EARTH DATA ANALYSIS

Date: 13 October 2024

Table of Contents

Introduction

In recent years, the significance of understanding and analyzing Earth data has become paramount due to the increasing impact of climate change, natural disasters, and resource consumption. This report summarizes the analysis of five critical datasets: precipitation data, temperatures, earthquake occurrences, CO2 emissions, and world energy consumption. These datasets provide insights into environmental trends, energy use, and climatic variations across different regions.

The analysis utilizes Python programming, leveraging libraries such as Pandas for data manipulation, Matplotlib and Seaborn for data visualization, thereby enabling comprehensive insights into the interactions between these variables and their implications for global sustainability.

Datasets Overview

This analysis utilizes five datasets, two of which were sourced from NASA Earth Data, while three were obtained from Kaggle.

1. **Precipitation Dataset**
   Columns:
   - Country Name
   - Country Code
   - Indicator Name
   - Indicator Code
   - Yearly data from 1961 to 2020 (1961, 1962, ..., 2020)
2. **Temperatures Dataset**
   Columns:
   - Region
   - Country
   - City
   - Month
   - Day
   - Year
   - AvgTemperature

3. **Earthquakes Dataset**
   Columns:
   - Time
   - Place
   - Status
   - Tsunami
   - Significance
   - Data Type
   - Magnitude
   - State
   - Longitude
   - Latitude
   - Depth
   - Year
   - Month
4. **CO2 Emissions Dataset**

   Columns:
   - Country
   - Code
   - Year
   - Annual Carbon Dioxide Emissions
5. **World Energy Consumption Dataset**
   Columns:
   - Country
   - Yearly data from 1965 to 2023 (1965, 1966, ..., 2023)

Methodology

Data Collection

The datasets were sourced from reputable organizations, including NASA Earth Data and Kaggle. Each dataset was selected for its relevance to understanding environmental changes and trends. The data were imported into Python using the Pandas library, facilitating efficient data handling and analysis.

Data Cleaning and Preprocessing

Before analysis, each dataset underwent a rigorous cleaning process to address missing values, correct data types, and remove duplicates. For example:

- In the precipitation dataset, years with NaN values were handled by either filling them with zeros or dropping those entries based on the analysis requirements.

Data Analysis Techniques

Various analytical techniques were employed across the datasets:

- **Descriptive Statistics**: Summary statistics were calculated to understand the general trends within each dataset.
- **Visualization**: Matplotlib and Seaborn were used to create engaging visual representations of the data, including:
  - Line charts to illustrate trends over time (e.g., average temperature changes).
  - Bar charts to compare annual precipitation by country.
  - Pie charts to represent the distribution of data.

Insights Extraction

The analysis aimed to extract meaningful insights, such as:

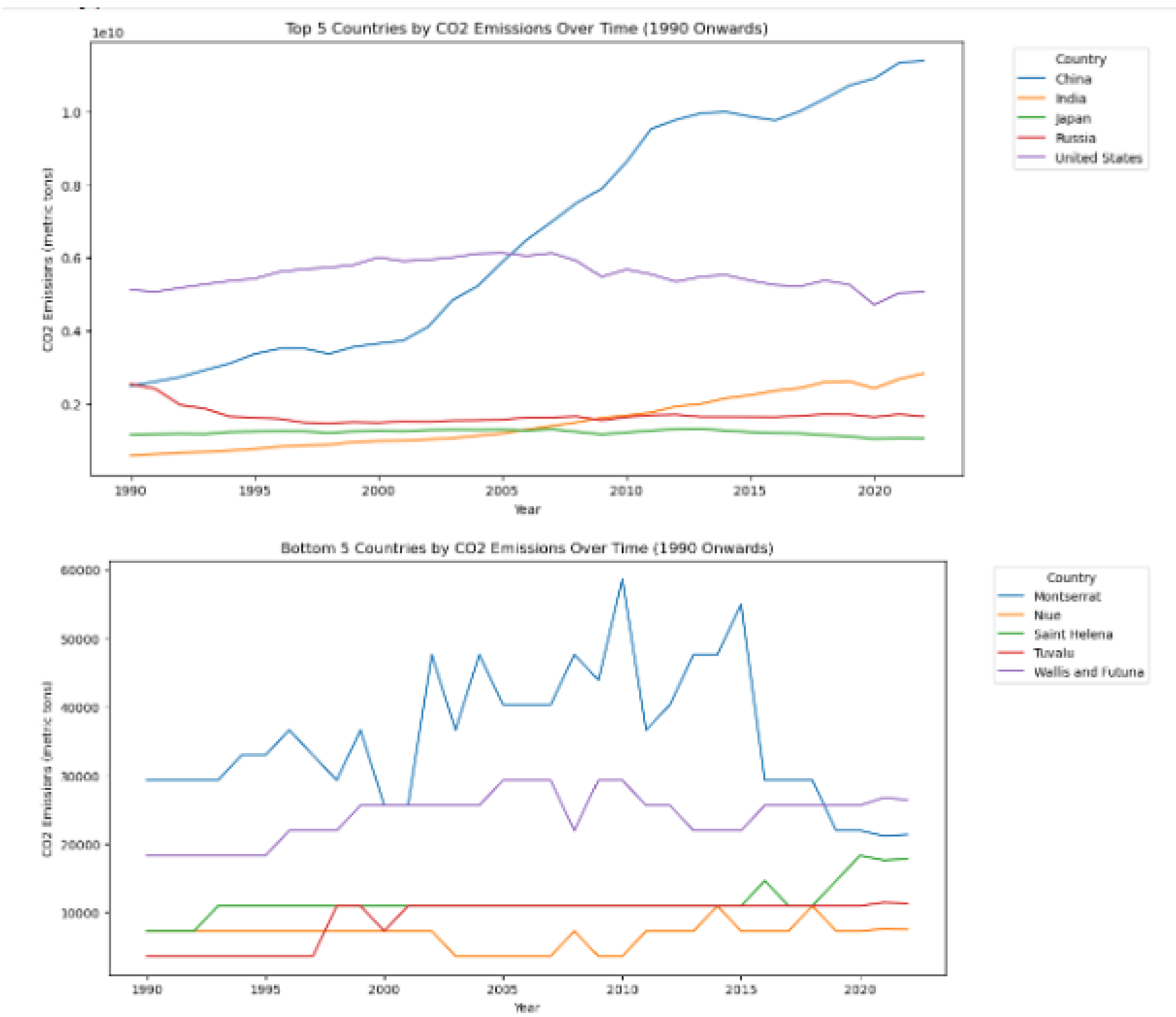
- Identifying countries with the highest and lowest precipitation levels.
- Analyzing temperature trends over decades.
- Evaluating the correlation between CO_2 emissions and temperature changes.
- Assessing earthquake frequency and intensity across various regions.

