Environmental Degradation Effects on Public Health in the U.S.

I. Executive Summary

This comprehensive study delves into the intricate relationships between various environmental factors, such as resource depletion, emissions, and air pollution, and crucial health indicators, including fertility rate, disease prevalence, and mortality rate. By analyzing extensive datasets and employing advanced statistical techniques, the study reveals significant correlations that underscore the profound impact of environmental conditions on public health.

Preliminary findings indicate that Carbon Dioxide Damage is negatively correlated with Female Survival to Age 65 (-0.70) and Life Expectancy (-0.61), suggesting a direct detrimental effect of increased carbon emissions on female longevity and overall life expectancy. Similarly, Annual Freshwater Withdrawals show a positive correlation with Adult Male Mortality Rate (0.60), highlighting the adverse consequences of excessive water use on male mortality. Methane Emissions are negatively correlated with Male Survival to Age 65 (-0.56), indicating that higher methane levels reduce male survival rates.

Cross-validation using various regression models, including linear regression, regression trees, random forest regression, and gradient boosting regression, revealed that environmental factors significantly influence health outcomes, though these relationships are often complex and non-linear. Key findings include:

- **Death Rate:** Significantly impacted by carbon dioxide damage, CO2 emissions, and methane emissions.
- **Life Expectancy:** Influenced by net forest depletion and mineral depletion, alongside other factors such as energy depletion and natural resource depletion.
- **Fertility Rate:** Primarily affected by particulate emission damage, net forest depletion, and mineral depletion.
- **Mortality from CVD, Cancer, Diabetes, or CRD:** Strongly influenced by nitrous oxide emissions, water stress, and air pollution.

In conclusion, the study highlights the urgent need for policymakers to implement stricter regulations on emissions and resource depletion to safeguard public health. Future research should aim to explore the combined effects of environmental and demographic factors on health outcomes, utilizing more sophisticated analytical techniques and comprehensive datasets.

II. Introduction

The context of this study is situated at the intersection of environmental science and public health, focusing on how resource depletion, emissions, and air pollution affect human health. Previous research has consistently shown that environmental degradation is linked to adverse health outcomes. For example, studies have demonstrated that air pollution increases the incidence of “cardiopulmonary morbidity and mortality as well as reproductive disorders and cancers” (Kumar Yadav et. al, 2023). Similarly, resource depletion, such as deforestation and water scarcity, has been associated with negative health impacts, including malnutrition and increased disease transmission (Goyal, 2022).

This study aims to build on these findings by specifically examining the relationships between various environmental factors and health indicators in the United States. The central hypothesis is that environmental degradation significantly contributes to increased mortality rates and overall health decline. By analyzing extensive datasets from the World Bank's World Development Indicators, this report seeks to elucidate the specific environmental factors that most strongly impact public health and provide actionable insights for policymakers.

III. Data and Methodology

The dataset utilized in this analysis, "World Development Indicators," was sourced from The World Bank, encompassing a wide range of development indicators for numerous countries over several years (World Bank Group, 2024). This study focuses on data relevant to the United States to assess the impact of environmental factors on health, fertility rates, and mortality rates.

Data Cleaning and Preparation: Initial data cleaning involved filtering out data from countries other than the United States. Numerical data with missing values were imputed using the K-Nearest Neighbors (KNN) algorithm due to its robustness in handling missing data compared to other methods. Categorical data were found to be complete with no missing values. The cleaned dataset was exported to a CSV file for further processing in Excel, where relevant indicators were selected based on the study's focus.

