Interconnected Environmental Challenges: CO2 Emissions, Renewable Energy, and Air Pollution in the Context of UN Sustainability Goals

Jessie Gao, Lu Liu, Chunyi Xu, and Lauren Xu

Spring 2025

Contents

List of Tables

List of Figures

```r
knitr::opts_chunk$set(echo = TRUE)
# Set your working directory
library (here)
##Checking working directory is the project folder
here()

# Load your packages
library(readxl)
library (ggplot2)
library(tidyverse)
library (lubridate)
library(dplyr)
library(writexl)


# Set your ggplot theme
my.default.theme <- theme_classic(base_size = 15) +
  theme(plot.background = element_rect(color='white',fill='white'),
        plot.title = element_text(color='black', size = 15),
        axis.text = element_text(color = 'black',size = 10),
        panel.grid.minor = element_line(size = 0.5),
        panel.grid.major = element_line(size = 0.5),
        legend.background = element_rect(color='white', fill = 'white'),
        legend.position = "right",
        legend.title = element_text(color='black',size=10),
        legend.text = element_text(size = 10))

#Setting the theme as the default theme
theme_set(my.default.theme)

# Load your datasets
Sus_Env <- read_excel("/Users/xuchunyi/Desktop/GaoLuXuXu_ENV872_EDA_FinalProject/GaoLuXuXu_ENV872_EDA_F

#Viewing the structure of the dataset
dim (Sus_Env)
view(Sus_Env)
colnames(Sus_Env)

# Data Wrangling
##Removing unnecessary rows and replacing missing values with N/A
Sus_Env_Step1 <- Sus_Env %>%
  na.omit() %>%
  mutate(`2020` = na_if(`2020`, '..'))

##Changing the year column to character
Sus_Env_Step1 <- Sus_Env_Step1 %>%
  mutate(across(4:ncol(Sus_Env_Step1), as.character))

##Reorganizing the dataset and filtering Series Name to keep variables of interest (PM, CO2, and Renew)
Sus_Env_Processed <- Sus_Env_Step1 %>%
  pivot_longer(cols = -c("Country Code", "Country Name", "Series Name"),
               names_to = "Year",
               values_to = "Value") %>%
```

```r
  filter(`Series Name` %in% c("PM", "CO2", "Renew"))


##Checking the class of the variable Value, and the result is character
class (Sus_Env_Processed$Value)

##Converting the Value to a numeric variable
Sus_Env_Processed <- Sus_Env_Processed %>%
  mutate(Value = as.numeric(Value))

##Checking the class of the variable Value, and the result is numeric
class (Sus_Env_Processed$Value)

#Saving the processed dataset
output_file <- "/Users/xuchunyi/Desktop/GaoLuXuXu_ENV872_EDA_FinalProject/GaoLuXuXu_ENV872_EDA_FinalPro
write_xlsx(Sus_Env_Processed, path = output_file)
```

Rationale and Research Questions

The global climate crisis has intensified over the past decade, with rising temperatures, air pollution, and energy insecurity threatening the health and stability of communities worldwide. Carbon dioxide (CO_2) emissions from fossil fuel combustion are a primary driver of climate change, while exposure to fine particulate matter (PM2.5) has been strongly linked to adverse health outcomes such as respiratory and cardiovascular disease. Meanwhile, the transition to renewable energy represents a critical pathway for reducing both greenhouse gas emissions and harmful air pollutants. Given the importance of environmental indicators in evaluating national progress toward sustainability, this study aims to explore the interplay between CO_2 emissions, PM2.5 pollution, and renewable energy adoption in the context of global development goals.

We selected data from the World Bank's World Development Indicators, which offer a reliable, standardized, and accessible source for cross-country comparisons. The four countries in our study, which are Germany, the United States, China, and Colombia, were chosen to provide a balanced perspective between developed and developing economies. These nations vary significantly in their energy structures, industrial activities, environmental policies, and climate vulnerabilities. The three metrics— CO_2 emissions per capita, PM2.5 air pollution exposure, and renewable energy share in total consumption— were selected due to their relevance to key UN Sustainable Development Goals (SDGs 3, 7, and 13) and their interconnected impacts on environmental and human health.

To guide our analysis, we propose the following research questions:

1) What is the relationship between renewable energy adoption and CO_2 emissions across countries and over time?

2) Are there observable correlations between CO_2 emissions and PM2.5 pollution exposure?

3) Which factor better predicts changes in CO_2 emissions: renewable energy consumption, exposure to PM2.5 pollution, or a combination of both?

4) Which countries demonstrate best practices or policy approaches that may serve as models for others in addressing these environmental challenges?