By synthesizing these analyses, the report aims to deliver a comprehensive overview of how these datasets contribute to a broader understanding of global sustainability challenges.

Key Findings

CO_2 Emissions Findings

1. Top 5 and Bottom 5 Countries by CO2 Emissions Over Time (1990-2020)



Top 5 Countries by CO2 Emissions Over Time (1990 Onwards)

The line chart illustrates the trends in CO2 emissions for the top five emitting countries: China, India, Japan, Russia, and the United States from 1990 to 2020.

- **China** has shown a significant increase in emissions, becoming the largest emitter by a considerable margin. This upward trend highlights China's rapid industrialization and economic growth over the past few decades.
- **India** also demonstrates a rising trend in emissions, reflecting its growing economy and energy demands.

- **The United States** and **Japan** show relatively stable emissions, with the U.S. emissions declining slightly over the years, likely due to energy efficiency improvements and a shift towards cleaner energy sources.
- **Russia's** emissions have fluctuated but have not increased substantially, indicating a stabilization in its industrial output.

This visualization underscores the urgent need for global cooperation in addressing CO2 emissions, particularly from the largest contributors.

Bottom 5 Countries by CO2 Emissions Over Time (1990 Onwards)

The second chart focuses on the bottom five countries by CO2 emissions, including Montenegro, Saint Helena, Tuvalu, and others.

- The emissions of these countries remain consistently low compared to the top emitters.
- **Montenegro** shows a slight increase over time, but overall, the emissions for these nations are negligible on a global scale.
- Countries like **Saint Helena** and **Tuvalu** maintain very low emissions, reflecting their smaller industrial bases and populations.

This visualization highlights the disparity in emissions between the largest and smallest contributors, emphasizing that while some countries are making minimal contributions to global emissions, the focus remains on major emitters for effective climate action.

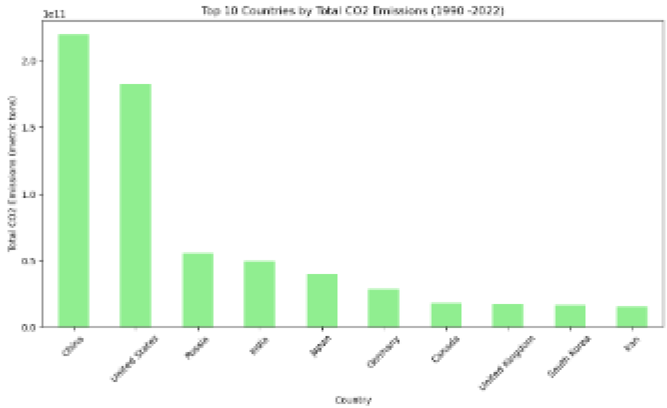
2. **Top 10 and Bottom 10 Countries by Total CO2 Emissions (1990-2022)**

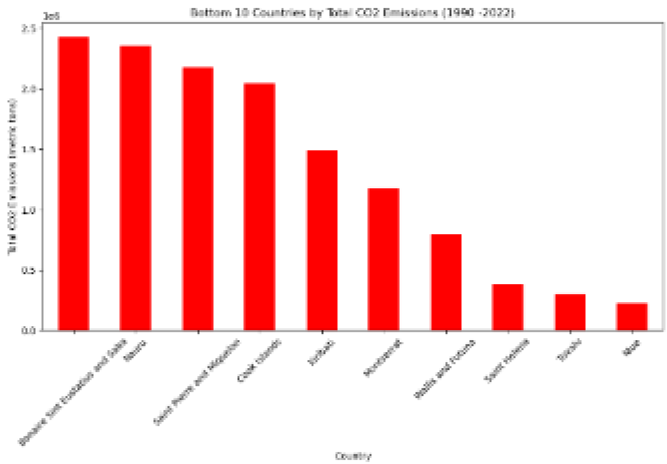
Top 10 Countries by Total CO2 Emissions (1990-2022)

This chart shows that **China** is the largest emitter, significantly outpacing the **United States**, which ranks second. Other contributors like **India** and **Russia** also have notable emissions, highlighting their industrial activities. The data underscores the concentration of emissions among a few major countries.

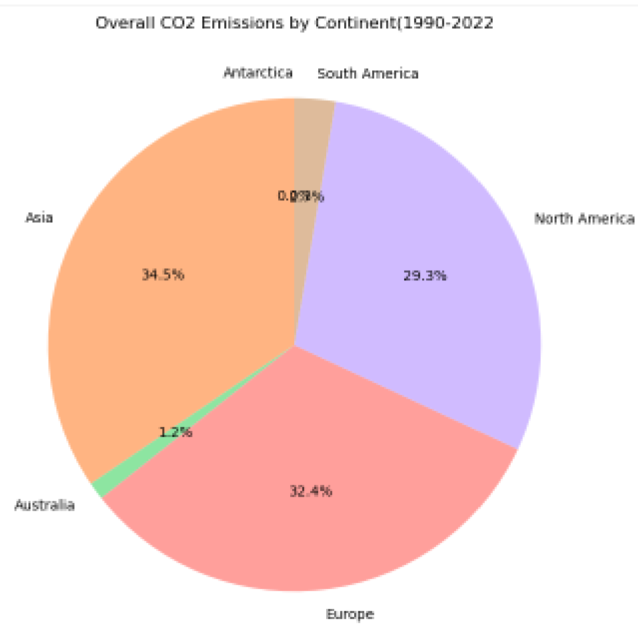
Bottom 10 Countries by Total CO2 Emissions (1990-2022)

In contrast, the bottom ten countries exhibit very low total emissions, with nations like the **Bahamas** and **Maldives** contributing minimally. This chart illustrates the stark difference in emissions between high and low contributors. It emphasizes that while some countries have negligible impact, major emitters drive the global emissions landscape.





