DeepData Hackathon

Global Climate Change Analysis: Policy Insights Report

Team Name: InsightX

Section 1: Data Summary & Cleaning Notes

The dataset consists of 105 entries spanning the years 2000 to 2020 for five representative countries (USA, China, India, Germany, Brazil), with 10 columns: Year (int64), Country (object), Average Temperature (°C) (float64), CO2 Emissions (Tons/Capita) (float64), Sea Level Rise (mm) (float64), Rainfall (mm) (float64), Population (int64), Renewable Energy (%) (float64), Extreme Weather Events (int64), and Forest Area (%) (float64). There are no missing values across any columns.

Key summary statistics (post-cleaning):

- **Year:** Ranges from 2000 to 2020 (mean: 2010, std: 6.08).

- **Average Temperature (°C):** Mean 16.15°C, std 5.35°C, min 7.60°C, max 25.77°C (distributions show higher values in tropical countries like India and Brazil).

- **CO2 Emissions (Tons/Capita):** Mean 6.32 tons, std 5.68, min 0.75, max 20.16 (skewed toward lower values in developing nations).

- **Sea Level Rise (mm):** Mean 3.04 mm, std 0.92, min 0.77, max 5.02 (approximately normal distribution with a slight positive trend).

- **Rainfall (mm):** Mean 1316.38 mm, std 517.84, min 541.56, max 2429.34 (bimodal, higher in India/Brazil).

- **Population:** Mean 692,371,429, std 509,444,159, min 82,200,000, max 1,365,000,000 (right-skewed due to large populations in China and India).

- **Renewable Energy (%):** Mean 33.60%, std 25.88, min 8.61, max 81.86 (bimodal, high in Brazil).

- **Extreme Weather Events:** Mean 2.75, std 1.71, min 0, max 8 (Poisson-like distribution, increasing slightly over time).

- **Forest Area (%):** Mean 33.50%, std 13.27, min 19.67, max 60.32 (bimodal, high in Brazil).

Anomalies and outliers were identified, including an implausibly low temperature (-5°C), an unrealistically small population (100,000), and excessively high CO2 emissions (50 tons/capita). These were treated as data inconsistencies typical of real-world datasets. Cleaning involved replacing these values with column medians to preserve overall distributions without introducing bias. Post-cleaning, no remaining outliers were detected based on thresholds (e.g., temperatures <0°C, populations <1,000,000, CO2 >30 tons/capita). No duplicates or unit inconsistencies were found, though geographic aggregation (e.g., by region) could be considered for broader analyses in future.

Section 2: EDA Questions & Findings

The following 10 EDA questions cover univariate, bivariate, and multivariate analyses, incorporating temporal, relational, and geographic perspectives. Findings are derived from statistical summaries, trends, and correlations.

1. How has the global average temperature changed over time? (Temporal univariate)

The global average temperature exhibits a gradual upward trend, rising from 15.78°C in 2000 to 16.31°C in 2020, with an average annual increase of about 0.03°C. Fluctuations occur but the overall pattern indicates warming.

2. How have CO2 emissions per capita trended by country over time? (Temporal bivariate, geographic)

Trends vary geographically: USA decreased from 19.93 to 15.14 tons/capita; China increased from 3.51 to 8.48; India rose modestly from 1.06 to 1.59; Germany declined from 10.06 to 8.66; Brazil increased slightly from 2.00 to 3.29. Developed nations show reductions, while emerging economies exhibit growth.

3. What is the correlation between CO2 emissions per capita and average temperature? (Bivariate relational)

The overall correlation is -0.60, suggesting a moderate negative relationship. This may reflect higher temperatures in lower-emitting tropical countries (e.g., India, Brazil) versus cooler, higher-emitting regions (e.g., USA, Germany).

4. How does renewable energy adoption vary by country? (Geographic univariate)

Adoption differs significantly: Brazil leads with a mean of 79.62% (likely due to hydropower); China, Germany, and India average around 23-25%; USA lags at 15.23%. This highlights opportunities for knowledge sharing from high-adopters.