6

Dataset Information

The data for this project were collected from the World Bank Databank, specifically from the World Development Indicators (WDI) dataset. These indicators are sourced from officially recognized international organizations and represent some of the most comprehensive and reliable global development data available. The WDI aggregates national, regional, and global statistics on topics related to environmental sustainability, public health, and economic growth.

The final dataset includes annual values for the following four countries: United States, Germany, China, and Colombia, covering four key indicators: (1) CO2 Emissions (metric tons per capita), (2) Renewable Energy Consumption (% of total final energy consumption), and (3) PM2.5 Pollution Exposure (mean annual exposure in micrograms per cubic meter).

These indicators were selected for their direct relevance to SDG 3 (Good Health and Well-being), SDG 7 (Affordable and Clean Energy), and SDG 13 (Climate Action). Together, these variables form the foundation for analyzing how environmental and energy infrastructure development correlate with greenhouse gas emissions.

By leveraging this dataset, this project can investigate the relationship between CO2 emissions and two environmental and public health variables—renewable energy consumption, and PM2.5 air pollution across two developed countries (Germany and the United States) and two developing countries (China and Colombia). The goal is to understand whether higher renewable energy adoption, cleaner air, and improved access to clean fuels are associated with lower carbon emissions in these national contexts from 2010 to 2020.

Table 1: Dataset Structure Summary

Variable	Description
Country Name	Name of the Country
Country Code	Three-letter Country Code
Series Name	Indicator Name
Year (2010-2020)	Values for the respective indicator for each year

Dataset Wrangling

Data preparation was carried out using the R programming language, with heavy use of the tidyverse, dplyr, and readxl packages. The data wrangling process involved importing the original World Bank Excel dataset for each of the three variables of four countries using readxl::read_excel(). Unnecessary rows were removed, and missing values were replaced with "N/A" using the na.omit() function and the na_if() function. Then, the year column was converted to a character variable to facilitate data analysis. The dataset was then reshaped to a single panel structure by consolidating year columns into a single "Year" column, reorganizing the corresponding values into a "Value" column, and maintaining key identifiers like "Country Code," "Country Name," and "Series Name" via the pivot_longer () function. Subsequently, the "Value" column was converted to a numeric variable to ensure it is a format suitable for analysis. Finally, the processed dataset was saved as an Excel file in the Processed_Data folder using the write_xlsx() function. In this way, the final dataset was reshaped into a tidy format (long or wide as needed) for visualization and statistical analysis.

Exploratory Analysis

The following summary statistic table provides an overview of the three indicators from 2010 to 2020, as part of the data exploration.

Table 2: Summary Statistics for CO2, PM, and Renew

Series Name	Observations	Mean	Max	Min	SD	Range
CO2	44	8.30	17.43	1.43	5.05	16.00
PM	44	24.95	63.29	7.38	18.56	55.91
Renew	44	16.88	32.38	7.44	8.49	24.94

The following three summary statistic tables provide an overview of the three indicators for each country from 2010 to 2020, as part of the data exploration.

Table 3: Summary Statistics for CO2 Emissions (metric tons per capita)

Country Name	Observations	Mean	Max	Min	SD	Range
China	11	7.21	7.76	6.34	0.39	1.42
Colombia	11	1.61	1.75	1.43	0.10	0.32
Germany	11	8.88	9.62	7.26	0.72	2.37
United States	11	15.49	17.43	13.03	1.15	4.40

Table 4: Summary Statistics for Renewable Energy Consumption (% of total final energy consumption)

| Country Name	Observations	Mean	Max	Min	SD	Range
China	11	12.65	14.81	11.34	1.18	3.47
Colombia	11	30.85	32.38	28.74	1.31	3.64
Germany	11	14.65	18.60	11.61	2.00	6.99
United States	11	9.36	11.16	7.44	1.03	3.72

8

Table 5: Summary Statistics for PM2.5 Pollution Exposure (mean annual exposure in micrograms per cubic meter)

Country Name	Observations	Mean	Max	Min	SD	Range
China	11	54.88	63.29	47.67	5.28	15.61
Colombia	11	23.44	25.95	22.01	1.35	3.94
Germany	11	13.14	16.09	11.78	1.60	4.31
United States	11	8.32	9.26	7.38	0.67	1.88