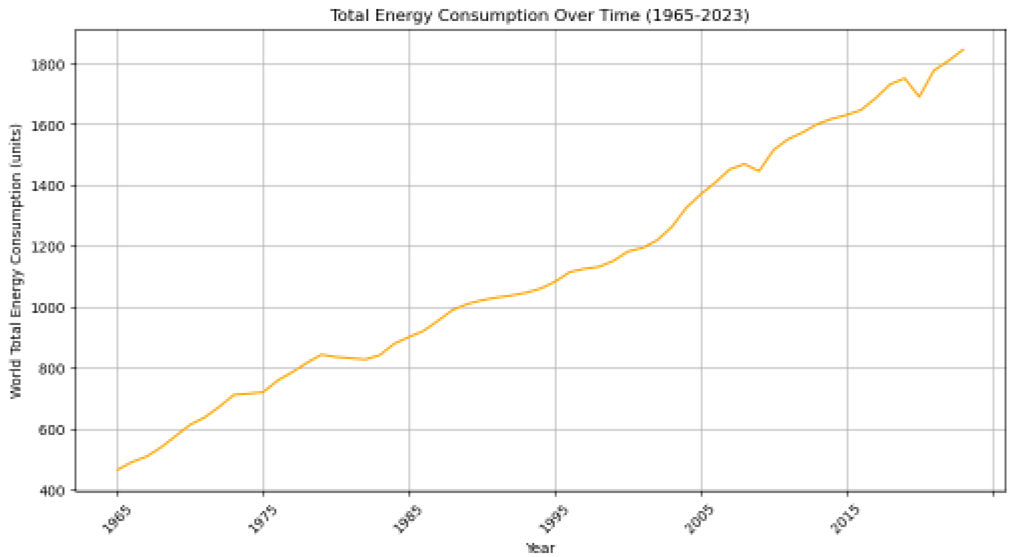
3. Overall CO2 Emissions by Continent (1990-2022)



The pie chart reveals that Asia is the largest contributor to CO2 emissions, accounting for 34.5% of the total. North America follows closely with 29.3%, while Europe contributes 32.4%. The minimal emissions from Australia and Antarctica highlight the significant disparity in emissions among continents, with a majority stemming from industrialized regions.

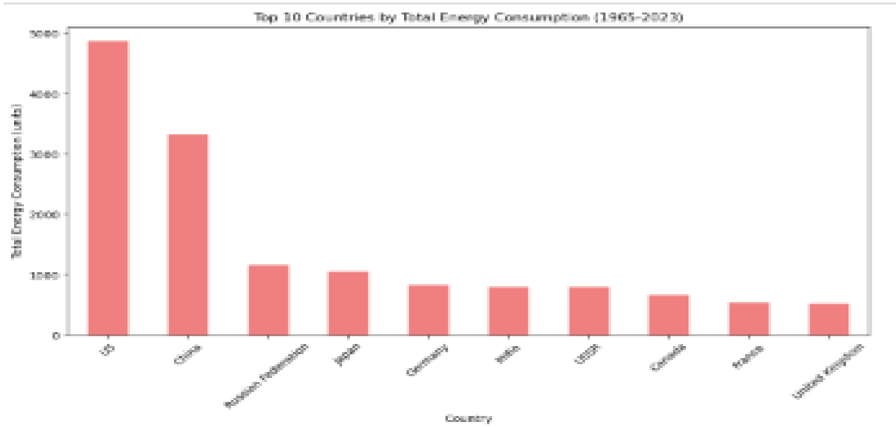
World Energy Consumption

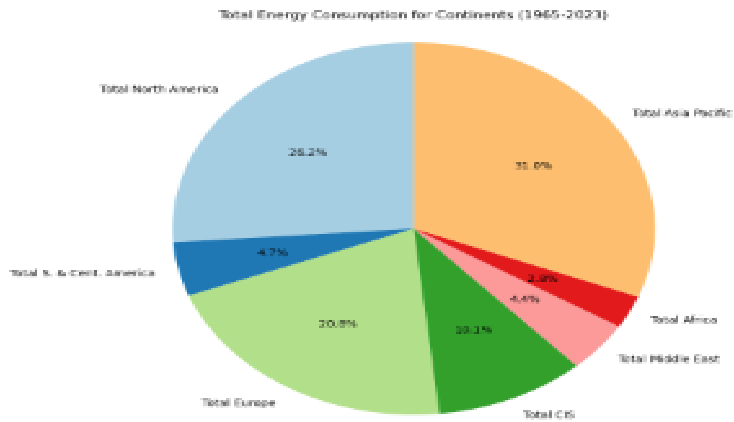
1. **Total Energy Consumption over Time (1965-2023)**



This line chart illustrates the steady increase in global energy consumption from 1965 to 2023. The trend shows a significant upward trajectory, indicating growing energy demands due to population growth and industrialization. Notable spikes may correlate with economic developments and technological advancements. Overall, the data highlights the persistent rise in energy consumption, emphasizing the need for sustainable energy solutions.

2. **Global Energy Consumption by Country and Continent (1965-2023)**
   The bar chart displays the top ten countries by total energy consumption, with the United States leading significantly. This stark contrast highlights the substantial energy demands of industrialized nations. The pie chart complements this by breaking down total energy consumption by continent, revealing that Asia-Pacific dominates with 31.8%, followed by North America at 25.2%. These visualizations underscore the disparities in energy consumption patterns and the need for equitable energy policies.





Precipitation Analysis

1. Top 5 and Bottom 5 Countries Total Precipitation

```
Top 5 countries with highest overall precipation (1961-2020)
              Country Name  Total Precipitation
219  Sao Tome and Principe             192080.0
189       Papua New Guinea             188520.0
209        Solomon Islands             181680.0
185                 Panama             175680.0
48              Costa Rica             175560.0

Bottom 5 countries by total precipitation:
             Country Name  Total Precipitation
67       Egypt, Arab Rep.               1086.0
132                 Libya               3360.0
205          Saudi Arabia               3540.0
200                 Qatar               4440.0
8    United Arab Emirates               4680.0
```

<u>Top 5 Countries by Total Precipitation</u>

1. **Sao Tome and Principe**: 192,080 mm
2. **Papua New Guinea**: 188,280 mm
3. **Solomon Islands**: 181,680 mm
4. **Panama**: 175,680 mm
5. **Costa Rica**: 175,560 mm

