**INTRODUCTION**

The global issue of climate change has brought to the forefront the need to understand the complex interplay between environmental factors and human health. In this portfolio project, we embark on a journey to investigate the impact of CO2 emissions on life expectancy in the West African region. This analysis is motivated by the pressing concerns surrounding climate change and its potential consequences on public health in vulnerable areas.

Project Objectives

Our primary objectives in this project are as follows:

Assess the Relationship: Explore the relationship between CO2 emissions and life expectancy in West Africa through data-driven analysis.

Data Analysis: Utilize regression analysis techniques to quantify the impact of CO2 emissions on life expectancy.

Informative Insights: Provide valuable insights into the potential implications of environmental policies and initiatives on public health outcomes.

Rationale

West Africa, with its diverse socio-economic landscape and varying levels of industrialization, presents an intriguing case study for examining the connection between CO2 emissions and life expectancy. By understanding this relationship, we aim to contribute to the ongoing discourse on sustainable development and environmental stewardship in the region.

Methodology

To achieve our objectives, we will employ regression analysis techniques, specifically focusing on linear regression models. This statistical approach will allow us to model the relationship between CO2 emissions and life expectancy, accounting for potential confounding factors.

Tools and Software

Our analysis will be conducted using Stata, a powerful statistical software package widely used for data analysis and visualization. Stata provides a comprehensive suite of tools for regression analysis, making it well-suited for this project. We will use it to perform statistical tests, build regression models, and generate insightful visualizations.

Structure of the Project

This project is structured as follows:

Introduction: Providing an overview of the project, its objectives, and rationale (you are currently reading this section).

Data Collection: Explaining the data sources and preprocessing steps.

Regression Analysis: Introducing the statistical methods and models used, with a focus on how Stata is utilized.

Results and Interpretation: Presenting the outcomes and their significance.

Discussion: Offering insights and discussing potential implications and limitations.

Conclusion: Summarizing key findings and the project's contribution.

**DATA COLLECTION**

In this phase of our project, we acquired the essential data required to examine the impact of CO2 emissions on life expectancy in West Africa. We sourced our data from the World Bank, a globally recognized institution renowned for its comprehensive and reliable datasets. Specifically, we retrieved two primary datasets from the World Bank:

CO2 Emissions Data: We accessed annual CO2 emissions data for the West African countries of interest. This dataset provides metrics of CO2 emissions in metric tons for each country within the region, offering insights into the environmental aspect of our analysis.

Life Expectancy Data: Life expectancy data, a critical variable for our analysis, was also obtained from the World Bank for the West African countries under investigation. This dataset provides annual information on life expectancy at birth, enabling us to evaluate the health-related aspect of our analysis.

To ensure the data's suitability for rigorous analysis, we executed essential data preprocessing steps. These steps encompassed data cleaning to address missing values and inconsistencies, ensuring the integrity and completeness of our datasets. Additionally, we integrated the CO2 emissions and life expectancy datasets, aligning data points by year and country to facilitate the examination of the relationship between these two variables over time.

With our datasets thoroughly collected and preprocessed, they form the cornerstone of our analysis. In the following section, we will delve into the specifics of our regression analysis, performed using Stata, highlighting the chosen methodology and the steps taken to quantitatively assess the influence of CO2 emissions on life expectancy in West Africa

**REGRESSION ANALYSIS**

In this section, we embark on the core of our project, where we employ regression analysis techniques to investigate the impact of CO2 emissions on life expectancy in West Africa. We have selected Stata as our statistical software for this analysis, leveraging its robust capabilities for regression modeling.

Methodology:

To comprehensively explore the relationship between CO2 emissions and life expectancy, we have chosen to employ both fixed and random effects models. These models offer distinct advantages in addressing potential sources of bias and capturing unobserved heterogeneity across countries in our dataset.

Fixed Effects Model: The fixed effects model allows us to control for country-specific variations that are constant over time. By including fixed effects for each country in our regression, we can account for unobservable factors that may impact life expectancy but do not change over time.

Random Effects Model: The random effects model, on the other hand, takes into account both time-invariant and time-varying unobserved heterogeneity across countries. This model acknowledges that certain factors may vary over time but are still specific to individual countries.

Fixed-effects (within) regression

Number of obs = 320

Number of groups = 20

R-sq: Obs per group:

within = 0.1390 min = 16

between = 0.7572 avg = 16.0

overall = 0.0958 max = 16

F (1,299) = 48.25

corr (u\_i, Xb) = -0.0590 Prob > F = 0.0000

------------------------------------------------------------------------------

LIFEEXP | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

CO2 | -0.0000806 .0000116 -6.95 0.000 -0.0001034 -0.0000578

\_cons | 59.4701 .2965509 200.54 0.000 58.88651 60.05369

-------------+----------------------------------------------------------------

sigma\_u | 2.6664306

sigma\_e | 4.9334557

rho | 0.22607678 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

F test that all u\_i=0: F (19, 299) = 4.66 Prob > F = 0.0000

**INTERPRETATION OF FIXED EFFECT**

R-squared (R-sq):

Within R-squared: 0.1390,Between R-squared: 0.7572 and Overall R-squared: 0.0958

These R-squared values assess the goodness of fit of the model. "Within R-squared" represents the proportion of variation in the dependent variable (LIFEEXP) that is explained by the independent variables within each group (country). "Between R-squared" indicates the proportion of variation that is explained by the independent variables between groups (countries). The "Overall R-squared" combines both within and between effects.