The exploratory analysis of CO2 emissions, renewable energy usage, and PM2.5 exposure data involved initial visualizations of each indicator across four countries (United States, Germany, China, and Colombia) from 2010-2020. After preliminary graphs revealed scaling challenges due to wide variations between nations, the y-axis was adjusted to better display trends while preserving comparisons. Examining differences in each indicator's levels, trajectories, and volatility helped identify potential patterns worthy of deeper investigation in subsequent analysis. In this way, this section could examine how CO2 emissions, renewable energy usage, and PM2.5 exposure have evolved from 2010 to 2020 in Germany, the United States, China, and Colombia.
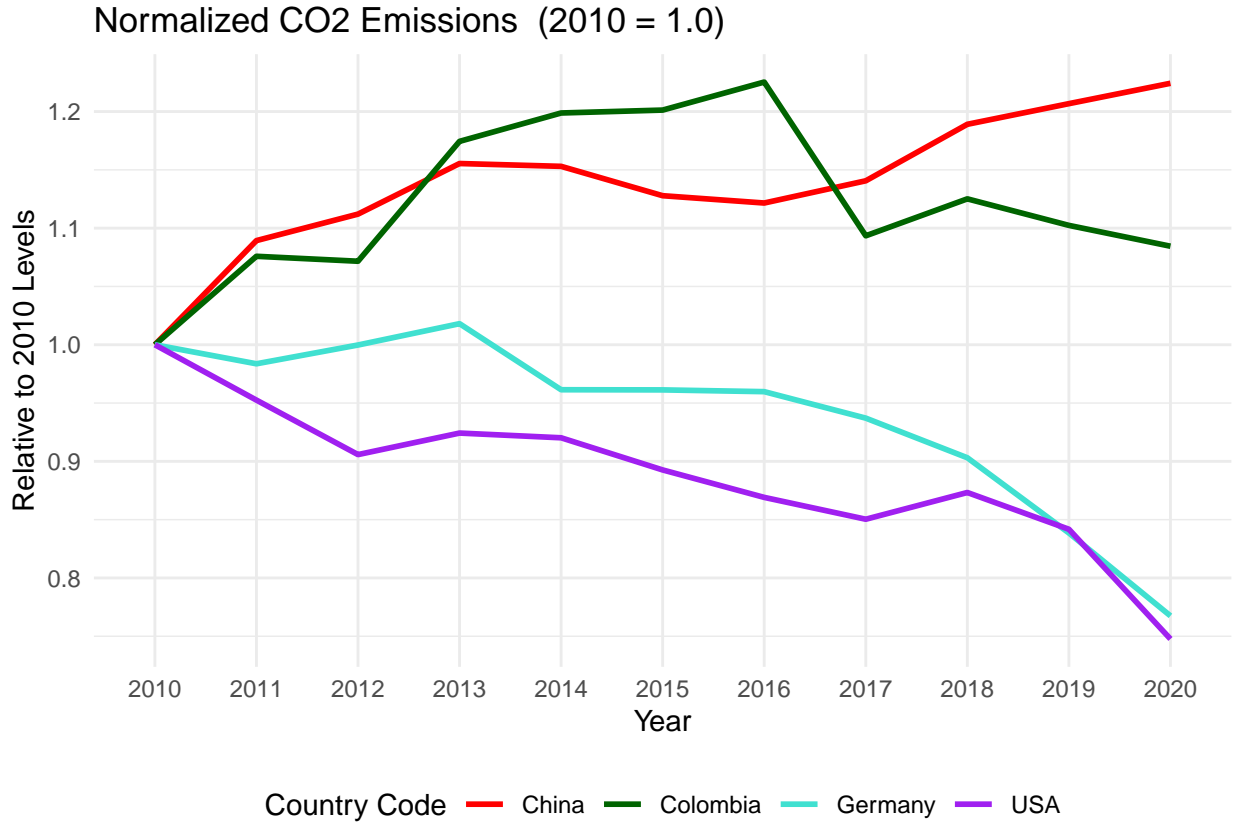
Between 2010 and 2020, Germany and the United States exhibited a downward trend in CO2 emissions, whereas China showed an upward trend. Colombia experienced an increase in emissions from 2010 to 2016, followed by a decline in the subsequent years. This normalized graph highlights that while absolute emissions varied greatly, throughout 2010 to 2020: - The U.S. achieved the largest relative reduction (-25%) - China's emissions grew the most (+22%) - Colombia remained stable (+0.08%) - Germany showed a moderate decline (-23%)