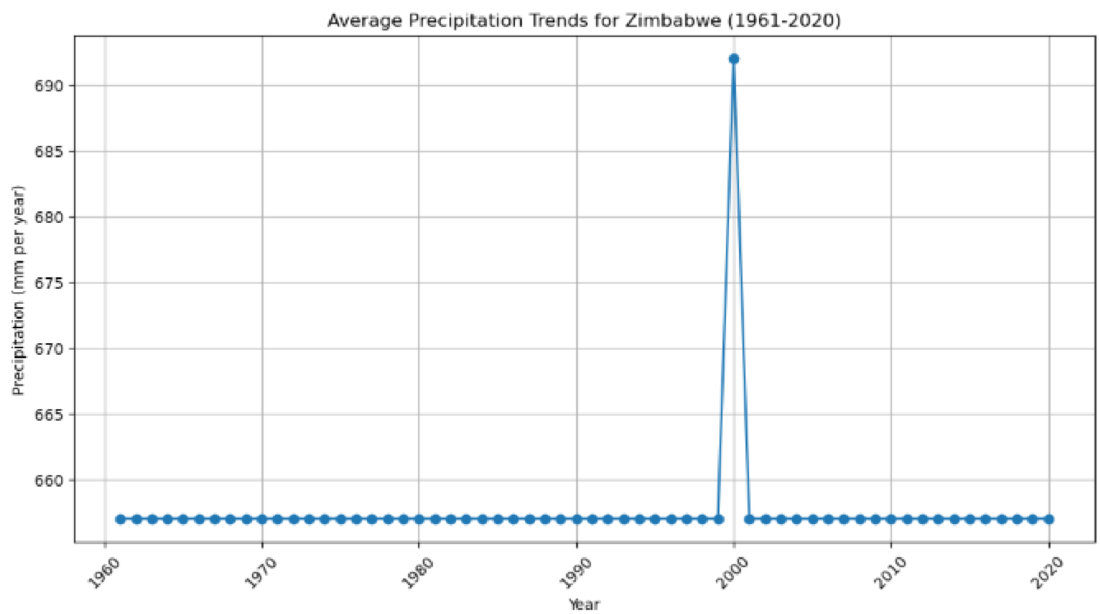
These countries exhibit significantly high levels of precipitation, likely due to their tropical climates and geographical features.

Bottom 5 Countries by Total Precipitation

1. **Egypt, Arab Republic**: 1,806 mm
2. **Libya**: 2,173 mm
3. **Saudi Arabia**: 3,604 mm
4. **Qatar**: 4,200 mm
5. **United Arab Emirates**: 4,680 mm

In contrast, these countries experience much lower precipitation, predominantly due to arid and semi-arid climates.

2. **Average Precipitation Trends for Zimbabwe**



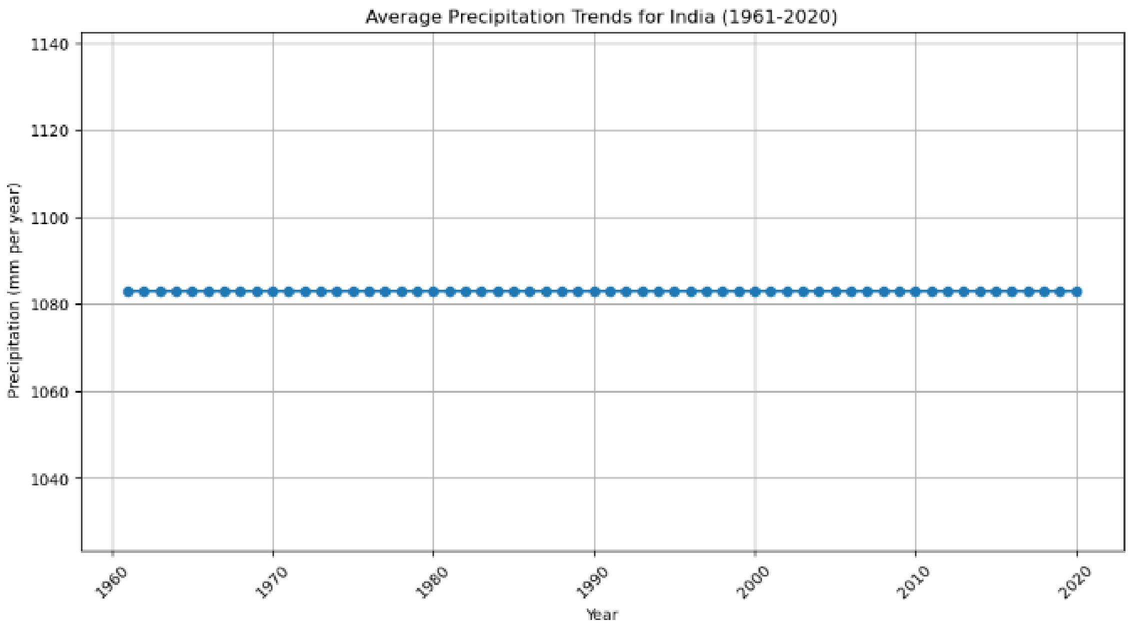
The chart illustrates the average annual precipitation in Zimbabwe from 1961 to 2020. Key observations include:

- **Overall Stability**: Most years show relatively consistent precipitation levels, fluctuating around 660-670 mm per year.
- **Significant Spike**: There is a notable spike in precipitation around the year **2000**, which stands out against the otherwise stable trend.

Implications

The spike might indicate unusual weather patterns or significant climate events during that period, which could have impacted agriculture, water resources, and overall ecological health in Zimbabwe. Further analysis would be necessary to understand the causes of this anomaly and its effects.

2. **Average Precipitation Trends for India (1961-2020)**



The chart displays the average annual precipitation in India over the period from 1961 to 2020. Key observations include:

- **Consistent Levels**: Average precipitation remains stable around **1080 mm per year** throughout the entire period.
- **Minimal Variation**: There are no significant fluctuations or trends, indicating a relatively steady climate in terms of precipitation during these decades.

Implications
The stability in precipitation may suggest effective water management practices or consistent monsoon patterns that have maintained average rainfall levels. However, it is also essential to consider regional variations within India, as different areas may experience diverse climatic conditions. Further investigation into localized trends could provide a more comprehensive understanding of precipitation impacts across the country.