F-Statistic: F (1,299) = 48.25

Prob > F = 0.0000 The F-statistic tests the overall significance of the regression model. In this case, it suggests that the model is statistically significant, as the p-value (Prob > F) is very close to zero.

Coefficient Estimates: CO2: -0.0000806 (Coefficient for CO2 emissions), \_cons (Intercept): 59.4701 (Intercept)

These coefficients represent the estimated impact of CO2 emissions on life expectancy within the fixed-effects model. In this context, the CO2 coefficient (-0.0000806) indicates that for each unit increase in CO2 emissions, life expectancy is estimated to decrease by approximately 0.0000806 years (or about 0.0294 days).

Standard Errors: Std. Err. (Standard Error) provides a measure of the precision of the coefficient estimates.

t-Statistic: The t-statistic measures the number of standard errors the coefficient estimate is away from zero. Larger absolute values of t indicate greater significance.

P-value (P>|t|): The p-value associated with each coefficient tests the null hypothesis that the coefficient is equal to zero. Smaller p-values suggest stronger evidence against the null hypothesis.

sigma\_u and sigma\_e: These values represent the estimated standard deviations of the individual-specific effects (sigma\_u) and the idiosyncratic errors (sigma\_e).

rho (rho): rho measures the fraction of the variance in the dependent variable (LIFEEXP) due to individual-specific effects (u\_i).

F-test for Group Effects: F test that all u\_i=0: F (19, 299) = 4.66

Prob > F = 0.0000

This F-test assesses whether there are significant differences in life expectancy between the different groups (countries). The low p-value suggests that there are significant group effects, indicating that individual countries have different life expectancy levels.

Random-effects GLS regression

Number of obs = 320

Number of groups = 20

R-sq: Obs per group:

within = 0.1390 min = 16

between = 0.7572 avg = 16.0

overall = 0.0958 max = 16

Wald chi2(1) = 34.88

corr (u\_i, X) = 0 (assumed) Prob > chi2 = 0.0000

------------------------------------------------------------------------------

LIFEEXP | Coef. Std. Err. z P>|z| [95% Conf. Interval]

-------------+----------------------------------------------------------------

CO2 | -0.0000747 0.0000127 -5.91 0.000 -.0000995 -.0000499

\_cons | 59.41495 0.337228 176.19 0.000 58.75399 60.0759

-------------+----------------------------------------------------------------

sigma\_u | .38319253

sigma\_e | 4.9334557

rho | 0.0059968 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

**INTERPRETATION OF RANDOM EFFECT**

R-squared (R-sq):

Within R-squared: 0.1390

Between R-squared: 0.7572

Overall R-squared: 0.0958

These R-squared values assess the goodness of fit of the model. Similar to the previous analysis, "Within R-squared" represents the proportion of variation in the dependent variable (LIFEEXP) explained by the independent variables within each group (country), while "Between R-squared" indicates the proportion of variation explained by the independent variables between groups (countries). The "Overall R-squared" combines both within and between effects.

Wald Chi-squared (chi2) Test:

Wald chi2(1) = 34.88

Prob > chi2 = 0.0000

The Wald chi-squared test assesses the overall significance of the regression model. In this case, it suggests that the model is statistically significant, as the p-value (Prob > chi2) is very close to zero.

Coefficient Estimates:

CO2: -0.0000747 (Coefficient for CO2 emissions)

\_cons (Intercept): 59.41495 (Intercept)

These coefficients represent the estimated impact of CO2 emissions on life expectancy within the random-effects GLS model. The CO2 coefficient (-0.0000747) indicates that for each unit increase in CO2 emissions, life expectancy is estimated to decrease by approximately 0.0000747 years (or about 0.0273 days).

Standard Errors:

Std. Err. (Standard Error) provides a measure of the precision of the coefficient estimates.

z-Statistic:

The z-statistic measures the number of standard errors the coefficient estimate is away from zero. Larger absolute values of z indicate greater significance.

P-value (P>|z|):

The p-value associated with each coefficient tests the null hypothesis that the coefficient is equal to zero. Smaller p-values suggest stronger evidence against the null hypothesis.

sigma\_u and sigma\_e:

sigma\_u represents the estimated standard deviation of the individual-specific effects (u\_i).

sigma\_e represents the estimated standard deviation of the idiosyncratic errors (sigma\_e).

rho (rho):

rho measures the fraction of the variance in the dependent variable (LIFEEXP) due to individual-specific effects (u\_i).