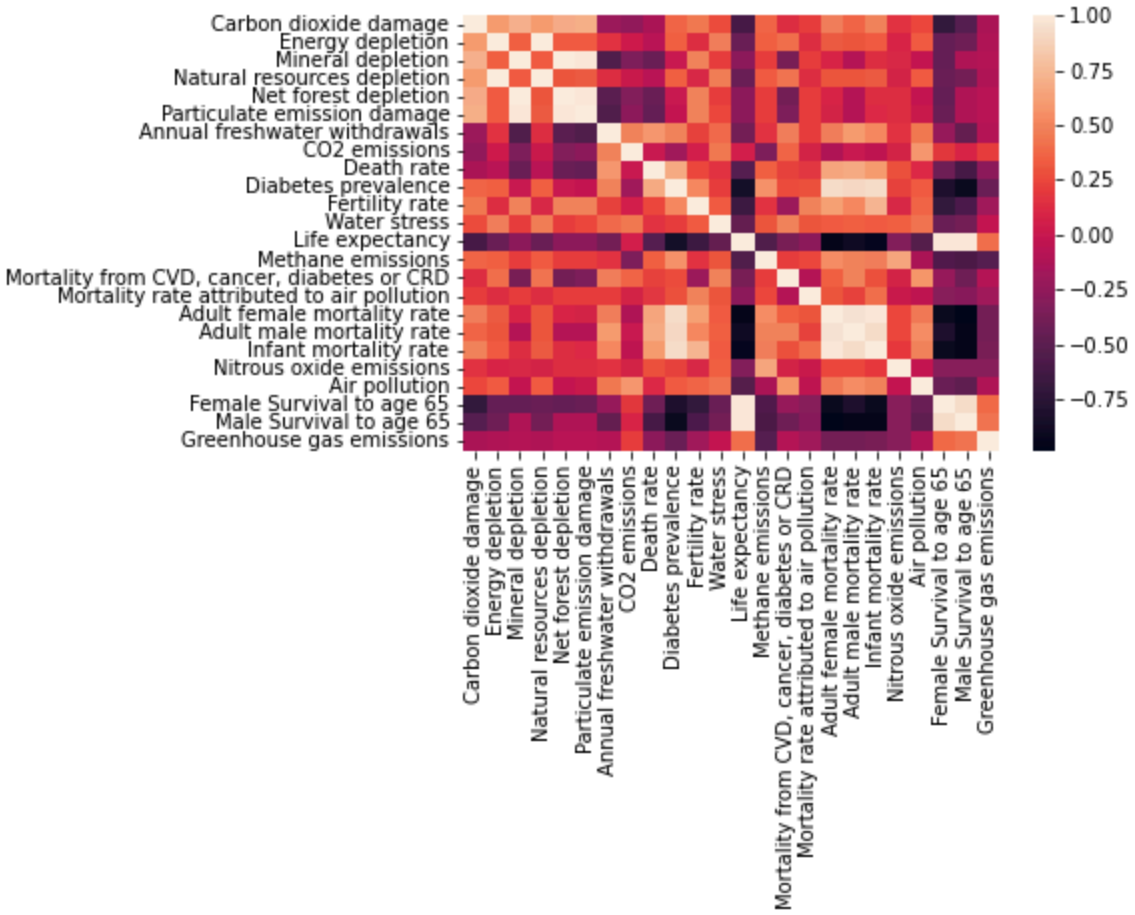
Selection of Indicators: Environmental factors considered in the study included Carbon Dioxide Damage, Energy Depletion, Mineral Depletion, Natural Resources Depletion, Net Forest Depletion, Particulate Emission Damage, Annual Freshwater Withdrawals, CO2 Emissions, Water Stress, Methane Emissions, Nitrous Oxide Emissions, Air Pollution, and Greenhouse Gas Emissions. Health indicators included Death Rate, Diabetes Prevalence, Fertility Rate, Life Expectancy, Mortality from CVD, Cancer, Diabetes or CRD, Mortality Rate Attributed to Air Pollution, Adult Female Mortality Rate, Adult Male Mortality Rate, Infant Mortality Rate, Female Survival to Age 65, and Male Survival to Age 65.

Data Processing: After selecting the relevant indicators, the dataset was read back into Python for further cleaning. Unnecessary columns such as country, country code, and indicator code were removed. The indicators, initially housed in one column, were transposed to become the names of each column, facilitating easier analysis.