Normalized Renewable Energy Consumption (2010 = 1.0)
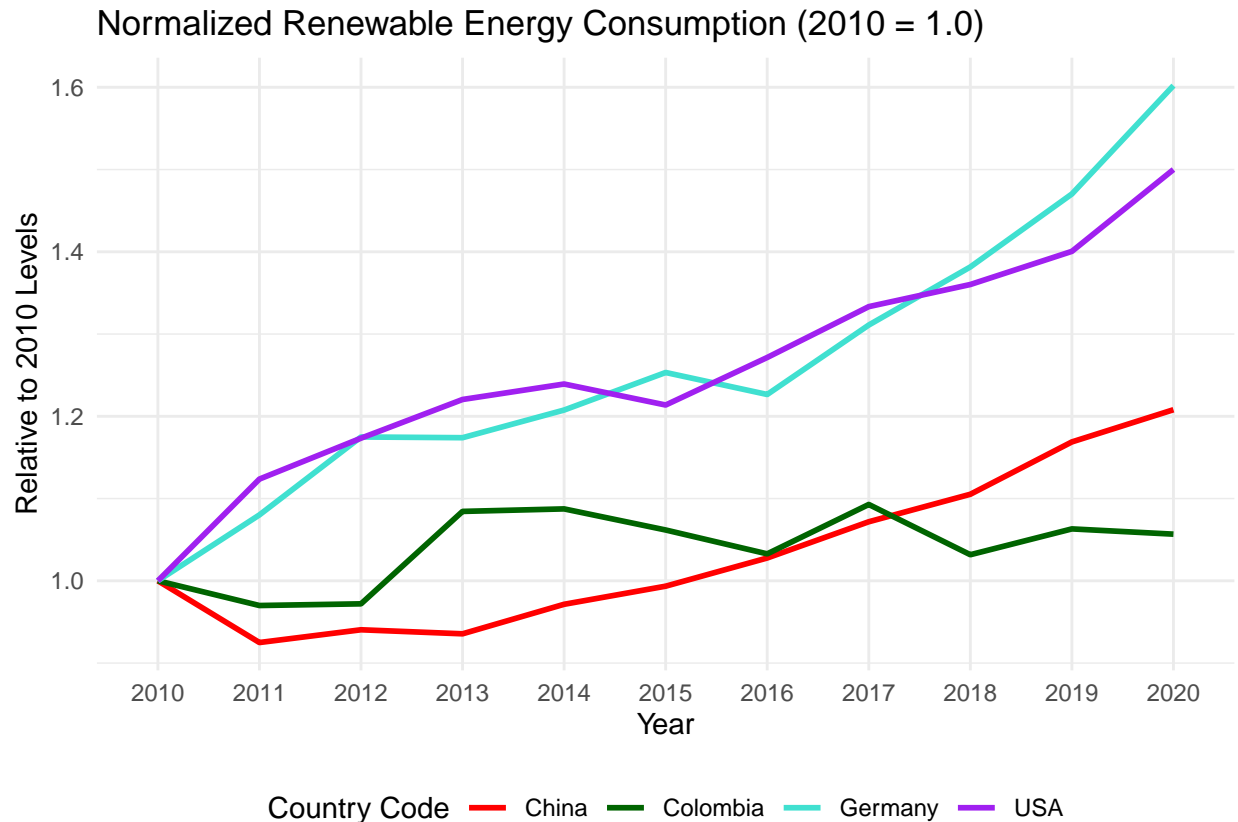
From 2010 to 2020, China, Germany, and the United States all demonstrated a notable upward trend in renewable energy consumption, while Colombia exhibited a modest overall increase. This normalized graph highlights that while absolute renewable energy consumption level varied greatly: - Germany achieved the largest relative increase (+60%) - The U.S. achieved an evident relative increase, too (+50%) - China showed a moderate increase (+21%) - Colombia remained stable (+0.06%)

```
## Warning: Removed 4 rows containing missing values or values outside the scale range
## ('geom_line()').
```

