

Project Overview

This project is structured into four components and leverages data visualization and statistical exploration techniques in **R Shiny** to analyze relationships between **air pollution, climate change, forest coverage, and human health outcomes**. The goal is to provide interactive insights into how these environmental factors interrelate across global regions, especially in large and industrially significant countries.

Data Sources

- Our World in Data – <https://ourworldindata.org/air-pollution>
 - Our World in Data – <https://ourworldindata.org/air-pollution#air-pollution-is-one-of-the-world-s-leading-risk-factors-for-death-analysis>
 - IMF Climate Data - <https://climatedata.imf.org/pages/climatechange-data>
 - Worldometers – <https://www.worldometers.info/geography/largest-countries-in-the-world/>
-

1. Data Preparation and Cleaning

Datasets used:

- **Air Quality Index (AQI)** values from a global air pollution dataset.
- **Death percentage from air pollution** from the Our World in Data repository.
- **Annual surface temperature change** sourced from the IMF climate dataset.
- **Forest area share** from a global forestry dataset.

Cleaning and Transformation Steps:

- Removed rows with missing or NA values to ensure accuracy.
- Normalized inconsistent country names across datasets to allow for proper merging.
- Converted year columns from wide to long format using `pivot_longer()` for time series analysis.

- Filtered datasets to include:
 - The **30 largest countries** by land area (for general comparisons).
 - The **G20 nations** (for focused analysis of industrial powers with global impact).
 - Created calculated averages (e.g., average AQI, average forest share) by country for comparison.
-

2. Exploratory Analysis and Aggregation in R

Summary statistics computed:

- **Average AQI** values across 30 largest countries.
- **Average forest coverage percentage** per country.
- **Death percentage from air pollution** over time (filtered to G20).
- **Temperature change trends** aligned with mortality rates.

Merged datasets:

- Combined temperature change and death rate by country and year to explore correlation.
 - Combined forest share and AQI to evaluate the impact of green coverage on pollution levels.
-

3. Interactive Dashboard in R Shiny

The **R Shiny dashboard** was developed using shinydashboard, ggplot2, and plotly to enable user-driven exploration. It features **seven dynamic visualizations**, organized across sidebar menu tabs:

KPI Visualizations:

- **AQI Bar Chart:** Displays average air pollution by country (top 30 largest countries).
- **Death % Over Time:** Line chart for mortality trends from pollution in G20 countries.
- **Temperature vs. Death %:** Scatter plot to evaluate correlation between warming and health impact.

- **Forest Share Bar Chart:** Compares average forest land share among the 30 largest nations.
 - **AQI vs Forest Scatterplot:** Explores the inverse relationship between forestation and pollution.
 - **Top AQI Selector:** Slider-controlled chart for top N countries by pollution levels.
 - **Top Forest Selector:** Numeric-controlled chart to visualize countries with the highest forest area.
-

4. Key Insights and Interpretations

- Countries like **India, Saudi Arabia, Pakistan, and Mauritania** exhibit **high AQI levels** and **low forest coverage**, reinforcing the link between deforestation and pollution.
 - G20 nations show **relatively stable death percentages** from pollution over the last 30 years, despite global initiatives, suggesting room for policy improvement.
 - Nations such as **Canada** and **Germany** maintain **low pollution-related deaths** even with moderate temperature changes, indicating that **healthcare infrastructure** also plays a key role.
 - Scatterplots reveal **only partial correlation** between warming, pollution, and deaths, emphasizing the **multifactorial nature** of climate impact.
 - The dashboard allows users to **interactively compare environmental indicators** across countries and time periods to support deeper exploration.
-