IV. Analysis and Visualizations

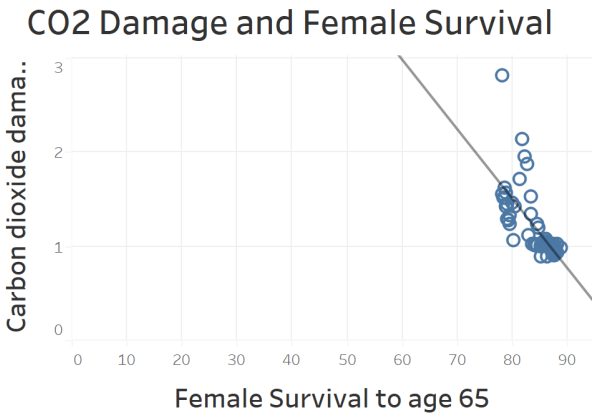
Exploratory Data Analysis: The cleaned dataset was subjected to exploratory data analysis (EDA) to uncover underlying patterns and relationships. A correlation matrix was constructed and visualized using a heatmap to identify the highest correlations between environmental and health factors. This heatmap revealed several significant correlations, which were further analyzed using scatterplots generated in Tableau.

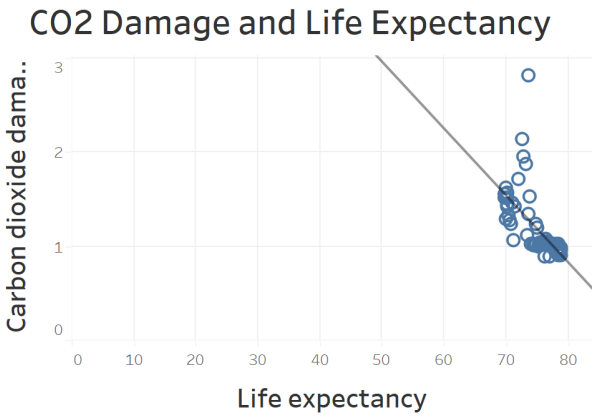
Visualizations:

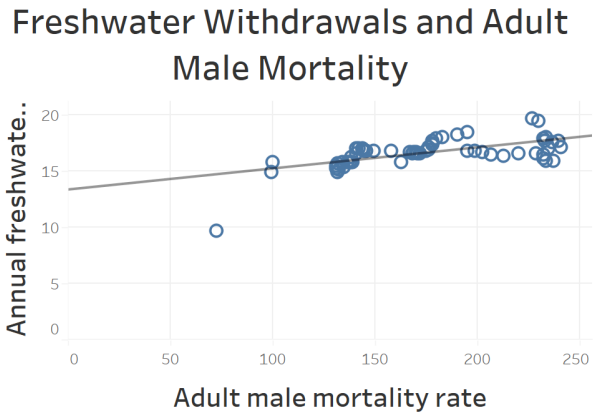


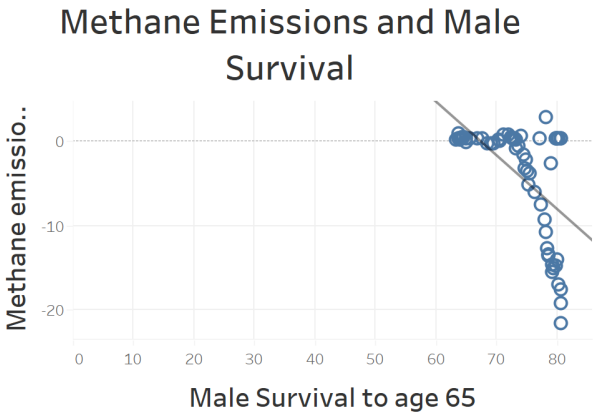
- **Heatmap:** The heatmap provided a comprehensive view of the correlation coefficients between all variables, highlighting significant relationships such as the negative correlation between Carbon Dioxide Damage and Life Expectancy, and the positive correlation between Annual Freshwater Withdrawals and Adult Male Mortality Rate.
- **Scatterplots:** Detailed scatterplots were created to visualize specific high-correlation pairs:

