Problem Statement -1

Analyzing Global Climate Change Data for Policy Insights

1. Introduction

In this track, you'll work with a dataset on global climate change indicators from Kaggle. This task simulates how data analysts in environmental organizations or government agencies explore data to inform policies that combat climate change and build resilience.

Your role: Act as a data consultant for an international environmental agency (e.g., like the UN Environment Programme). The agency needs actionable insights from this dataset to recommend policies that reduce climate risks and promote sustainability. Focus on exploratory data analysis (EDA) to uncover patterns, trends, and relationships and no machine learning models required. This keeps the emphasis on understanding the data deeply and creatively.

2. Dataset Overview

The dataset provides climate-related metrics across countries and years. Key columns include:

- **Year**: The year of the data point.
- **Country**: The country or region.
- **Average Temperature (°C)**: Average annual temperature.
- **CO2 Emissions (Tons/Capita)**: Per-person carbon dioxide emissions.
- **Sea Level Rise (mm)**: Annual Sea level increase.
- **Rainfall (mm)**: Total annual rainfall.
- **Population**: Total population.
- **Renewable Energy (%)**: Percentage of energy from renewable sources.
- **Extreme Weather Events**: Number of extreme events (e.g., storms, floods).
- **Forest Area (%)**: Percentage of land covered by forests.

Dataset Link - https://www.kaggle.com/datasets/bhadramohit/climate-change-dataset

The dataset may contain inconsistencies (e.g., outliers in temperatures or populations). treat this as part of real-world data challenge.

Problem Statement

Climate change is accelerating, with rising temperatures, extreme weather, and environmental degradation affecting billions. Governments and organizations need evidence-based policies to mitigate impacts, such as reducing emissions, protecting ecosystems, and enhancing resilience. Using this dataset, explore how factors like emissions, energy use, and natural resources interact across countries and over time. Your analysis should highlight opportunities for policy interventions, like promoting renewables or forest conservation, to build a more resilient world.

Objectives

- Perform EDA to understand the data's structure, quality, and key patterns.
- Develop at least 10 targeted EDA questions (e.g., "How have CO2 emissions trended by country over time?") to guide your exploration.
- Derive insights that could inform real-world policies, such as strategies to lower emissions while boosting renewable energy adoption.
- Visualize findings in an interactive dashboard for stakeholders.
- Present your work as if pitching to policymakers, emphasizing creativity in how insights translate to actionable recommendations.

Analysis Framework: Adapted EDA Roadmap

To structure your work like a professional data project, follow this simple framework (inspired by real-world analytics processes like those used in consulting firms). It provides a clear path but leaves room for your innovative approaches:

1. **Data Understanding**: Load the dataset, review its structure (e.g., data types, missing values), and summarize basic statistics (e.g., means, distributions). Identify any anomalies or cleaning needs.
2. **Data Preparation**: Handle issues like duplicates, outliers, or inconsistencies. For example, normalize units if needed or group data by regions/years for better analysis. Keep it lightweight—focus on enabling exploration.
3. **Question Formulation and Exploration**: Brainstorm and select at least 10 EDA questions. These should cover univariate (single variable), bivariate (relationships between two), and multivariate (multiple variables) analyses. Use statistical methods (e.g., correlations, trends) and visualizations (e.g., histograms, scatter plots, heatmaps) to answer them. Be creative: Think about geographic or temporal angles.
4. **Insight Generation**: From your EDA, extract 5-7 key insights. Link them to policy implications, e.g., "Countries with higher renewable energy adoption show lower emission growth - recommend incentives for solar/wind investments."

5. **Policy Recommendations and Presentation:** Translate insights into 3-5 policy proposals. Prepare a presentation that tells a story or if using PowerBI or tableau, share the dashboard file or link highlighting how your analysis supports resilience-building.

Files to be submitted: GitHub link which contains the following files -

Python Analytics File: A Jupyter Notebook (.ipynb) or Python script (.py)

Dashboard: An interactive visualization tool:

- If Python-based: A dashboard file (e.g., Streamlit app) or a presentation file (ppt). Use visuals from you EDA analysis
- If Power BI/Tableau: Shareable file or link (e.g., published report).

Guidelines and Tips

Tools: Python for analysis (libraries like pandas, numpy, matplotlib, seaborn). For dashboards: Python options, Power BI, or Tableau.

Creativity Space: We won't specify exact questions or policies but use your judgment to explore underrepresented angles, like equity across developing vs. developed countries.

Ethics and Realism: Ensure insights are grounded in data; avoid overgeneralizing. Think like a business analyst: Prioritize high-impact, feasible recommendations.

Collaboration: Teams of 3-4 – divide the work like one for the coding, one for the Research about domain to know exactly how each feature is related to climate change analysis which will help in better insights, one for visualization and one to present those insights with presentation or dashboards.

We're here for questions during the event. Good Luck!