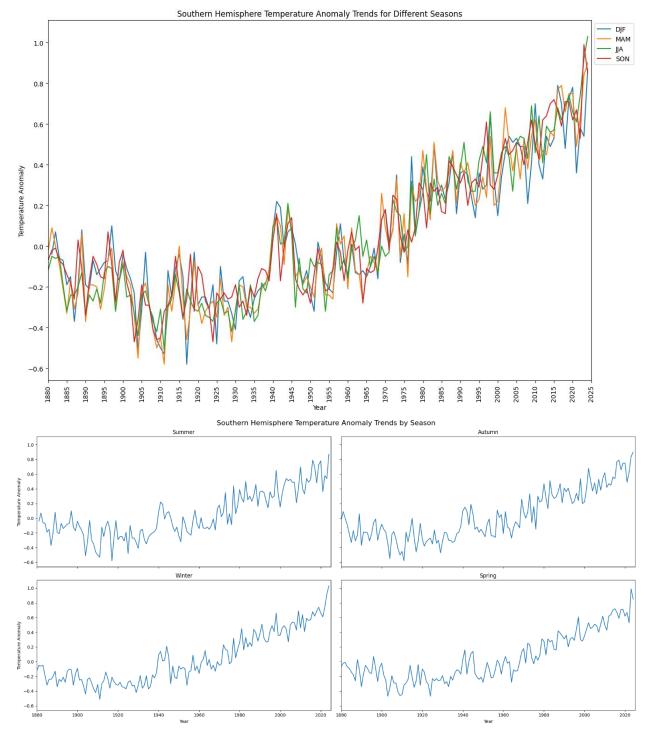




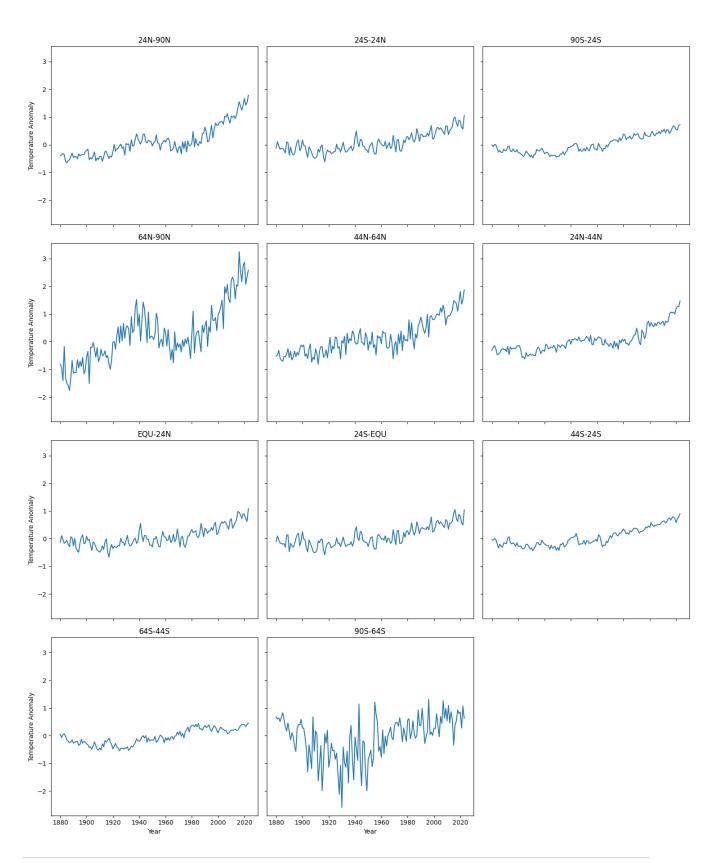


Furthermore, the effect of other variables, such as season and zone, on the temperature anomalies can be detected by using line graphs with different legends or subplots.