Top Correlations









1. **Carbon Dioxide Damage vs. Life Expectancy:** A clear negative trend was observed, indicating that increased carbon emissions are associated with reduced life expectancy.
2. **Annual Freshwater Withdrawals vs. Adult Male Mortality Rate:** A positive correlation suggested that higher freshwater withdrawals are linked to increased male mortality.
3. **Methane Emissions vs. Male Survival to Age 65:** Higher methane emissions were associated with lower male survival rates.
4. See Appendix B for additional scatterplots illustrating other significant relationships and emphasizing the strong impact of environmental factors on various health outcomes.

Model Fitting: To determine the best-fitting models, cross-validation was conducted using multiple regression techniques, including linear regression, regression trees, random forest regression, and gradient boosting regression. Models were evaluated based on Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared metrics. The analysis focused on key health outcomes such as death rate, fertility rate, life expectancy, and mortality from CVD, cancer, diabetes, or CRD. Results indicated:

- **Death Rate:** Best predicted by random forest regression, influenced by carbon dioxide damage, CO2 emissions, and methane emissions.
- **Fertility Rate:** Random forest regression provided the best fit, with particulate emission damage, net forest depletion, and mineral depletion being the most significant predictors.
- **Life Expectancy:** Linear regression was most suitable, with net forest depletion and mineral depletion as key factors.
- **Mortality from CVD, Cancer, Diabetes, or CRD:** Gradient boosting regression emerged as the best model, influenced by nitrous oxide emissions, water stress, and air pollution.

V. Results

The analysis revealed the most significant environmental factors affecting various health outcomes:

- **Death Rate:** Carbon dioxide damage, CO2 emissions, and methane emissions emerged as the most significant environmental predictors of death rate. This underscores the lethal impact of air pollution and greenhouse gas emissions on public health.
- **Life Expectancy:** Net forest depletion and mineral depletion were found to be the primary environmental factors influencing life expectancy, highlighting the importance of sustainable resource management in enhancing longevity.
- **Fertility Rate:** Particulate emission damage, net forest depletion, and mineral depletion were identified as the key environmental factors affecting fertility rates, suggesting that air pollution and resource depletion can have profound effects on reproductive health.
- **Mortality from CVD, Cancer, Diabetes, or CRD:** Nitrous oxide emissions, water stress, and air pollution were the most significant factors influencing mortality from chronic diseases, indicating the critical role of air quality and water availability in chronic disease prevalence.

VI. Discussion and Recommendations

The study provides compelling evidence of the significant impact of environmental factors on public health. However, it is essential to consider potential confounders and limitations. Demographic factors such as age, socioeconomic status, and lifestyle choices can also influence health outcomes and should be incorporated into future research to provide a more comprehensive understanding of these relationships.

Recommendations:

- **Policy Interventions:** Policymakers should enforce stricter regulations on emissions and resource depletion to mitigate the adverse health effects of environmental degradation. Policies promoting renewable energy, sustainable resource management, and pollution control are essential.
- **Further Research:** Future studies should incorporate a broader range of demographic data and utilize advanced methodologies, such as machine learning and multivariate analysis, to explore the combined effects of environmental and demographic factors on health outcomes.
- **Public Awareness:** Raising public awareness about the health impacts of environmental degradation is crucial. Education campaigns can help communities understand the importance of environmental sustainability and support for regulatory measures.

VII. References

Goyal, P. (2022, August 11). Depleting Resources & How it is affecting our health. Times of India Blog. https://timesofindia.indiatimes.com/blogs/voices/depleting-resources-how-it-is-affecting-our-health/

Virendra Kumar Yadav, Sangha Bijekar, Amel Gacem, Abdullah M Alkahtani, Krishna Kumar Yadav, Maha Awjan Alreshidi, Pankaj Kumar, Tathagata Ghosh, Rakesh Kumar Verma, Sunidhi Mishra, Ashish Patel, Nisha Choudhary, The Impact of Fine Particulate Matters (PM10, PM2.5) from Incense Smokes on the Various Organ Systems: A Review of an Invisible Killer, Particle & Particle Systems CharacterizationParticle & Particle Systems CharacterizationParticle & Particle Systems Characterization, 10.1002/ppsc.202300157, 41, 5, (2023).

