Unveiling the Climate Crisis: Quantifying the Impact of CO_2 Emissions on Climate-Related Disasters

1.1 Introduction

The alarming rise in climate-related disasters has become a pressing global concern. This project investigates a crucial question:

"To what extent does the increase in CO_2 emissions over the past few decades correlate with the increased frequency and severity of climate-related disasters globally?"

This data science investigation uses comprehensive datasets and advanced analytical techniques to examine the link between CO_2 emissions and climate-related disasters over the past few decades. The project also seeks to track the trajectory of CO_2 emissions, and determine their correlation with the frequency and intensity of climate disasters.

1.2 Data Sources

The project utilizes two primary datasets, each with specific licensing and usage obligations:

1. **CO_2 Emissions (CSV):** This comprehensive dataset, sourced from Kaggle and licensed under CC0: Public Domain, provides a detailed record of CO_2 emissions from 1750 to 2022. Its use requires attribution to the author, who intends for the data to support research on global warming.

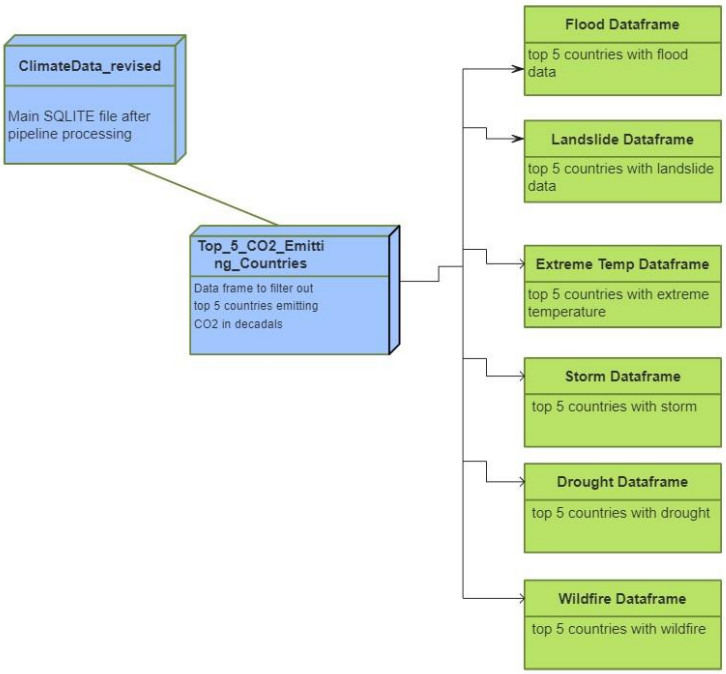
2. **Climate Change Data (IMF) (CSV):** This dataset, provided by the IMF and sourced from EM-DAT, CRED/UCLouvain, focuses on the relationship between climate change and natural disasters. It is available for non-commercial use only, allowing downloads and linking to the IMF website, but prohibiting reselling, redistribution, or the creation of derivative works.

1.2.1 Curated data for exploratory analysis:

Following the successful execution of the data pipeline, a SQLite file named *"ClimateData_revised"* is generated. Given the extensive time series data spanning from 1980 to 2020, the data has been aggregated to a decadal basis. According to the datasets, there are 6 disaster types:

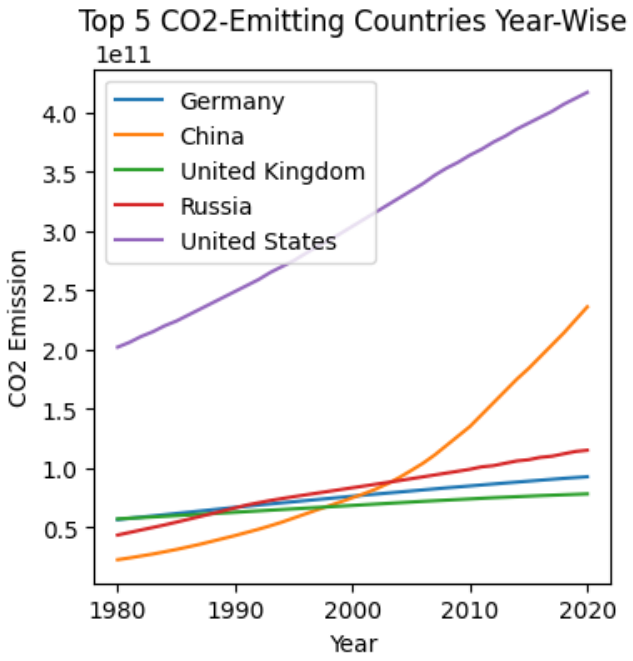
- Flood
- Landslide
- Storm
- Extreme Temperature
- Wildfire
- Drought