## Normalized PM2.5 Pollution Exposure (2010 = 1.0)
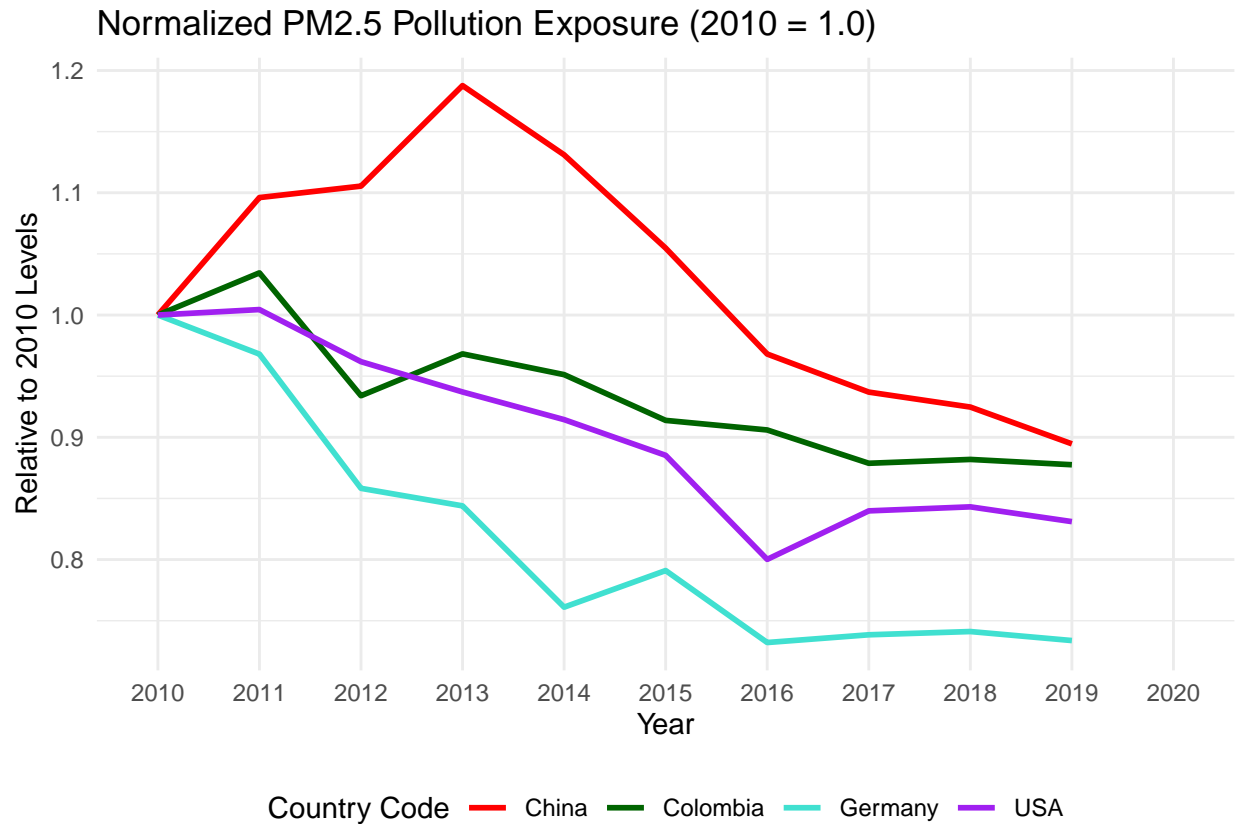
Between 2010 and 2020, all four countries exhibited a significant decline in PM2.5 pollution exposure, although China experienced an initial increase from 2010 to 2013. This normalized graph highlights that while absolute PM2.5 pollution exposure level varied greatly: - Germany achieved the largest relative reduction (-27%) - The U.S. achieved evident relative reduction, too (-17%) - China and Colombia both showed a moderate reduction (-11% and -12%, respectively)

These preliminary visualizations suggest fundamentally different CO2 emissions, renewable energy usage, and PM2.5 exposure trajectories across countries, which the subsequent analysis will examine in greater depth through correlation tests, statistical comparisons, and policy context.

Analysis

Question 1: What is the relationship between renewable energy adoption and CO2 emissions across countries and over time?