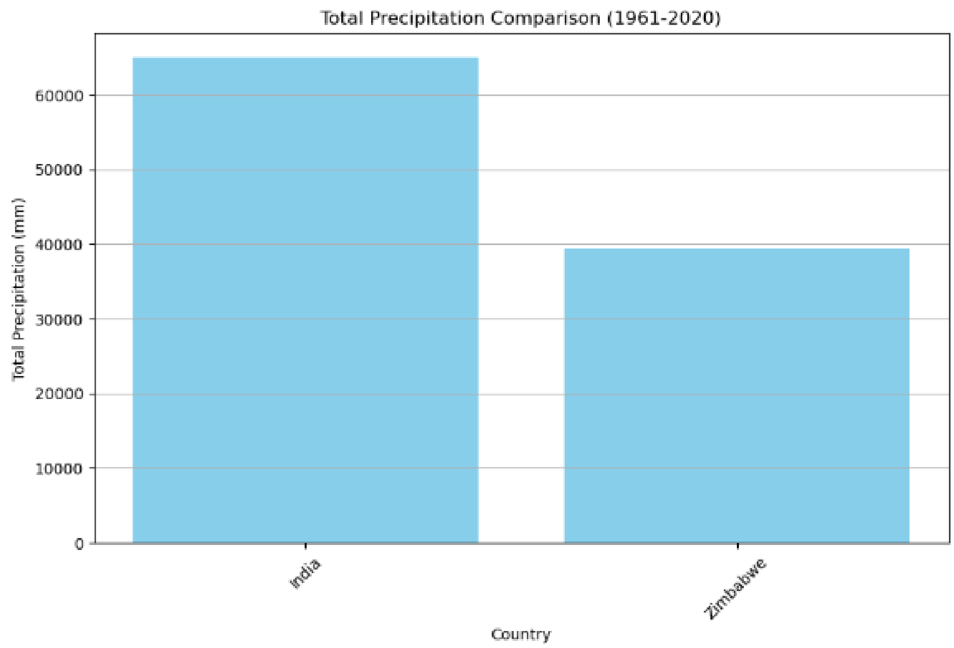
3. Total Precipitation Comparison (1961-2020)

The bar chart compares the total precipitation recorded in India and Zimbabwe from 1961 to 2020. Key observations include:

- **India**: The total precipitation is significantly higher, totaling around **60,000 mm** over the period.
- **Zimbabwe**: The total precipitation is markedly lower, at approximately **20,000 mm**.

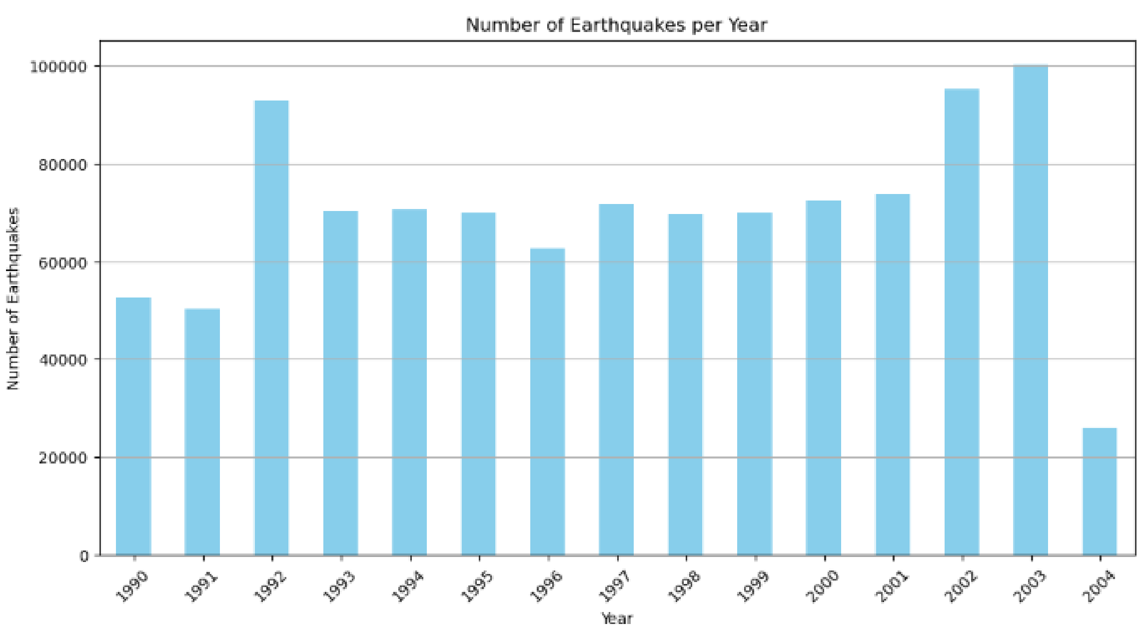
Implications
The substantial difference in total precipitation between the two countries could reflect various climatic, geographical, and seasonal factors. India, with its diverse monsoon system, typically receives more rainfall compared to Zimbabwe, which has a different

climate and seasonal patterns. This disparity can have significant implications for agriculture, water resources, and ecosystem health in both regions



World Earthquakes Analysis

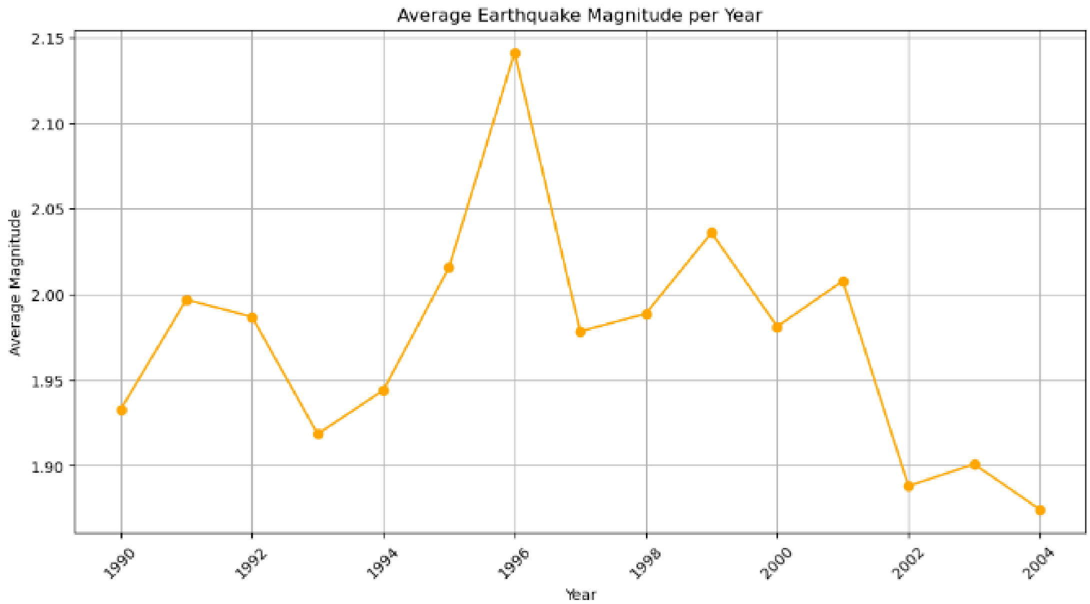
1. Annual Earthquake Frequency (1990-2004)



The bar chart illustrates the number of earthquakes recorded each year from 1990 to 2004. Key observations include:

- **General Trends**: The number of earthquakes fluctuates over the years, with certain peaks and troughs.
- **Peak Year**: The year **2003** shows the highest number of earthquakes, reaching around **100,000**.
- **Lower Years**: Some years, like **1992** and **1994**, exhibit significantly lower counts, around **20,000**.