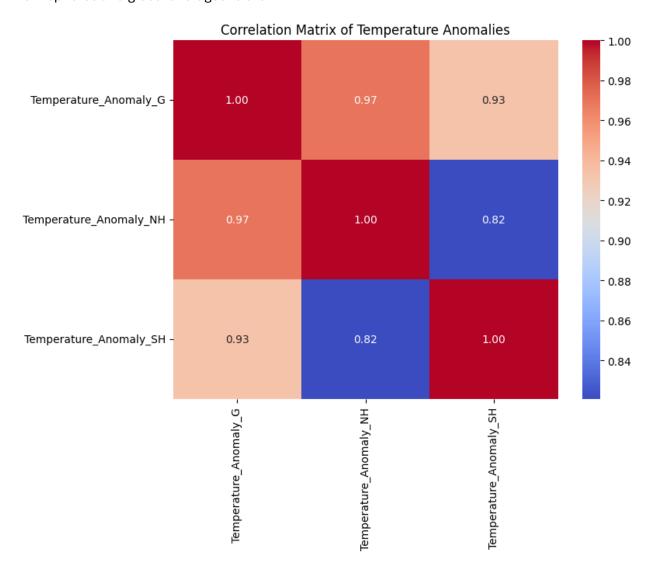


The above graphs show how winter is the most affected season as it shows higher anomalies towards 2024 than the other seasons.



Meanwhile, this graph shows how the highest anomalies are from the northern polar regions (64N-90N). This supports how the northern hemisphere showed higher temperature anomalies in the first time series analysis. The southern polar regions, on the other hand, show the most variation in these anomalies.

A correlation matrix combined with a heatmap can show how the anomalies of the two hemispheres and global averages relate.



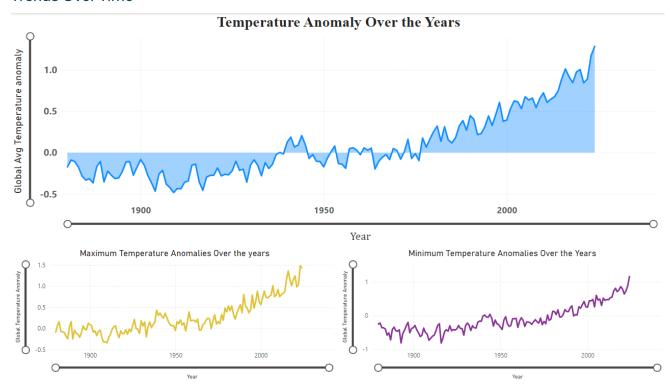
As expected they all have high correlation since the heating of one part of the globe effects the rest of it as well.

Data Storytelling

Introduction

Global warming has been a hot topic recently, but have you stopped to think about why it has gained so much attention in recent years? Just how much warming are we talking about? This analysis is based on answering the research question: How have Earth's surface temperature anomalies varied across different geographical zones, and what patterns can be observed in the monthly and seasonal variations of the surface temperature anomalies? For this, temperature anomaly data calculated based on 1951 to 1980 base averages have been acquired from NASA's Goddard Institute for Space Studies.

