# **Group 6: Global Temperature Analysis Report**

ALY 6010: Probability Theory and Introductory Statistics

CRN: 21059  
Professor Dr. Yvonne Leung

Due Date: 24th February, 2025

Team: Jhansi Devi Telu, Jinjie Hu and Kaushal Nilesh Parekh

**Global Temperature Analysis Report**

**Introduction**

This report examines global land temperature trends over the past 50 years using data from the *Climate Change: Earth Surface Temperature Data (2017)* dataset. The dataset includes 645,675 records with attributes such as date, average temperature, temperature uncertainty, state, and country. Using R, exploratory data analysis (EDA) was conducted to summarize and visualize temperature distributions through statistical summaries, frequency tables, and visualizations such as histograms, bar charts, and box plots. The study leverages ggplot2, dplyr, and gmodels to identify trends and regional temperature variations, contributing to the understanding of climate change impacts.

### **Data Preparation**

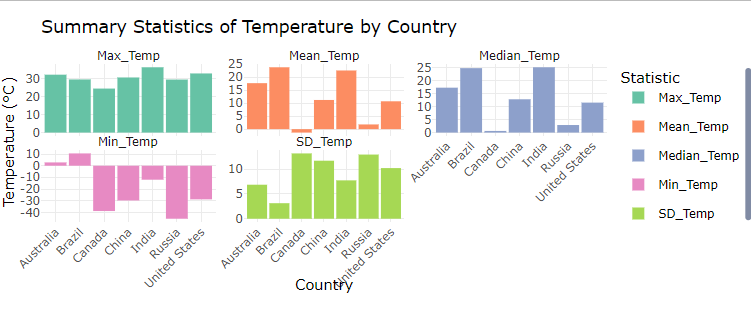
### The dataset was imported into R using the read.csv() function, followed by an initial exploration using head(), str(), and summary() to understand its structure and identify missing values. Data cleaning and transformation involved renaming columns for clarity, converting the Date column to Date format, and ensuring Avg\_Temp was numeric for analysis. Additionally, State and Country were converted to factors, missing values in Avg\_Temp were dropped, and the dataset was filtered to include only the last 50 years based on the most recent available year. Finally, the Year variable was extracted from the Date column for trend analysis.

### **Initial Analysis**

#### **Summary Statistics**

A summary statistics table was created for each country, including:

* Mean, Median, Standard Deviation (SD), Minimum, and Maximum Temperature.

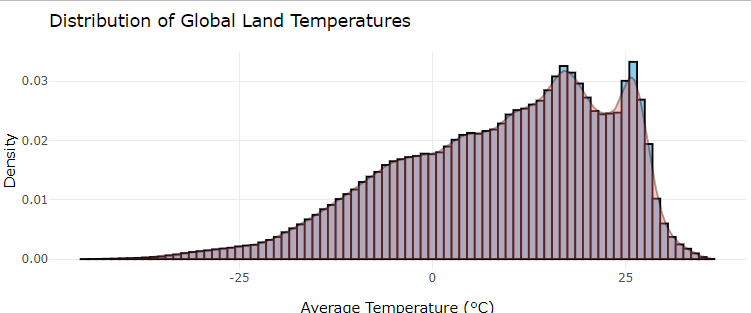


*Figure 1: Faceted Grouped Bar Chart of Summary Statistics*

* **Mean and median** highlight overall temperature trends, showing general warming patterns across regions.
* **Standard deviation** measures temperature variability, with higher values indicating greater fluctuations over time.
* **Minimum and maximum** values highlight extreme temperatures, showing the coldest and hottest recorded values.
* Countries with higher variability may experience stronger seasonal or climate shifts, while those with lower variability indicate more stable climate conditions.

**Distribution of Global Land Temperatures**

A histogram was generated to visualize the distribution of average temperatures across all regions.



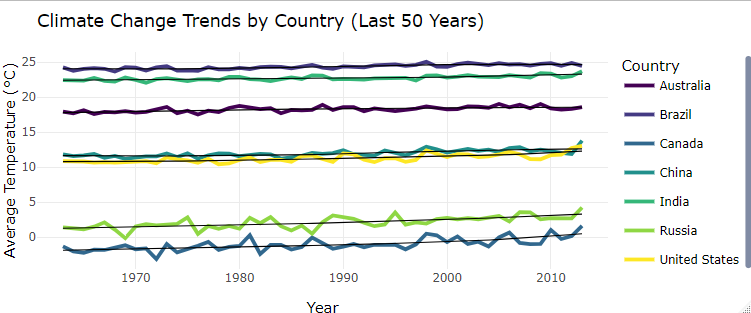
*Figure 2: Histogram of Average Temperatures of all regions*

* Purpose: This histogram with a density overlay illustrates the distribution of average temperatures globally, providing insights into common temperature ranges and extremities.
* Interpretation:

The distribution peaks between 10°C and 20°C, indicating these are the most common average temperatures worldwide.The presence of a long left tail shows that extremely cold temperatures are less common but significant, particularly highlighting regions that experience sub-zero averages.The density plot helps in visualizing the probability density of temperatures, confirming the skew towards more temperate conditions globally.

