Title: Analyzing the Pattern of Global Temperature Change

2025-12-05

Authors: Claire Nguyen, Katherine Mederos, Meghna Joshi

Abstract:

The main question we aim to answer in our project is: Do trends in global temperature over time on Earth show the occurrence of global warming? We collected data from 1850-2025 on the global land and ocean temperature from that time period. In analyzing such a large time frame, it was possible to observe trends in history. Based on the data collected and after we have cleaned it, we observed that there has been a consistent increase in global temperature as the years have passed. This was predicted before we collected the data, the data simply confirms our hypothesis to be true. After we have saved the dataset, cleaned it, and organized it, we performed an analysis by generating visualizations of the data to better make conclusions about our data. After conducting all of these steps, it is evident that our data suggest that global warming truly exists, due to the observable upwards trend in global temperature as the years have progressed.

Introduction:

The purpose of our analysis is to provide an answer to whether or not global warming exists on our planet. There are people that do not believe in the existence of this phenomenon, therefore our analysis exists to hopefully provide scientific evidence that global warming exists (given that increasing global temperature trends over the years exist). We are specifically aiming to test the hypothesis that global warming is in fact occurring and has been for years. Global warming is the rise in Earth's temperature over a long period of time due to human interactions and natural processes. This is a very controversial topic that has sparked debate over the years. Though a majority of the general population believe that global warming exists, there is a population of people who do not. With our analysis, we can successfully disprove those who do not believe in its existence by presenting concrete, scientifically proven, statistical evidence. By using tools like RStudio to develop our statistics skills, we can also face real-world problems that are seen as controversial!

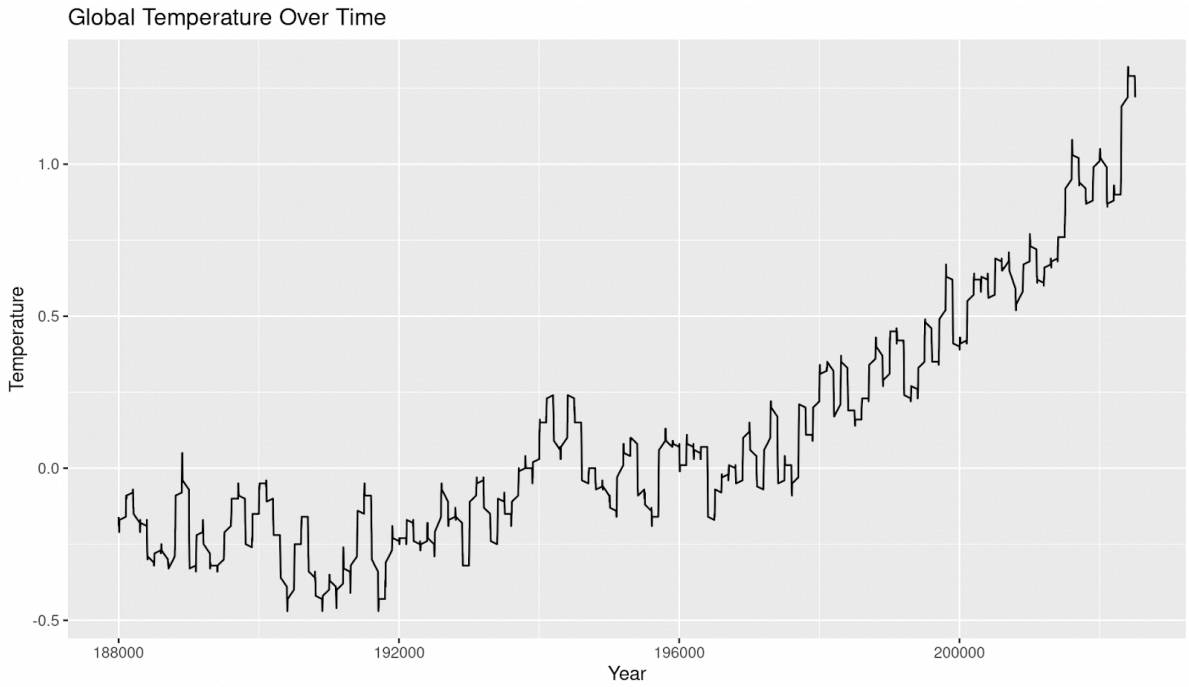
Data:

The data we are using for the analysis comes from the National Centers for Environmental Information (NCEI). This organization has tracked historical climate information for many years, and is a federal organization. The dataset obtained from their website is therefore credible, after researching the background behind the organization. The data we collected from the website includes global land and ocean temperatures from 1850-2025, using a time scale of 5 years. We used a time scale of five years in order to efficiently show long term patterns between a longer time frame, while still keeping the time intervals lower so that the trend may be easily observed. To clean the data, we imported the csv file containing the data and using the summary and head functions. We identified missing values using the is.na function and as.numeric() to correct it. There were some values in the dataset that were read as character strings, not numeric values. This was likely caused by issues with formatting on the dataset.