Trends Over Time

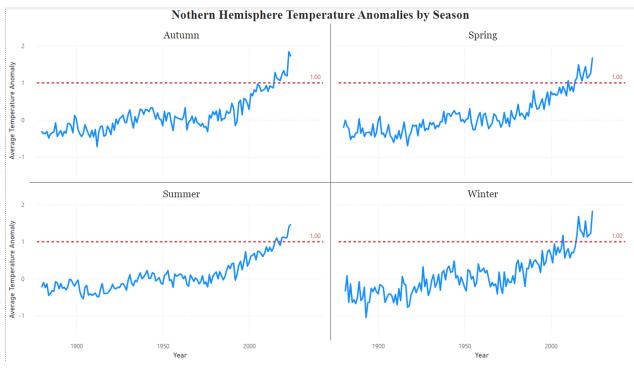


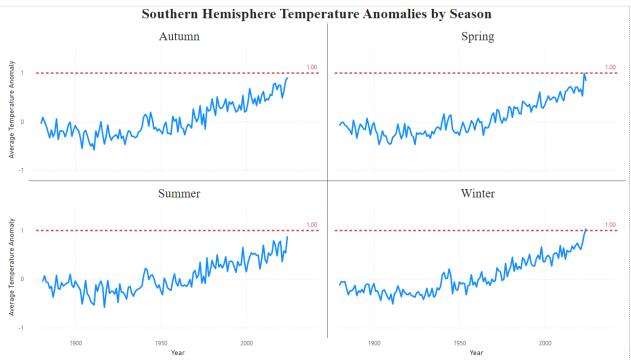
At a glance, by looking at this global average, maximum and minimum temperature anomalies over the years, it is clearly visible how, while there are short-term fluctuations, over the long-term following up to and after the base years, the temperature anomalies keep rising, getting warmer and warmer and in the most recent year rising so sharply that it's almost exponential growth. This exponential growth, along with the already alarmingly high anomalies, with the maximum anomalies reaching 1.5 degrees, is what has scientists and environmentalists most worried.

This is because a one-degree increase in global averages requires massive amounts of heat energy to be absorbed and, as such, can have a devastating impact on several different fronts. These could include heat waves, extreme precipitation, impact on biodiversity and ecosystems, droughts, food shortages, rising sea levels, and many more (NASA Science Editorial Team, 2019).

Seasonal Patterns

Having looked at an overview of the temperature anomalies across the years, we can also investigate how these variations have occurred over the different seasons. Have the summers been more affected? Or have winters become more warmer?

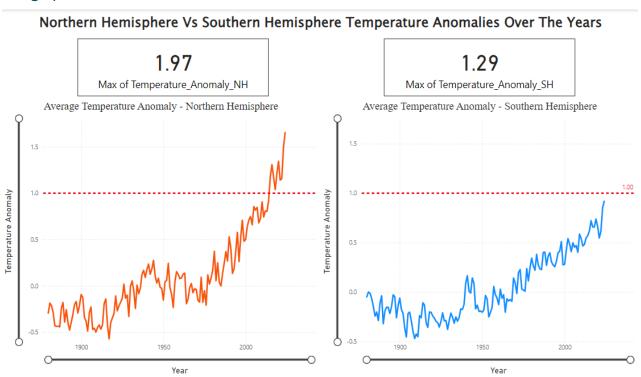




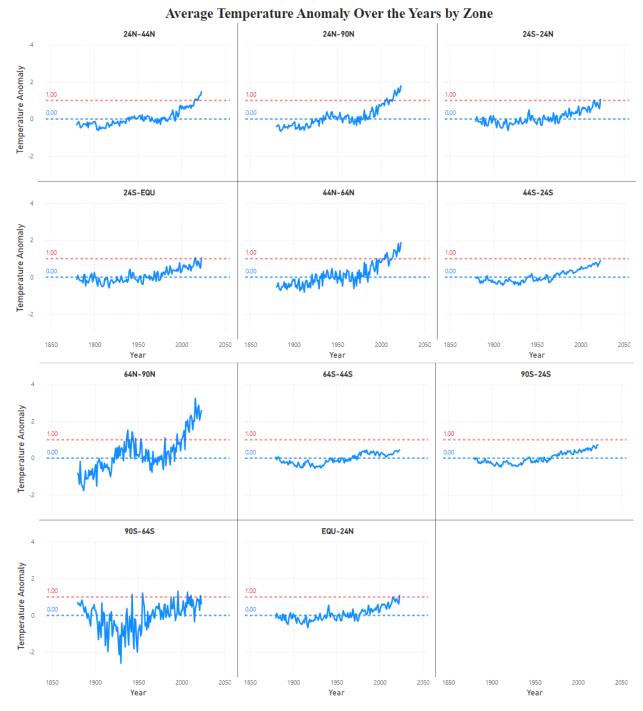
The above infographic depicts the temperature anomalies over the years for the four seasons of the northern and southern hemispheres, respectively, as each hemisphere faces these seasons during different parts of the year. It can be seen from these graphs that winters are most affected out of the four seasons, especially in the northern hemisphere, as anomalies have risen well above the critical one-degree mark. Even in the southern hemisphere, the winter season shows higher anomalies. In comparison, summer is the least affected season as its anomalies are lower than the rest, possibly due to summer having the warmest temperature out of the seasons.

Another keen observation from these plots is how the northern hemisphere as a whole seems to have higher temperature anomalies than the southern hemisphere. This indicates a geographical pattern that could be further studied.