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```


Cross Country Comparison of Renewable Energy Trends
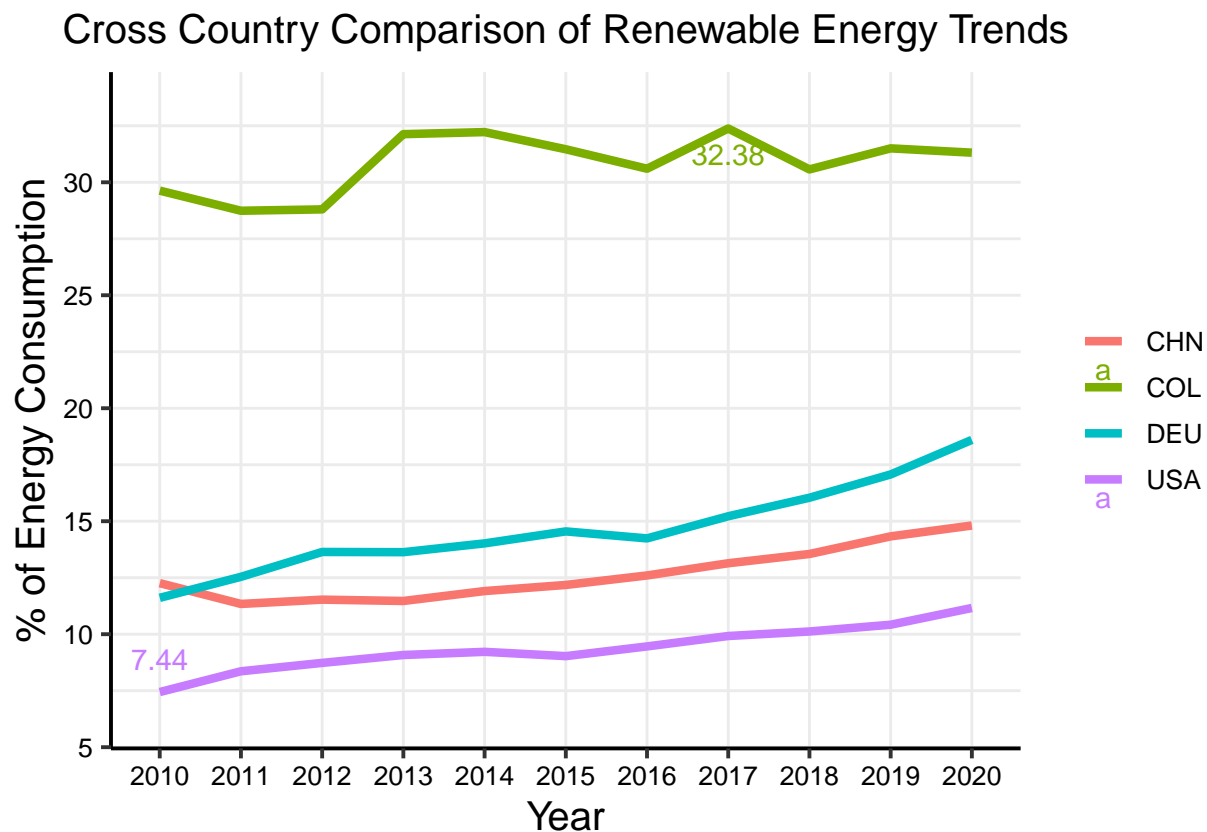
Overview of Renewable Energy Consumption The widespread adoption of renewable energy is essential to mitigating the impacts of climate change caused by human activity. Renewable energy sources, such as solar, wind, geothermal, and hydropower, hold the potential to transform global energy systems, particularly by providing electricity to the **760 million people** who currently lack access to it. Additionally, replacing fossil fuels like coal, oil, and natural gas with renewable energy is crucial for reducing greenhouse gas emissions. Although countries worldwide have made some progress in adopting renewable energy, the pace remains insufficient. In 2020, renewables accounted for just '12.6% of global energy consumption'. This rate of adoption is far from adequate to meet climate targets. To effectively limit global warming, it is necessary to 'triple renewable energy capacity by 2030'. Our analysis reveals that some nations have made more significant strides toward achieving this goal than others. The World Bank tracks renewable energy use as a percentage of total energy consumption. In 2020, the average renewable energy share across the four countries in our study was **16.88%**. Among these countries, **Colombia** stands out as a leader, reaching a peak of **32% renewable energy** in 2017. **Germany** follows, with impressive but still insufficient renewable energy adoption compared to Colombia. In stark contrast, **the United States** has been much slower to

transition to renewables, with a peak share of just **11% in 2020**, indicating substantial underperformance relative to global needs.

The Correlation Analysis

The correlation analysis reveals no universal relationship between renewable energy adoption and CO_2 emissions. While China shows a slight negative trend (-0.15), other countries (Colombia, USA, Germany) exhibit positive or neutral correlations (0.52, 0.24, and 0.12 respectively). The results suggest that renewables alone are insufficient to cut emissions without explicit fossil fuel phase-outs. This underscores that national context—energy mix, policy, and economic growth—drives whether renewables effectively mitigate CO_2.

The Linear Regression Analysis

```
##
## Call:
## lm(formula = CO2_COL$Value ~ Renew_COL$Value)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.12821 -0.06409  0.01401  0.03909  0.15624
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.39776    0.66444   0.599    0.564
## Renew_COL$Value  0.03921    0.02152   1.822    0.102
##
## Residual standard error: 0.08911 on 9 degrees of freedom
## Multiple R-squared:  0.2694, Adjusted R-squared:  0.1883
## F-statistic: 3.319 on 1 and 9 DF,  p-value: 0.1018


##
## Call:
## lm(formula = CO2_CHN$Value ~ Renew_CHN$Value)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.79198 -0.05934  0.07246  0.10398  0.36280
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.48841    1.06516   4.214  0.00226 **
## Renew_CHN$Value  0.21525    0.08389   2.566  0.03040 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3119 on 9 degrees of freedom
## Multiple R-squared:  0.4225, Adjusted R-squared:  0.3583
## F-statistic: 6.583 on 1 and 9 DF,  p-value: 0.0304


##
## Call:
## lm(formula = CO2_USA$Value ~ Renew_USA$Value)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -0.54899 -0.25971 -0.07488  0.31369  0.53670
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       25.4285     1.0742  23.672 2.04e-09 ***
## Renew_USA$Value   -1.0615     0.1142  -9.298 6.53e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3709 on 9 degrees of freedom
## Multiple R-squared:  0.9057, Adjusted R-squared:  0.8952
## F-statistic: 86.46 on 1 and 9 DF,  p-value: 6.534e-06


##
## Call:
## lm(formula = CO2_DEU$Value ~ Renew_DEU$Value)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.44241 -0.21300  0.05771  0.17353  0.40474
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      13.78287    0.64425  21.394 5.01e-09 ***
## Renew_DEU$Value  -0.33480    0.04361  -7.678 3.07e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2759 on 9 degrees of freedom
## Multiple R-squared:  0.8676, Adjusted R-squared:  0.8528
## F-statistic: 58.95 on 1 and 9 DF,  p-value: 3.069e-05


##
## Call:
## lm(formula = merged_data$Value_CO2 ~ merged_data$Value_Renew)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.3552 -1.5838 -0.2625  1.4720  4.2426
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)             17.04590    0.85031   20.05  < 2e-16 ***
## merged_data$Value_Renew -0.51838    0.04511  -11.49 1.51e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.512 on 42 degrees of freedom
## Multiple R-squared:  0.7587, Adjusted R-squared:  0.7529
## F-statistic: 132.1 on 1 and 42 DF,  p-value: 1.508e-14


##                          Country Coefficient Direction  p_value Significant
## Renew_COL$Value         Colombia       0.039         ↑ 1.02e-01          No
```

```
## Renew_CHN$Value         China      0.215    ↑ 3.04e-02     Yes
## Renew_USA$Value           USA     -1.062    ↓ 6.53e-06     Yes
## Renew_DEU$Value       Germany     -0.335    ↓ 3.07e-05     Yes
## merged_data$Value_Renew  Overall  -0.518    ↓ 1.51e-14     Yes
##                        R_squared
## Renew_COL$Value            0.27
## Renew_CHN$Value            0.42
## Renew_USA$Value            0.91
## Renew_DEU$Value            0.87
## merged_data$Value_Renew    0.76
```