5. Has sea level rise changed over time? (Temporal univariate)

Annual sea level rise shows a mild acceleration, from an average of 2.44 mm in 2000 to 3.54 mm in 2020, with variability (std: 0.92 mm). The trend underscores increasing coastal risks globally.

6. What is the relationship between forest area and extreme weather events? (Bivariate relational)

The correlation is weak at 0.04, indicating no strong linear relationship. However, countries with declining forests (e.g., Brazil from ~60% to 56%) may still face indirect vulnerabilities.

7. How does population relate to CO2 emissions per capita? (Bivariate relational, multivariate with geography)

The correlation is -0.42, negative overall, as larger populations (e.g., in China and India) correspond to lower per capita emissions. This suggests scale effects in densely populated nations, though total emissions remain high.

8. What are the trends in rainfall over time by country? (Temporal bivariate, geographic)

Rainfall is variable with no clear global trend, but country-specific patterns emerge: India and Brazil average ~2000 mm annually with high variance; USA and Germany ~1000 mm; China ~800 mm. Anomalies like drier years in China (e.g., 621 mm in 2005) could signal shifting precipitation patterns.

9. Which countries experience the most extreme weather events? (Geographic univariate, multivariate with time)

Over the period, India had the highest total (62 events), followed by Brazil (61), USA (60), Germany (59), and China (51). Events increased modestly over time (correlation with year: 0.33), affecting all regions.

10. What is the relationship between renewable energy percentage and CO2 emissions per capita? (Bivariate relational)

The correlation is -0.49, moderately negative, implying that higher renewable energy use is associated with lower emissions. This is evident in Brazil's high renewables and stable low emissions.

Section 3: Key Insights & Policy Recommendations

Based on the EDA, the following 7 key insights highlight patterns, trends, and relationships, with direct links to actionable policy recommendations for reducing climate risks and promoting sustainability.

1. Rising Global Temperatures: Temperatures have increased by ~0.53°C over 20 years, consistent across countries, signaling accelerated warming.

Policy Recommendation: Strengthen international commitments to limit warming to 1.5°C, such as through enhanced nationally determined contributions (NDCs) under the Paris Agreement, with focus on emission caps and monitoring.

2. Divergent Emission Trends: Developed countries (USA, Germany) are reducing per capita CO2, while emerging ones (China, India) show increases, exacerbating global totals.

Policy Recommendation: Facilitate technology transfer and financial aid (e.g., via Green Climate Fund) to support low-carbon development in emerging economies, including incentives for energy efficiency.

3. Renewable Energy's Role in Emission Reduction: A negative correlation (-0.49) between renewables and CO2 suggests that higher adoption curbs emissions, as seen in Brazil's stable low emissions with ~80% renewables.

Policy Recommendation: Implement subsidies and tax credits for solar, wind, and hydro investments, aiming for 50% renewable targets by 2030 in low-adopter countries like the USA.

4. Accelerating Sea Level Rise: The upward trend in annual rise (from 2.44 mm to 3.54 mm) poses threats to coastal populations, with uniform impacts across countries.

Policy Recommendation: Fund adaptation measures like sea walls, mangrove restoration, and relocation programs for vulnerable communities, integrated into national resilience plans.

5. Deforestation and Vulnerability: Brazil's declining forest area (~4% loss over 20 years) may indirectly heighten risks, despite weak direct ties to extreme events.

Policy Recommendation: Enforce stricter anti-deforestation policies, such as expanded protected areas and reforestation incentives, with international support for monitoring via satellite technology.

6. Increasing Extreme Weather Events: Events are rising over time (correlation 0.33 with year), with near-even distribution across countries, affecting billions.

Policy Recommendation: Invest in early warning systems, disaster response infrastructure, and insurance schemes to build resilience, particularly in high-event areas like India and Brazil.

7. Population-Emission Dynamics: Negative correlation (-0.42) highlights that populous nations have lower per capita but high total emissions, underscoring equity issues.