2. **Average Earthquake Magnitude per Year (1990-2004)**

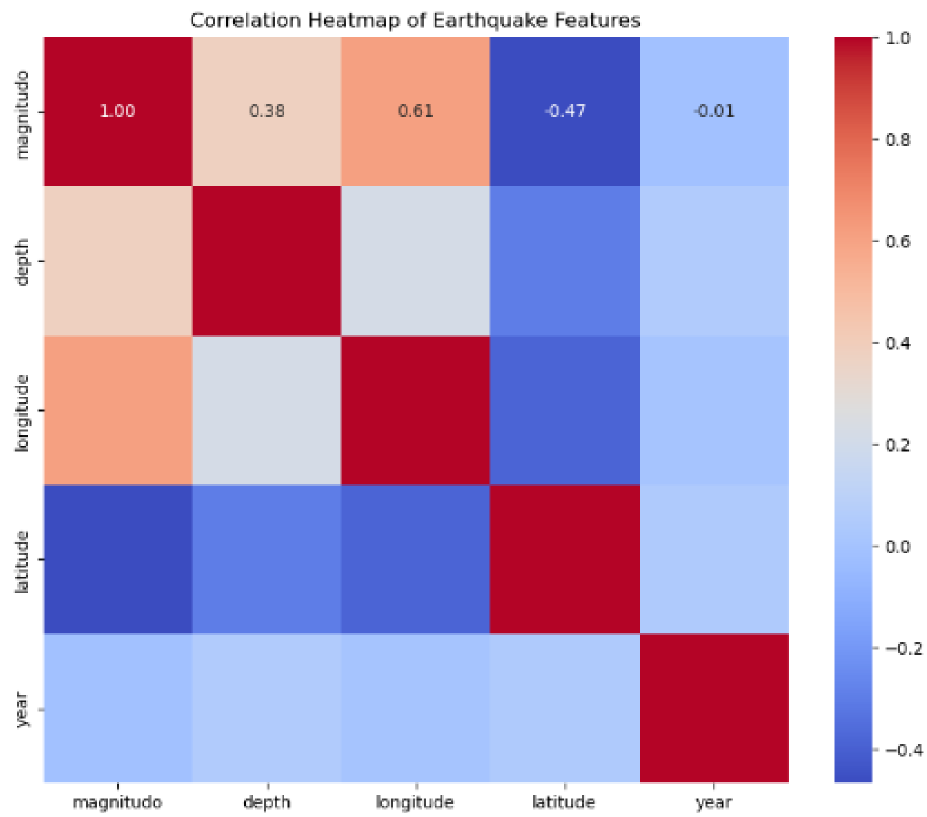


The chart displays the average magnitude of earthquakes recorded each year from 1990 to 2004. Key insights include:

- **Peak Magnitude**: The year **1996** shows a notable peak in average magnitude, reaching just above **2.15**.
- **Fluctuations**: The average magnitude exhibits fluctuations, with several years hovering around **1.90** to **2.00**.
- **Trend Analysis**: Despite the variations, there is no clear upward or downward trend over the entire period, suggesting that while earthquake frequency may vary, the average magnitude remains relatively stable.

These insights indicate that while the number of earthquakes may change significantly from year to year, the average strength of these earthquakes tends to be consistent. Understanding this relationship can aid in assessing risks and preparing for potential seismic events.

3. **Correlation Heatmap of Earthquake Features**

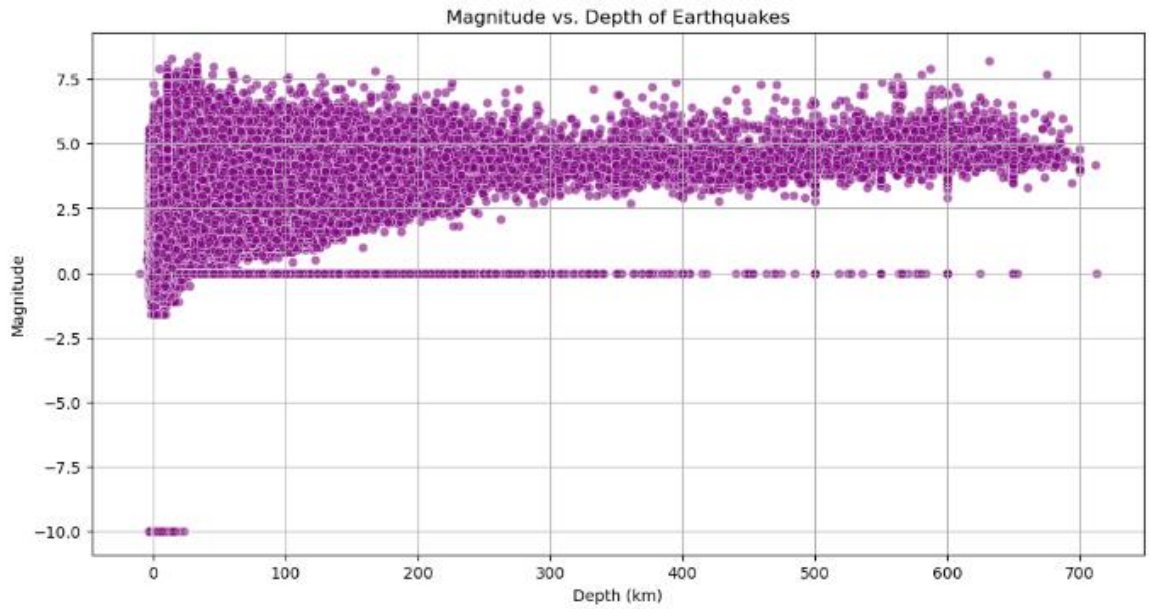


The heatmap illustrates the correlation between various features associated with earthquakes, including magnitude, depth, longitude, latitude, and year. Key insights include:

- **Strong Correlation**: There is a strong positive correlation (**0.61**) between **magnitude** and **depth**, indicating that deeper earthquakes may tend to have higher magnitudes.
- **Negative Correlation**: The correlation between **depth** and **latitude** is slightly negative (**-0.47**), suggesting that as depth increases, latitude may have a decreasing relationship, although this correlation is moderate.
- **Weak Relationships**: Other features, such as **year** and **longitude**, show weak correlations (near **0.00**), indicating little to no relationship between these variables.