Results: - Colombia: It is predicted 0.03921 metric tons per capita CO_2 emissions will increase for every 1% increase in the percentage of renewable energy consumption of total final energy consumption (df = 9, F = 3.319., p = 0.1018). The p-value is 0.1018, suggesting this result is statistically insignificant. The adjusted R-squared is 0.1883, it suggests 18.83% of the variability in CO_2 emissions is explained by changes in renewable energy consumption. - China: It is predicted 0.21525 metric tons per capita CO_2 emissions will increase for every 1% increase in the percentage of renewable energy consumption of total final energy consumption (df = 9, F = 6.583., p = 0.0304). The p-value is 0.0304 and less than 0.05, suggesting this result is statistically significant. The adjusted R-squared is 0.3583, it suggests 35.83% of the variability in CO_2 emissions is explained by changes in renewable energy consumption. - The U.S.: It is predicted 1.0615 metric tons per capita CO_2 emissions will decrease for every 1% increase in the percentage of renewable energy consumption of total final energy consumption (df = 9, F = 86.46., p = 6.534e-06). The p-value is 6.534e-06 and less than 0.05, suggesting this result is statistically significant. The adjusted R-squared is 0.8952, it suggests 89.52% of the variability in CO_2 emissions is explained by changes in renewable energy consumption. - Germany: It is predicted 0.33480 metric tons per capita CO_2 emissions will decrease for every 1% increase in the percentage of renewable energy consumption of total final energy consumption (df = 9, F = 58.95, p = 3.069e-05). The p-value is 3.069e-05 and less than 0.05, suggesting this result is statistically significant. The adjusted R-squared is 0.8528, it suggests 85.28% of the variability in CO_2 emissions is explained by changes in renewable energy consumption. - All four countries: It is predicted 0.51838 metric tons per capita CO_2 emissions will decrease for every 1% increase in the percentage of renewable energy consumption of total final energy consumption (df = 42, F = 132.1, p = 1.508e-14). The p-value is 1.508e-14 and less than 0.05, suggesting this result is statistically significant. The adjusted R-squared is 0.7529, it suggests 75.29% of the variability in CO_2 emissions is explained by changes in renewable energy consumption. Analysis: The linear regression results suggest varying impacts of renewable energy adoption on CO_2 emissions across countries. In the United States and Germany, there is a strong and statistically significant negative relationship between the renewable energy index and CO_2 emissions, indicating that increased renewable energy usage has contributed meaningfully to emissions reductions. Specifically, renewable energy explains over 90% of the variation in CO_2 emissions in the U.S. and 87% in Germany, highlighting effective decarbonization strategies. In contrast, China's regression shows a significant but positive relationship, implying that despite expanding renewables, CO_2 emissions have continued to rise, likely due to sustained industrial growth and fossil fuel reliance. For Colombia, the relationship is positive but not statistically significant, suggesting that renewables may not yet play a major role in reducing emissions, possibly due to their smaller share in the energy mix or concurrent growth in energy demand. These findings underscore the importance of national context in shaping the emissions impact of renewable energy development. However, across all four countries, the linear regression analysis reveals a statistically significant negative relationship between the adoption of renewable energy and CO_2 emissions.