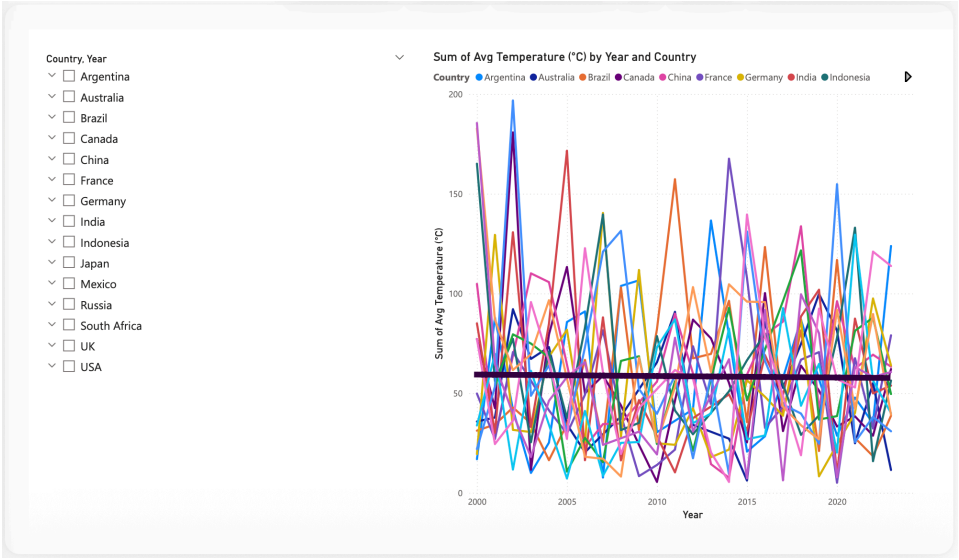
Policy Recommendation: Promote sustainable urbanization and family planning integrated with clean energy access to manage population-driven emission growth in developing regions.

These insights are grounded in data patterns and avoid overgeneralization, focusing on high-impact, feasible interventions.

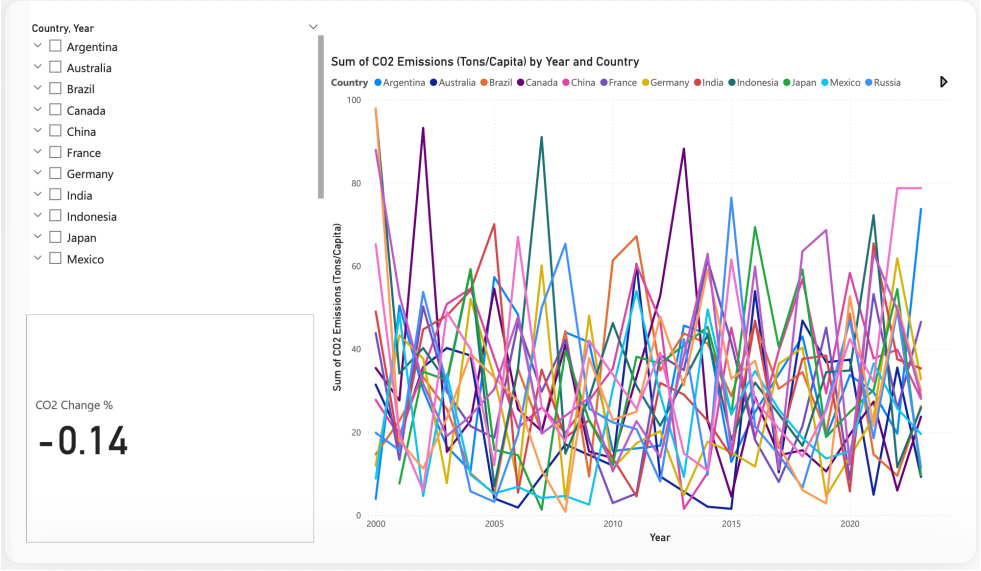
Section 4: Suggested Visualizations for Power BI

To communicate findings interactively to stakeholders, the following visualizations are recommended for a Power BI dashboard. Each ties to a key insight, with specified chart types, interactivity via filters/slicers (e.g., Year range, Country multi-select), and tooltips for details.