These insights can help researchers and seismologists understand the relationships between different earthquake characteristics, which may inform predictive models and risk assessments in seismic research.

4. Magnitude vs. Depth of Earthquakes



The scatter plot illustrates the relationship between the depth (in kilometers) and magnitude of earthquakes. Key insights include:

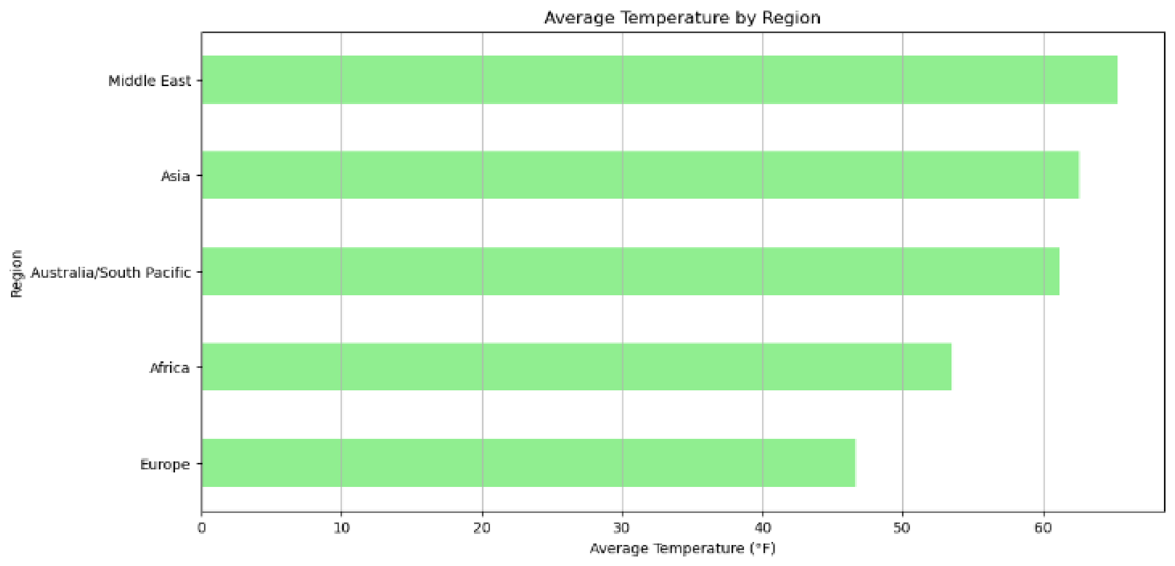
- **Inverse Trend**: There is a noticeable trend where deeper earthquakes tend to have lower magnitudes. As depth increases, the majority of magnitudes cluster below **3.0**, indicating that deeper events are often less powerful.
- **Concentration of Data Points**: Most data points are concentrated at shallower depths (0-100 km) with a wider range of magnitudes, suggesting that shallow earthquakes are more frequent and can vary significantly in strength.
- **Outliers**: A few outliers exist with high magnitudes at shallow depths, indicating that exceptionally strong earthquakes can occur near the surface.
- These insights highlight the inverse relationship between the depth and magnitude of earthquakes, providing valuable information for understanding seismic activity and potential impacts.

World Temperatures Analysis

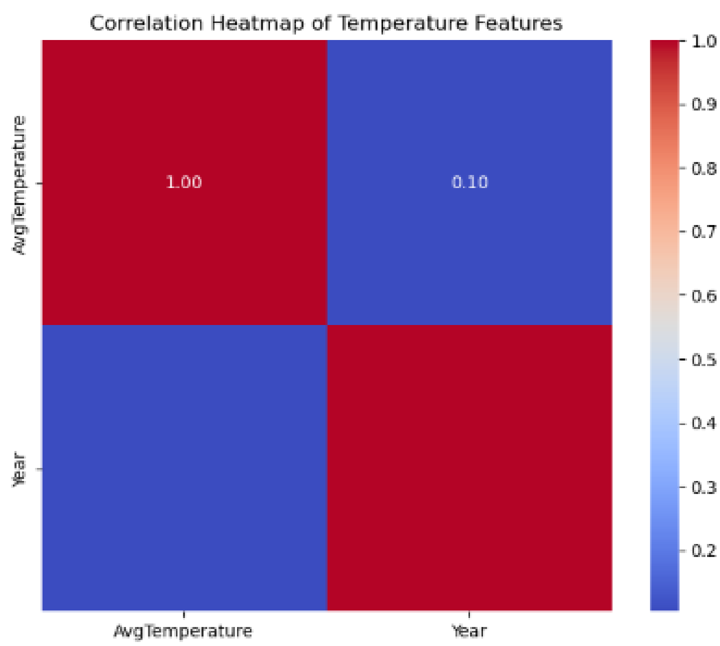
1. **Average Temperature by Region**
   The bar chart presents the average temperatures across different global regions. Key insights include:

   - Middle East: This region has the highest average temperature, exceeding 60°F, indicating a generally warmer climate.
   - Asia: Following the Middle East, Asia also shows a high average temperature, reflecting its diverse climates but generally warmer conditions.
   - Australia/South Pacific: This region has a moderate average temperature, lower than Asia and the Middle East, but still significant.

- Africa: The average temperature is slightly lower than that of Australia/South Pacific, suggesting a variety of climates across the continent.
- Europe: This region has the lowest average temperature among those displayed, indicating cooler climatic conditions.



2. Correlation Heatmap of Temperature Features



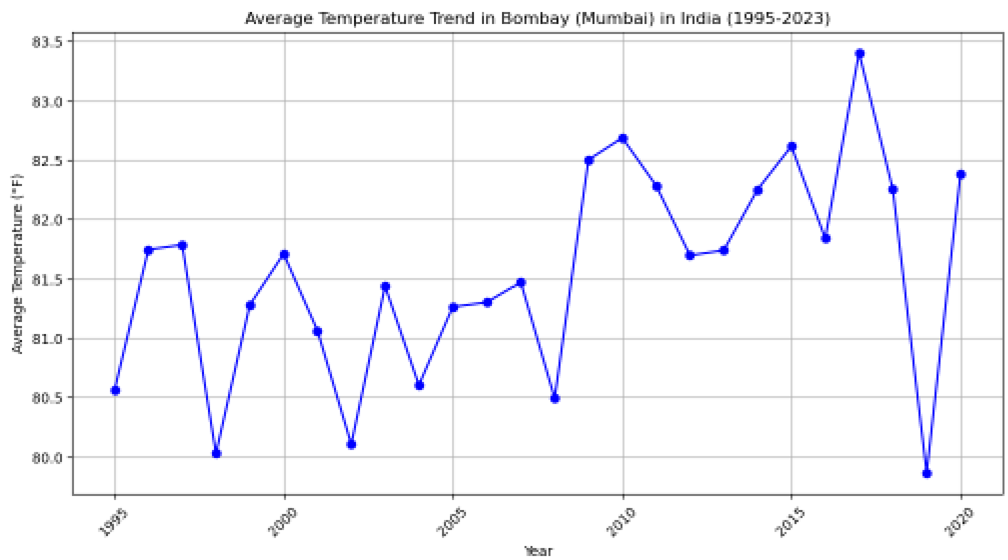
The heatmap illustrates the correlation between average temperature and year. Key insights include:

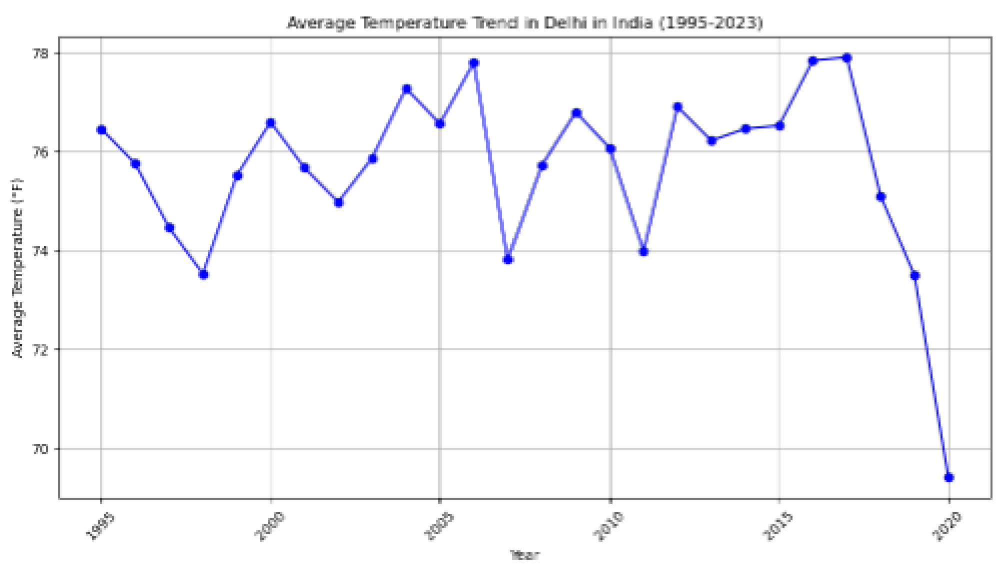
- **Strong Positive Correlation**: The correlation value of **1.00** between average temperature and itself indicates perfect correlation, as expected.

- **Weak Correlation with Year**: The correlation value of **0.10** between average temperature and year suggests a very weak positive relationship. This implies that changes in average temperature over the years are minimal and not significantly influenced by time.

These insights indicate that while average temperatures remain relatively stable over the observed period, there is little evidence of a clear trend or change year over year. This could be important for understanding long-term climate patterns and variability.

3. Average Temperature Trends in Mumbai and Delhi (1995-2023) India





The graphs display the average temperature trends for two major cities in India, Mumbai and Delhi, from 1995 to 2023.

Mumbai:

- **Fluctuating Temperatures**: The average temperatures show significant fluctuations over the years, with peaks and troughs indicating variability in climate.
- **Rising Trend**: Overall, there appears to be a slight upward trend in average temperatures, particularly noticeable in the latter years of the period.

Delhi:

- **Gradual Decline**: The temperature trend in Delhi reveals a more gradual decline over the years, especially noticeable towards the end of the observed period.
- **Less Variation**: Compared to Mumbai, Delhi shows less fluctuation, indicating a more stable climate pattern during this timeframe.

Conclusion
These trends highlight the different climatic responses of Mumbai and Delhi, which can be vital for urban planning, public health, and environmental policy in response to climate change.

Conclusion

1. **Interconnected Climate Dynamics**: The analysis of the five datasets—precipitation, temperatures, earthquakes, CO2 emissions, and world energy consumption—reveals a complex interplay among these variables. Understanding these relationships is crucial for addressing global sustainability challenges.
2. **Temperature Variability**: The temperature trends in Mumbai and Delhi indicate significant regional differences. While Mumbai experiences fluctuations with a slight upward trend, Delhi shows a gradual decline, suggesting varying climatic impacts that could inform urban planning and public health strategies.
3. **CO2 Emissions and Climate Change**: The findings underscore the urgent need for global cooperation in reducing CO2 emissions, particularly from the largest contributing countries. The data highlights stark disparities between high and low emitters, reinforcing the importance of targeted climate action.
4. **Precipitation Trends**: The precipitation analysis indicates significant variations across countries, with notable extremes in both high and low precipitation areas. This variability can impact water resources and agricultural practices, necessitating adaptive management strategies.
5. **Seismic Activity Insights**: The earthquake analysis provides essential insights into the frequency and magnitude of seismic events, emphasizing the need for preparedness in vulnerable regions. The correlation between earthquake features offers valuable data for risk assessment and mitigation strategies.

Recommendations

1. **Enhanced Data Monitoring**: Establish continuous monitoring systems for climate variables, including temperature, precipitation, CO_2 emissions, and seismic activity, to inform policy decisions and enhance preparedness for climate-related challenges.
2. **Sustainable Practices**: Promote the adoption of renewable energy sources and energy-efficient technologies to mitigate CO_2 emissions. Encourage countries to transition towards sustainable energy practices that can reduce environmental impact.
3. **Water Resource Management**: Develop comprehensive water management strategies to address changing precipitation patterns, ensuring water security for agriculture and urban areas. Implement programs to enhance water conservation and efficiency.
4. **Disaster Preparedness and Response**: Strengthen disaster preparedness plans, particularly in regions prone to earthquakes and extreme weather events. Invest in early warning systems and community education programs to improve resilience.
5. **Collaborative Efforts**: Foster partnerships among governments, NGOs, and the private sector to tackle climate change collectively. Encourage cross-border collaborations to address shared environmental challenges and develop joint solutions.
6. **Public Awareness Campaigns**: Increase public awareness of climate change impacts and encourage community engagement in sustainable practices. Educational initiatives can empower individuals to contribute to climate action at the local level