World Bank Group. (2024, May 30). World Development Indicators. The World Bank. https://datacatalog.worldbank.org/search/dataset/0037712/World-Development-Indicators

VIII. Appendix

Appendix A:

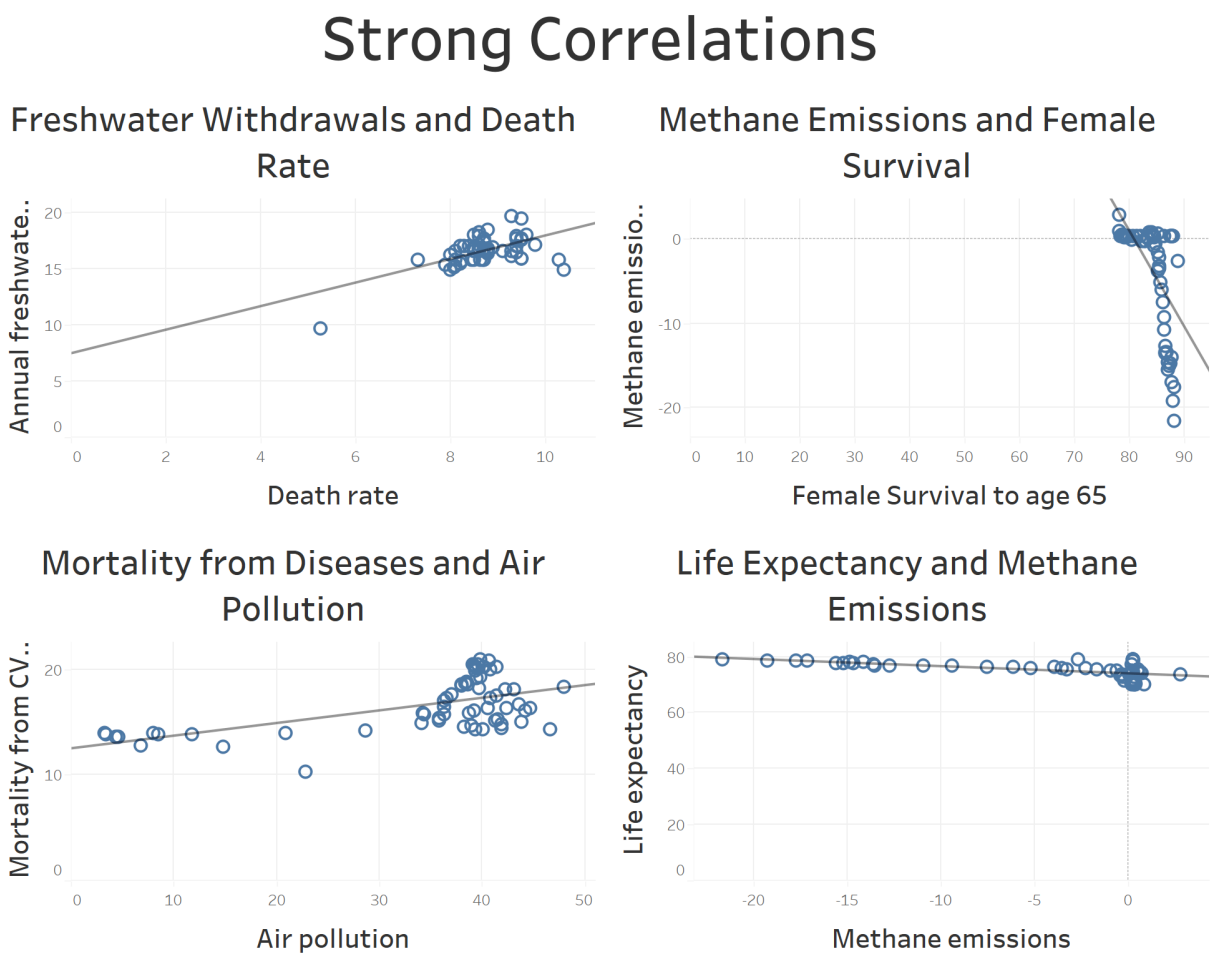
Top Correlations

The table below presents the absolute values of the most significant correlation coefficients identified in the study, providing a clear ranking of the top correlations between environmental factors and health outcomes.

Indicator	Indicator	Corr
Carbon dioxide damage	Female Survival to age 65	0.70139
Carbon dioxide damage	Life expectancy	0.608756
Annual freshwater withdrawals	Adult male mortality rate	0.596998
Methane emissions	Male Survival to age 65	0.591154
Annual freshwater withdrawals	Death rate	0.574973
Methane emissions	Female Survival to age 65	0.572576
Mortality from CVD, cancer, diabetes or CRD	Air pollution	0.568377