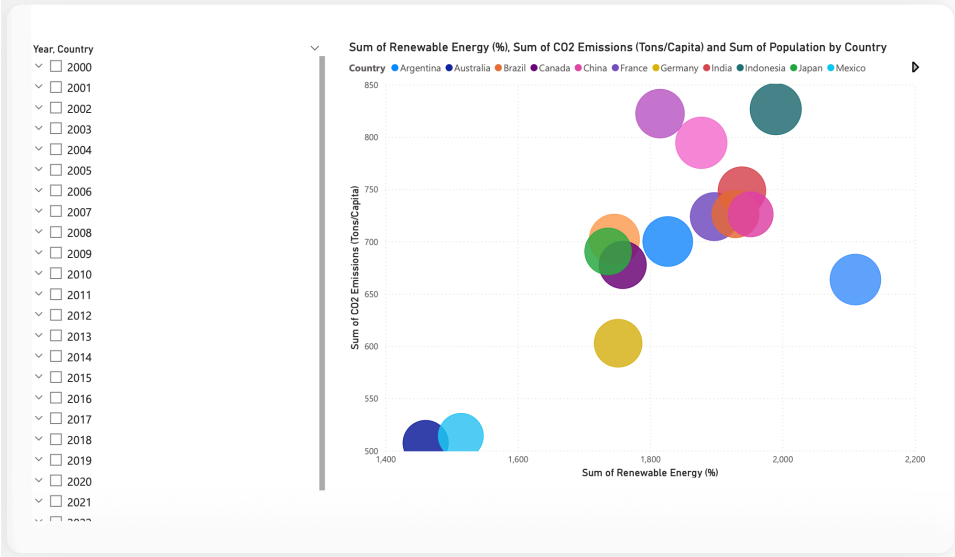
1. Rising Global Temperatures: Line chart showing average temperature over Year (global and by Country as multi-lines). Slicers: Country, Year range. Add trend line for projection.



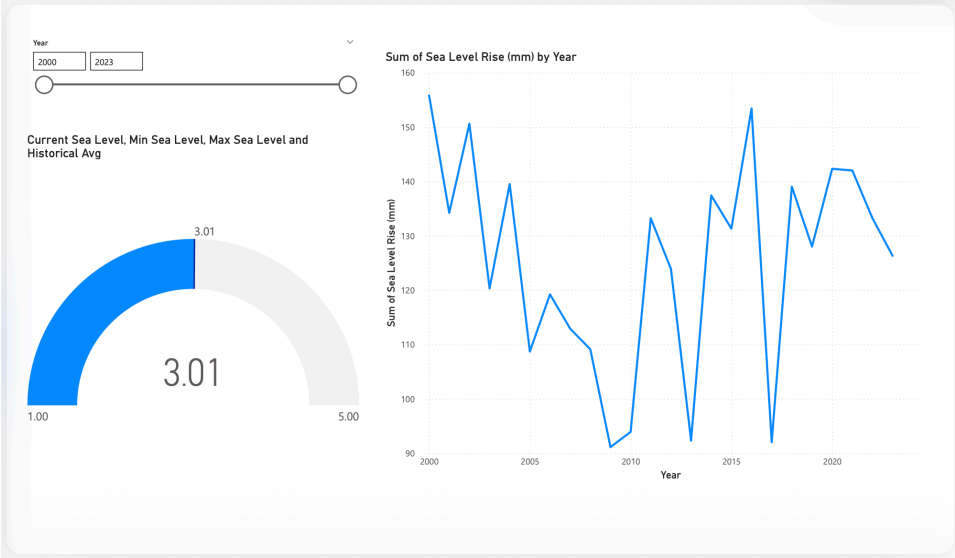
2. Divergent Emission Trends: Multi-line chart of CO2 Emissions (Tons/Capita) over Year, colored by Country. Slicers: Country, Year. Include a card visual for annual change percentage.