This aggregation involves calculating the average CO_2 emissions per decade and the total sum of each disaster type within each decade. Additionally, the top five CO_2 emitting countries across the entire timeframe have been identified. The SQL table is first filtered to include only these top five emitting countries. This filtered dataset is then used as a reference to categorize the data by disaster type, with corresponding CO_2 emissions for each decade. The detailed methodology employed in this analysis is illustrated in the schematic diagram below.



1.3 Analysis

After aggregating the data to calculate the total disaster count and average CO_2 emissions per decade, the following insights were obtained:

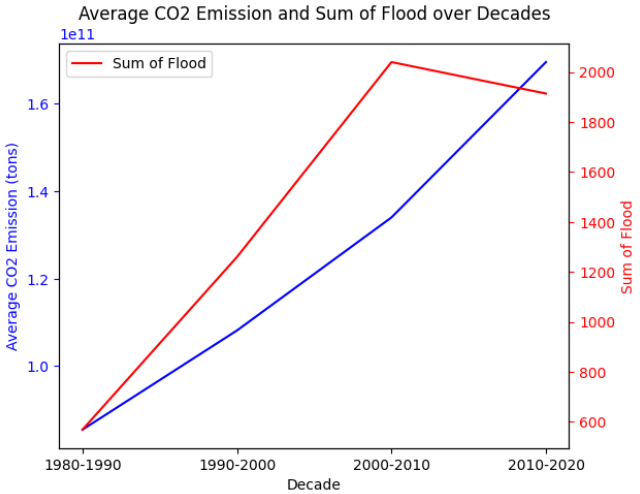


The diagram illustrates that the United States was the highest CO_2 emitter across the decades spanning from 1980 to 2020.

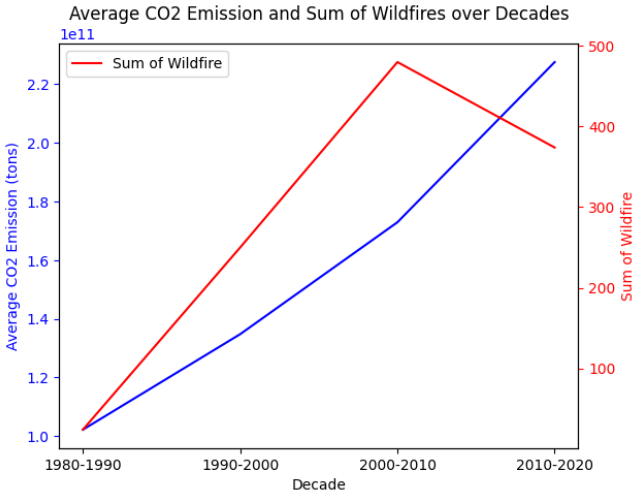
For each disaster type, a visual representation illustrates the relationship between the total number of disaster occurrences and the corresponding CO_2 emissions per decade.

1. **Average CO_2 emission vs Flood over Decades:**

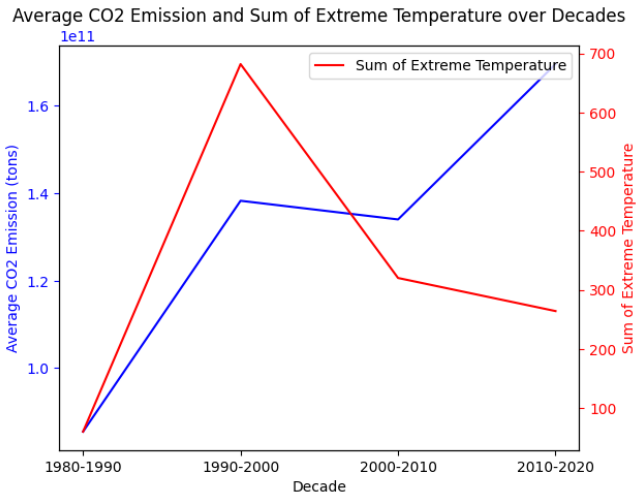
The data indicates a significant increase in flood occurrences until the 2010s, followed by a slight decline.