#### **Climate Change Trends**

A line plot was created to visualize temperature trends over the last 50 years for each country. The findings indicate an upward trend in average temperatures across most countries, suggesting climate warming.



*Figure 3: Line Plot of Climate Change trends by country*

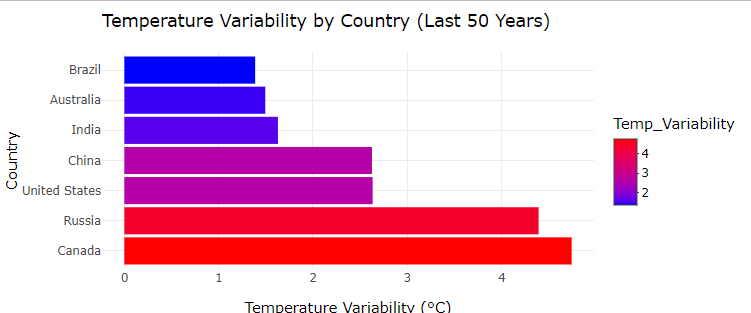
#### Purpose: This line graph tracks the average temperature changes over the last 50 years for each country, illustrating trends in global warming.

#### Interpretation:

#### A general upward trend in temperatures across all countries, with variations in the rate of increase, suggesting differential impacts of global warming.Notably sharp increases in countries like India and Brazil emphasize regions experiencing faster rises in temperature.The smooth lines help to identify long-term trends despite yearly fluctuations, supporting analyses of climate change progression.

#### **Temperature Variability by Country**

A bar chart was generated to illustrate temperature variability (difference between maximum and minimum recorded temperatures) by country. The results show significant variation in temperature trends across different regions.



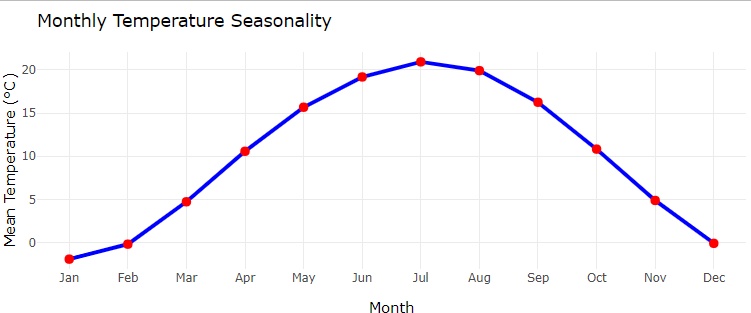
*Figure 4:Bar chart of Temperature Variability by Country Over Last 50 years*

* Purpose: This bar chart measures the range between the highest and lowest recorded temperatures in each country over the last 50 years, offering insights into climatic stability.
* Interpretation:

Countries like Brazil and Australia show higher variability, indicating a significant difference between summer and winter temperatures.Lower variability in countries like Canada suggests a narrower range of temperature fluctuations, which may indicate less extreme seasonal changes.The chart effectively underscores how geographic and environmental factors contribute to temperature variability.

#### **Seasonal Temperature Patterns**

A line graph was used to analyze temperature seasonality. The data revealed recurring seasonal fluctuations, with higher temperatures during mid-year months and lower temperatures during the end and beginning of the year.



*Figure 5: Line Plot of Seasonal Temperatures*

* Purpose: This line graph displays the mean temperatures for each month, revealing patterns of seasonal variation.
* Interpretation:

The graph shows a clear sinusoidal pattern, peaking in July and reaching a minimum in January, consistent with the expected behavior of temperatures in the Northern Hemisphere.The plotted points (red dots) on the peaks and troughs help emphasize the highest and lowest average temperatures of the year, facilitating a quick visual assessment of seasonal extremes.This visualization is crucial for understanding how temperatures shift throughout the year, aiding sectors like agriculture and energy in planning and optimization.

### **Conclusion**

This analysis confirms a rising trend in global temperatures over the past 50 years, with notable regional variations. The findings highlight increasing averages, significant temperature variability in some countries, and clear seasonal patterns. These insights emphasize the need for proactive climate policies and adaptation strategies to mitigate the effects of climate change.

**References**

*Climate change: Earth surface temperature data*. (2017, May 1). Kaggle. https://www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data

Kabacoff, R. I. (2022). *R in Action, Third Edition: Data analysis and graphics with R and Tidyverse*. Simon and Schuster.

Smart Student. (2024, September 29). *Format a student paper in APA 7th Edition using Microsoft Word* [Video]. YouTube. <https://www.youtube.com/watch?v=EE-IWUYcvVo>

OpenAI. (2025). ChatGPT (Feb 24 version) [Large language model]. <https://chat.openai.com>