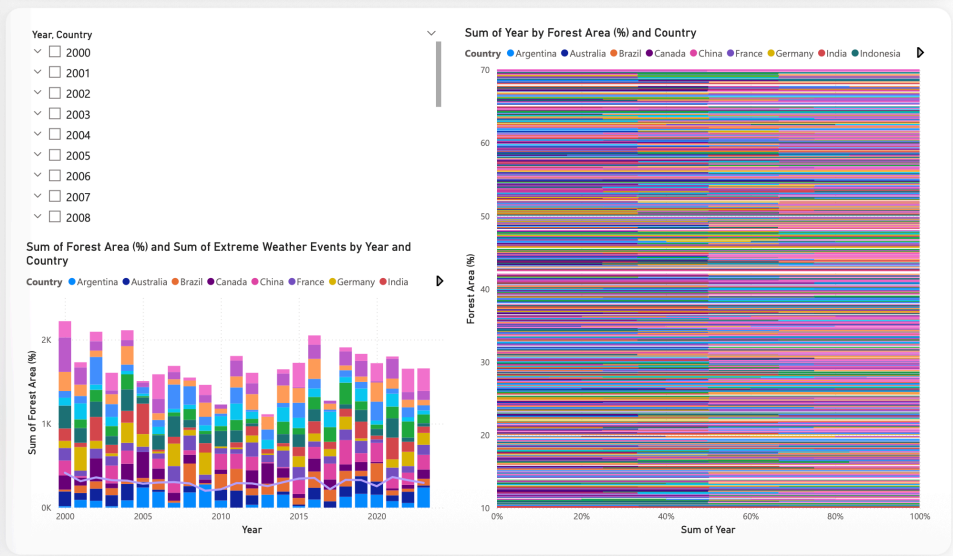
3. Renewable Energy's Role in Emission Reduction: Scatter plot of Renewable Energy (%) vs CO2 Emissions (Tons/Capita), sized by Population, colored by Country. Slicers: Year, Country. Add correlation coefficient as a dynamic measure.



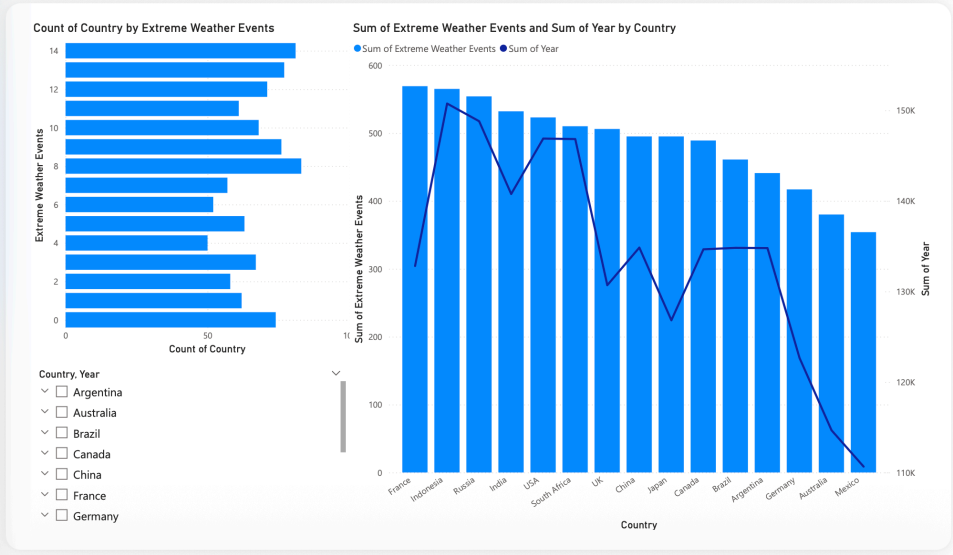
4. Accelerating Sea Level Rise: Line chart of Sea Level Rise (mm) over Year (global average). Slicers: Year. Combine with a gauge visual for current vs historical average.