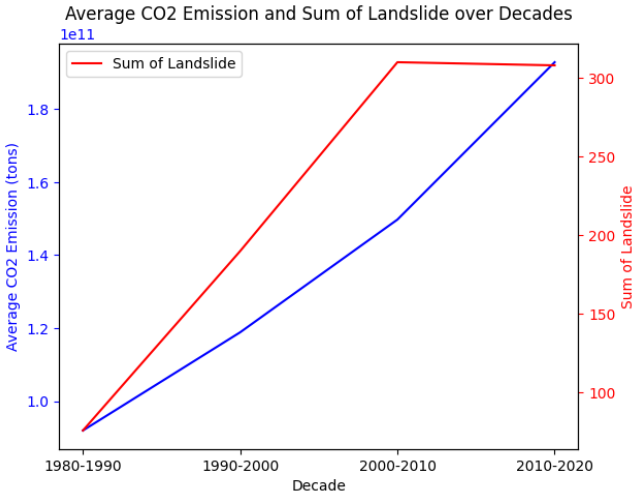
2. **Average CO_2 emission vs Wildfires**:
The data shows a significant rise in wildfire occurrences leading up to the 2010s, followed by a relative decline.



3. **Average CO_2 emission vs Extreme Temperature** :
The figure demonstrates a correlation between rising CO_2 levels and an increase in extreme temperature events until the 2010s, followed by a sharp decline afterwards.

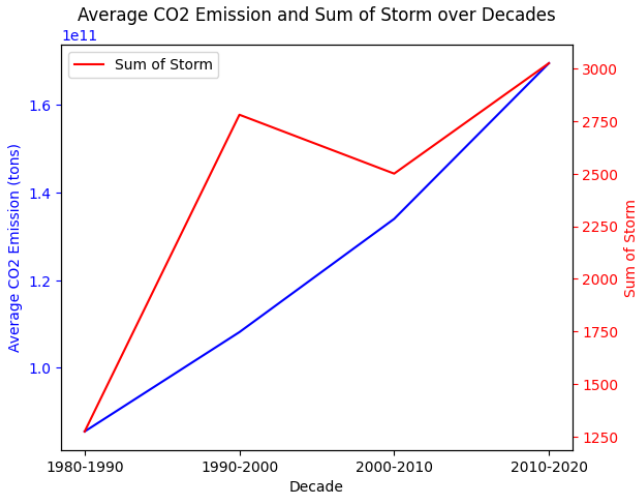


4. **Average CO_2 emission vs Landslide**:
The data shows a significant rise in landslide occurrences leading up to the 2010s, followed by period of relative stability.



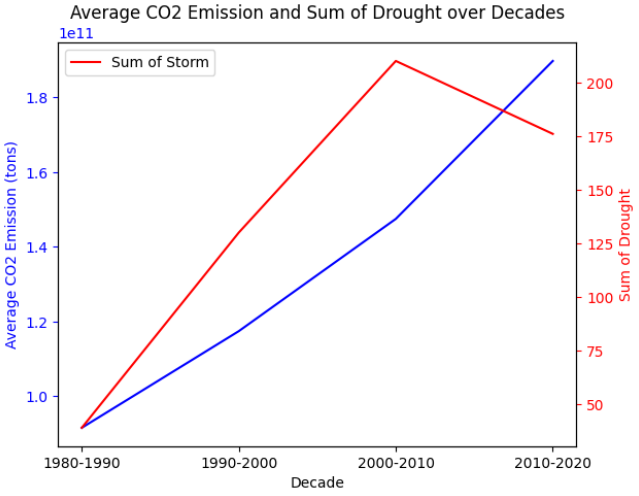
5. **Average CO_2 emission vs Storm**

The data shows a significant rise in storm occurrences leading up to the 2000s, followed by period of decline till 2010s and then again sharp rise.



6. **Average CO_2 emission vs Drought**

The data indicates a significant increase in storm occurrences until the 2010s, followed by a slight decline.