Data Cleaning:

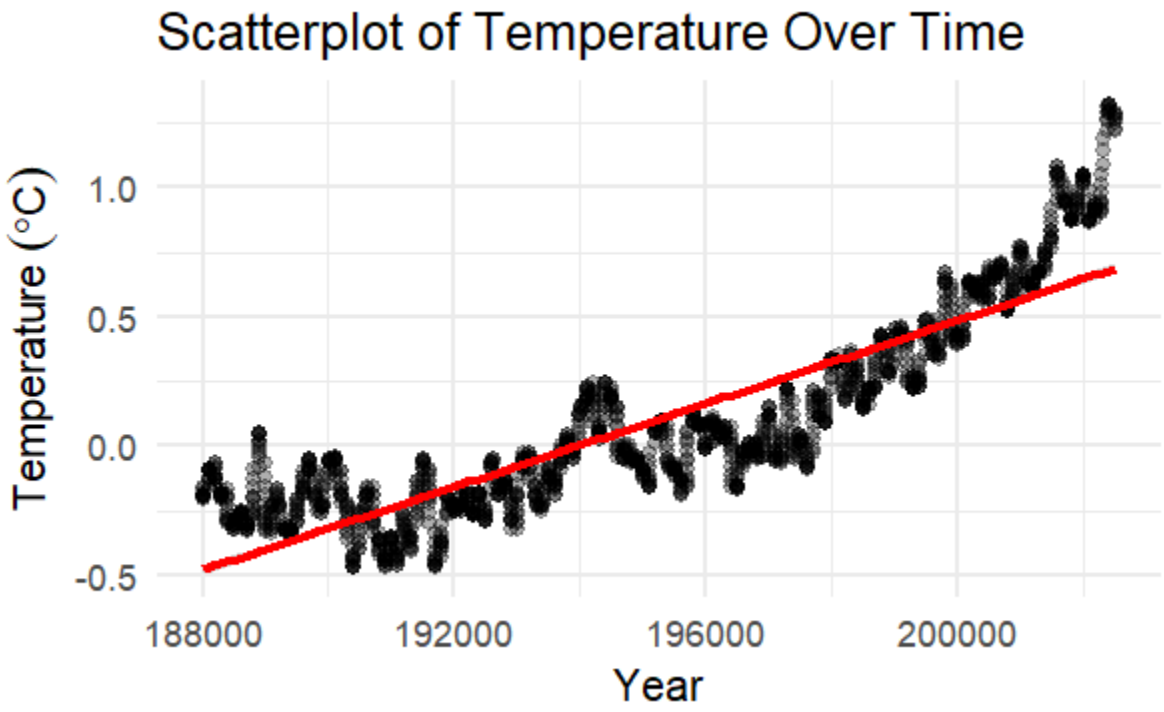
After noticing that there were some values that were incorrectly formatted (they were character strings when read, not numeric values), the function as.numeric() was used to convert the temperature format into numeric values. We removed all entries of data that were missing values/were corrupted. If values could not be corrected, they were also removed. After removing all of the unwanted values from the dataset, we were ready to proceed. The dataset now had consistent values that would result in a clean analysis on Rstudio. After looking through all the values and confirming that they were properly cleaned, we proceeded.

Visualization:

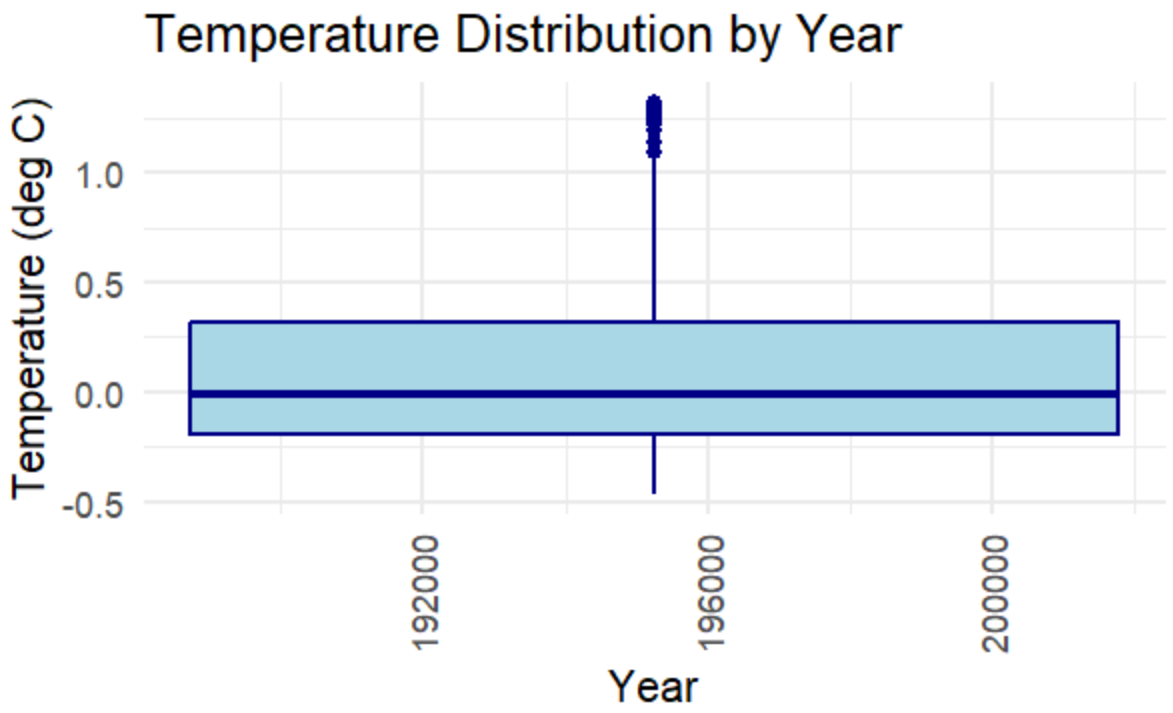


The visualization shows that global temperatures have risen steadily over time, with a sharp increase in recent decades. This supports our main question: whether trends in global temperature over time show evidence of global warming. After cleaning up the data, we observe a steady and consistent increase in the global temperature as the years progress, aligning with our predictions beforehand. Overall, the data suggests that global warming is occurring, as the long-term temperature trend rises significantly rather than remaining stable or randomly fluctuating. In addition to these observations, the visualization gives us a better understanding of the broader context and how temperature has changed over time. When examining the earlier years on the graph, we can see mild fluctuation. However, these fluctuations are normally expected due to climate variability, but they do not overshadow the clear upward trend that can be observed. This dramatic rise in temperature further emphasizes the impact of global warming, as discussed by climate scientists over the past few decades. Another important aspect of the visualization is that it allows us to identify patterns that cannot be easily observed in a raw

dataset. Therefore, by plotting the data, the trends become easier to interpret and helps us to identify the rise in temperature and intervals of stabilization. The consistent upward trend in the graph helps demonstrate that global warming is a persistent phenomenon. Furthermore, the visualization confirms our understanding of how human activity correlates with rising temperatures.



We also included a scatterplot to clearly display the line of best fit for our data. It is apparent that the temperature is experiencing a general increase throughout the years, with a noticeable steepness that starts approximately after the 1960s. This method of analyzing our data proved to be more clear and noticeable, thus being a more effective way of displaying the upward trend in global temperature.



Another method we used to analyze our data was by using a boxplot. Although it is a useful tool in statistical analysis, in our experiment, a boxplot is not the most effective method in analyzing our data. As shown above, the trends in our data are not observable. Therefore, the preferred methods of analysis for our experiment were the scatterplots. Though we wanted to try out different ways of displaying our data, we discovered that this particular method was not useful, which is okay. Ultimately, we deduced that the scatterplots were the best way to approach our analysis, with the one including the line of best fit as the most useful.

Analysis:

For our analysis, we planned to examine whether global temperature increases over time. To do so, we took a look at how temperature values change over time and used a line graph to help identify an upward trend. This visualization allows us to determine whether temperatures rise, fall, or stay constant as time progresses. Based on the shape of the plotted line, we expected to see a clear increase in global temperature over the years. We observed a steady, long-term upward trend of temperature. In more recent years, it was observed that the slope has gotten steeper, indicating the rate in which global temperature is increasing is getting higher. This means that the effects of human activity on Earth is making global warming more severe. This is a cause for concern, the presence of global warming is concerning in itself, but noticing that the

rate has increased as of late means that the issue is only getting worse. The visualization attached to the report earlier indicates an increase in the rate of global warming around 1960's period. This would make a lot of sense with the increasing population and overall human activity that only continues to increase. A line of best fit would indicate an upward trend to reflect increasing temperatures, with a strong linear positive correlation.

Conclusion:

After conducting our analysis on the global temperature data from NCEI, we have concluded that global temperatures have indeed increased as the years have progressed (from about a century and a half ago). This is evident in our visualizations of our data that show a clear upward trend on the graph, with a noticeable increase in the rate of global warming in recent years. We had predicted that this trend would be observed before conducting our analysis, which simply confirms our suspicions to be correct. When presenting our findings to the general population, those who have doubted the existence of global warming can quite clearly notice the trend of increase of temperature that is in front of them. It is also important to bring to the audience's attention that since the rate of global warming is actively increasing, this demonstrates a bigger issue that needs to be addressed. With the current observable trends, this issue is only going to continue getting worse based on historic data and action needs to be taken to slow the effects of global warming.