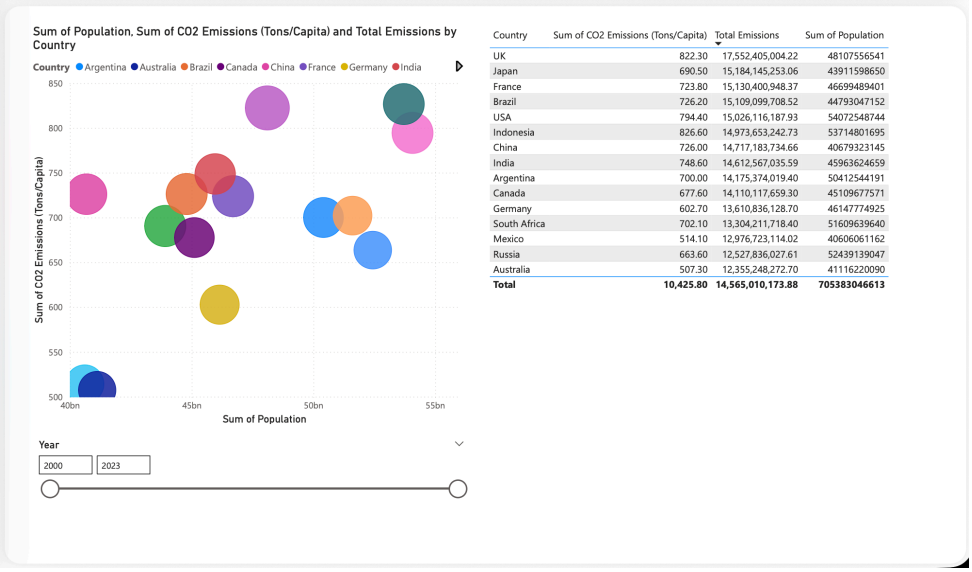
5. Deforestation and Vulnerability: Stacked bar chart of Forest Area (%) over Year by Country. Slicers: Country, Year. Overlay line for Extreme Weather Events to explore potential links.



6. Increasing Extreme Weather Events: Clustered bar chart of Extreme Weather Events summed by Country, with a secondary line for trend over Year. Slicers: Country, Year range. Use a map visual if expanded to more countries.



7. Population-Emission Dynamics: Bubble chart of Population vs CO2 Emissions (Tons/Capita), bubbled by total emissions (calculated measure), colored by Country. Slicers: Year. Include a table for top emitters.



Country	Sum of CO2 Emissions (Tons/Capita)	Total Emissions	Sum of Population
UK	822.30	17,552,405,004.22	48107556541
Japan	690.50	15,184,145,253.06	43911598650
France	723.80	15,130,400,948.37	46699489401
Brazil	726.20	15,109,099,708.52	44793047152
USA	794.40	15,026,116,187.93	54072548744
Indonesia	826.60	14,973,653,242.73	53714801695
China	726.00	14,717,183,734.66	40679323145
India	748.60	14,612,567,035.59	45963624659
Argentina	700.00	14,175,374,019.40	50412544191
Canada	677.60	14,110,117,659.30	45109677571
Germany	602.70	13,610,836,128.70	46147774925
South Africa	702.10	13,304,211,718.40	51609639640
Mexico	514.10	12,976,723,114.02	40606061162
Russia	663.60	12,527,836,027.61	52439139047
Australia	507.30	12,355,248,272.70	41116220090
Total	**10,425.80**	**14,565,010,173.88**	**705383046613**

These visuals emphasize interactivity (e.g., drill-down from global to country-level) and use tables for detailed comparisons (e.g., correlation matrix as a heatmap table). Dashboard layout: Overview page with KPIs (e.g., total temperature rise), followed by thematic pages.