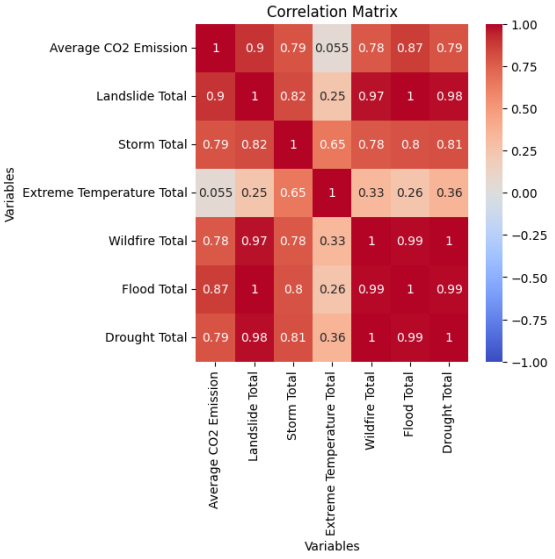


Correlation Results

The major objective of the project is to figure out whether there exists the correlation between average CO_2 emissions and rate of climate related disasters. Based on Pearson Correlation, following correlation results were obtained:

Correlation Between CO_2 Emissions and	**Correlation Value (p)**
Drought	0.79
Flood	0.87
Landslide	0.90
Storm	0.79
Wildfire	0.78
Extreme Temperature	0.05

Correlation Matrix:



Correlation Interpretation:

To assess the relationship between CO_2 emissions and the frequency of climate-related disasters, we employed Pearson correlation analysis. This statistical method quantifies the strength and direction of the linear relationship between two variables, ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation). A value of 0 indicates no linear relationship. By calculating the Pearson correlation coefficient for each disaster type against CO_2 emissions, we aimed to determine the extent to which changes in CO_2 levels are associated

with changes in disaster frequency. This analysis helps us understand the potential impact of CO_2 emissions on specific types of climate-related disasters and inform targeted mitigation strategies.

The Pearson correlation analysis reveals a varying degree of association between CO_2 emissions and the frequency of climate-related disasters:

Strong Positive Correlation:

Landslides: The correlation of 0.90 suggests a strong positive relationship, indicating that higher CO_2 emissions are strongly associated with an increased frequency of landslides.

Floods: A correlation of 0.87 signifies a strong positive relationship between CO_2 emissions and flood occurrences.

Drought: CO_2 emissions and drought occurrences exhibit a strong positive correlation of 0.79, suggesting a significant link between rising emissions and drought frequency.

Storm: With a correlation of 0.79, there is a strong indication that rising CO_2 emissions are linked to an increased frequency of storms.

Wildfire: Rising CO_2 emissions are strongly associated with an increased occurrence of wildfires, as indicated by a correlation of 0.78.

Very Weak Correlation:

Extreme Temperature: The correlation of 0.05 indicates a very weak, almost negligible association between CO_2 emissions and the frequency of extreme temperature events. This suggests that other factors likely play a more significant role in driving extreme temperature events.

1.4 Conclusions

Our analysis reveals a strong positive correlation between rising CO_2 emissions and the frequency of specific climate-related disasters, particularly floods, landslides, droughts, wildfires and storms. This suggests a significant causal link between human induced greenhouse gas emissions and the increased occurrence of these destructive events. However, the observed relationship does not extend to all disaster types, as evidenced by the very weak correlation between CO_2 emissions and extreme temperature events. This indicates that other factors, such as natural climate variability or regional weather patterns, may play a more significant role in driving these events. Furthermore, the data reveals an alarming trend: the United States consistently ranked as the highest CO_2 emitting country across all decades in the study period.

While our findings contribute to the growing body of evidence linking CO_2 emissions to climate-related disasters, several limitations warrant acknowledgment. First, the available data may not fully capture the true extent of these disasters, particularly in less developed regions with limited reporting capabilities. Second, our analysis focused on the correlation between CO_2 and disaster frequency, not on the severity or intensity of individual events. Further research is needed to explore this aspect. Finally, while the correlation analysis suggests a causal link, definitively proving causality requires more sophisticated statistical modeling and a deeper understanding of the underlying mechanisms.

Additionally, our reliance on a limited number of datasets may have introduced bias into our analysis. Conducting correlation tests using a wider variety of datasets could potentially yield different results.