Geographical Patterns

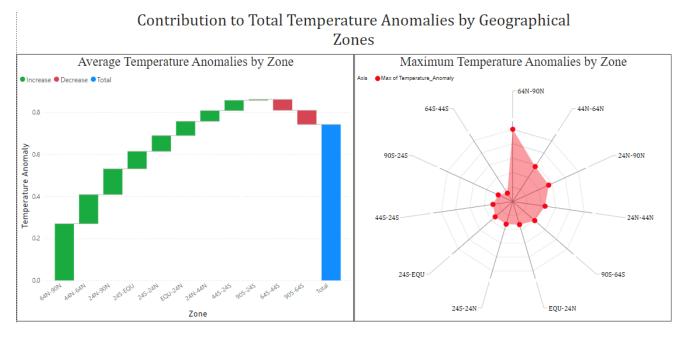


It can be clearly observed here how the northern hemisphere is more severely affected than the southern hemisphere; this means that more of those before-mentioned impacts can be observed within the northern hemisphere region. This does not, however, mean that the southern hemisphere is in the clear, as it can be observed that the temperature anomalies are sharply rising there as well. Furthermore, while higher anomalies are detected in the northern hemisphere, their impacts will be felt throughout the world, not just in the northern hemisphere, as sea levels rise and extreme dry weather bring about food shortages.



This allows closer observation of how temperature anomalies have occurred in smaller geographical zones, and it clearly shows that the latitude zones in the northern hemisphere show higher anomalies, with the highest anomalies from the northern polar region (64N-90N). On the other hand, the southern polar region (90S-64S) shows the most fluctuation in its anomalies. This coincides with the melting of the Arctic ice caps, which has become a significant concern since the northern polar region is heating up the fastest.

This can also be looked at from another perspective. How do these different zone's temperature anomalies contribute to the global average?



This infographic explores that question. It shows how each zone has contributed to the global average, thus telling us just how much each zone has been affected. The observation from these also further validates how the northern hemisphere has been more critically affected, as the highest of these contributions come from zones in the northern hemisphere.

Conclusion

Global warming is one of the biggest challenges we face today, and it's something that affects us all. Initiatives such as the Paris Agreement show that countries want to work together to tackle this issue, aiming to keep global temperature rise below 2°C compared to preindustrial levels (United Nations). Even with new policies and a push for renewable energy, we're still not meeting the targets set by this agreement, and the temperature shifts are becoming even more alarming, especially in the northern hemisphere.

This data story dives into surface temperature changes and highlights different patterns that vary by region and season, making it clear just how urgent this crisis is. Complicating matters, some political leaders, such as US President-Elect Donald Trump, choose to deny the existence of this crisis (Cheung, 2020) (Noor, 2024). It's crucial that awareness is raised and people are motivated to take collective action to tackle global warming and its impacts head-on. By understanding these patterns, we can foster informed decision-making and advocacy, ultimately working towards a sustainable future for ourselves and generations to come.

References

- Cheung, H. (2020), What does Trump actually believe on climate change? BBC News, 24
 January, Available at: https://www.bbc.com/news/world-us-canada-51213003 [Accessed: 05/01/2025]
- NASA Science Editorial Team (2019), A Degree of Concern: Why Global Temperatures
 Matter, NASA, Available at: https://science.nasa.gov/earth/climate-change/vital-signs/a-degree-of-concern-why-global-temperatures-matter/#hds-sidebar-nav-15 [Accessed: 03/01/2025]
- Noor, D. (2024), Trump continues to deny climate crisis as he visits hurricane-ravaged
 Georgia: Ex-president refers to climate crisis as 'one of the great scams' and plans to
 attend two fundraisers in oil-rich Texas, The Guardian, 1 October, Available at:
 https://www.theguardian.com/us-news/2024/oct/01/trump-visits-georgia-denies-climate-crisis-after-hurricane-helene [Accessed: 05/01/2025]
- United Nations, The Paris Agreement, un.org, Available at: https://www.un.org/en/climatechange/paris-agreement [Accessed: 05/01/2025]