Life expectancy	Methane emissions	0.563594
Diabetes prevalence	Methane emissions	0.563541
Adult male mortality rate	Air pollution	0.543916
Life expectancy	Air pollution	0.539378
Methane emissions	Adult female mortality rate	0.537842
Carbon dioxide damage	Male Survival to age 65	0.521687
Annual freshwater withdrawals	Infant mortality rate	0.504771
Methane emissions	Adult male mortality rate	0.503458
Net forest depletion	Fertility rate	0.500366

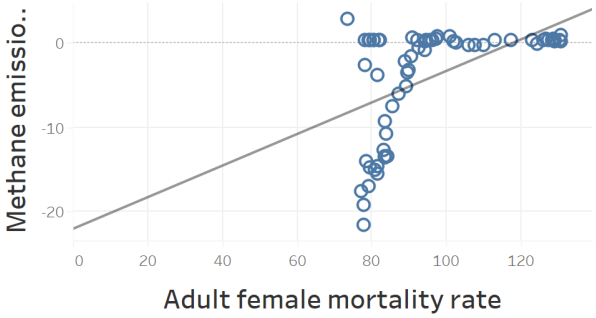
Appendix B:

Significant Correlations Among Other Environmental and Health Factors

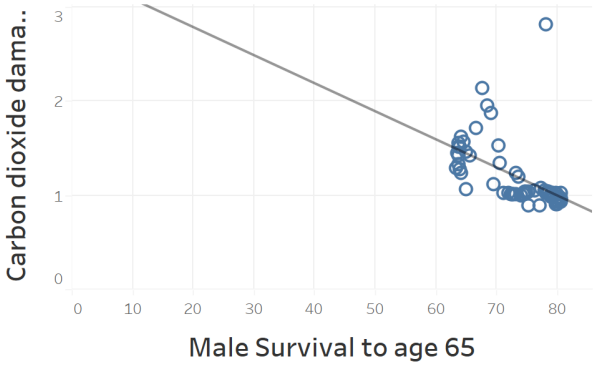


Significant Correlations

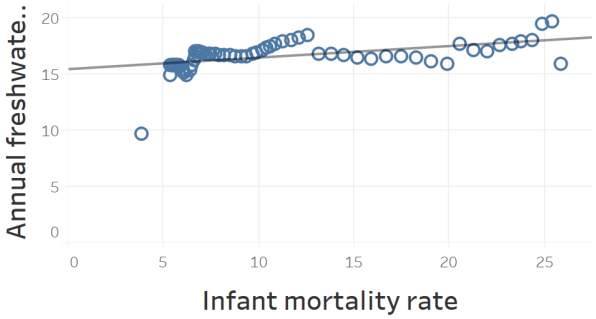
Methane Emissions and Adult Female Mortality



CO2 Damage and Male Survival



Freshwater Withdrawals and Infant Mortality



Methane Emissions and Adult Male Mortality