In the context of a random-effects GLS model, the key distinction lies in the treatment of individual-specific effects (u\_i). Unlike fixed-effects models, which account for individual-specific effects by including fixed effects for each group (country), random-effects models assume that these effects are uncorrelated with the independent variables and are distributed randomly.

**RESULTS AND INTERPRETATION**

In this section, we present the results of our regression analysis and provide an interpretation of the findings. Our analysis employed both fixed and random-effects models to assess the impact of CO2 emissions on life expectancy in West Africa. Here are the key results:

Fixed-Effects Model:

Within R-squared: 0.1390

Between R-squared: 0.7572

Overall R-squared: 0.0958

In the fixed-effects model, we observed that CO2 emissions were negatively associated with life expectancy. Specifically, for each unit increase in CO2 emissions, life expectancy was estimated to decrease by approximately 0.0000806 years (or about 0.0294 days). This result was statistically significant, as indicated by the low p-value (0.0000).

Random-Effects GLS Model:

Within R-squared: 0.1390

Between R-squared: 0.7572

Overall R-squared: 0.0958

In the random-effects GLS model, we also found a statistically significant negative association between CO2 emissions and life expectancy. The estimated coefficient for CO2 emissions indicated a decrease in life expectancy by approximately 0.0000747 years (or about 0.0273 days) for each unit increase in CO2 emissions.

Both models yielded consistent results, suggesting that higher CO2 emissions are associated with reduced life expectancy in West African countries. However, it's essential to consider the practical significance of these findings and potential confounding variables not included in the models.

Interpretation:

These results have important implications for public health and environmental policies in the West African region. They suggest that efforts to reduce CO2 emissions may have a positive impact on life expectancy by mitigating the adverse health effects associated with higher levels of pollution and climate change.

It's important to note that while the statistical relationship is significant, the practical impact of CO2 emissions on life expectancy may be relatively small in terms of years or days. Nonetheless, this analysis provides valuable insights into the complex interplay between environmental factors and public health.

**DISCUSSION**

In the "Discussion" section, we delve deeper into the implications of our findings regarding the impact of CO2 emissions on life expectancy in West Africa. This section aims to provide context, draw meaningful conclusions, and offer insights for policymakers, researchers, and stakeholders.

Implications of the Findings

Environmental Health: The consistent negative association between CO2 emissions and life expectancy underscores the critical importance of environmental factors in public health. Higher CO2 emissions are linked to increased air pollution, which can lead to respiratory illnesses and other health issues. Policymakers should consider these findings when formulating environmental and health policies.

Vulnerable Populations: Vulnerable populations in West Africa, such as those living in densely populated urban areas and near industrial centers, may be disproportionately affected by the health consequences of higher CO2 emissions. Targeted interventions and healthcare access improvements are crucial for these communities.

Climate Change Mitigation: Addressing CO2 emissions is not only essential for mitigating climate change but also for safeguarding public health. Our results emphasize the interconnectedness of environmental sustainability and well-being, reinforcing the need for comprehensive climate action plans.

**LIMITATIONS AND CONSIDERATIONS**

Causation vs. Correlation: While our analysis identifies a statistical association between CO2 emissions and life expectancy, it's essential to acknowledge that correlation does not imply causation. Other factors not included in our models, such as healthcare infrastructure, socio-economic status, and lifestyle factors, may contribute to the observed effects.

Data Limitations: Our analysis relies on data availability and quality. Data gaps or inconsistencies could affect the robustness of our results. Future studies should seek to address these limitations by utilizing more comprehensive datasets and exploring additional covariates.

Temporal Considerations: The impact of CO2 emissions on life expectancy may have long-term effects that extend beyond the scope of our analysis. Longitudinal studies could provide insights into the evolving relationship between environmental factors and public health outcomes.

**POLICY RECOMMENDATIONS**

Emissions Reduction Policies: Policymakers in West African countries should prioritize the development and implementation of policies aimed at reducing CO2 emissions. This includes promoting cleaner energy sources, improving public transportation, and enhancing industrial emission controls.

Healthcare Infrastructure: Investments in healthcare infrastructure, particularly in areas affected by high levels of pollution, can help mitigate the health consequences of environmental factors. Access to quality healthcare services is essential for addressing respiratory and other pollution-related illnesses.

Public Awareness: Public awareness campaigns can educate communities about the health risks associated with high CO2 emissions. Promoting sustainable practices and encouraging citizen engagement in environmental protection efforts can contribute to improved public health.

**FUTURE RESEARCH DIRECTIONS**

Causality Studies: Conduct further research to explore causal relationships between CO2 emissions and specific health outcomes, considering potential mediators and moderators.

Regional Analyses: Extend the analysis to sub-regions within West Africa to account for geographical variations in CO2 emissions and their health impacts.

Longitudinal Studies: Long-term studies tracking changes in CO2 emissions and life expectancy can provide a more comprehensive understanding of the relationship.

In conclusion, our analysis highlights the significance of addressing CO2 emissions not only for environmental sustainability but also for the health and well-being of West African populations. The findings presented here underscore the urgency of coordinated efforts to reduce emissions and protect public health in the face of ongoing climate change.