Scatterplot with Regression Line

```
## `geom_smooth()` using formula = 'y ~ x'
```

Renewable Energy Adoption vs. CO2 Emissions (2010–2020)
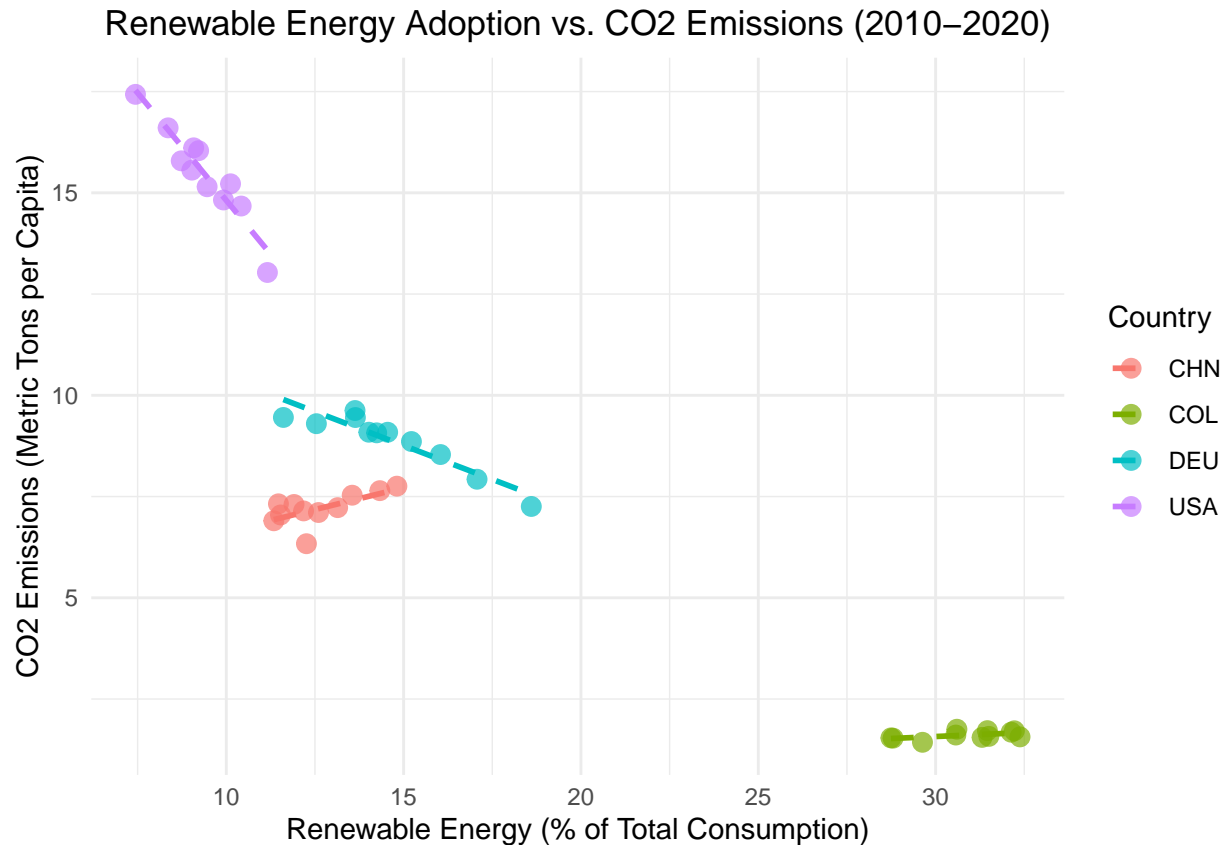
Based on the graph showing renewable energy adoption versus CO2 emissions across the USA, Germany, China, and Colombia from 2010-2020, there's a generally negative correlation between the percentage of renewable energy and carbon emissions per capita. Countries with higher renewable energy adoption tend to have lower per capita CO2 emissions, as demonstrated by Colombia having the highest renewable percentage (28-30%) and lowest emissions (about 2 metric tons per capita), while the USA shows the opposite pattern with low renewable adoption (8-12%) and high emissions (15-18 metric tons). The downward trend lines for individual countries (particularly visible for the USA and Germany) further suggest that as nations increase their renewable energy share over time, their emissions decrease accordingly, though other factors like overall energy consumption, economic development stage, and industrial activity also influence emission levels, as evidenced by China's somewhat different pattern.

Question 2: Are there observable correlations between CO2 emissions and PM2.5 pollution exposure?

```
## Warning: Removed 4 rows containing missing values or values outside the scale range
## ('geom_line()').
```

Cross Country Comparison of PM2.5 Air Pollution
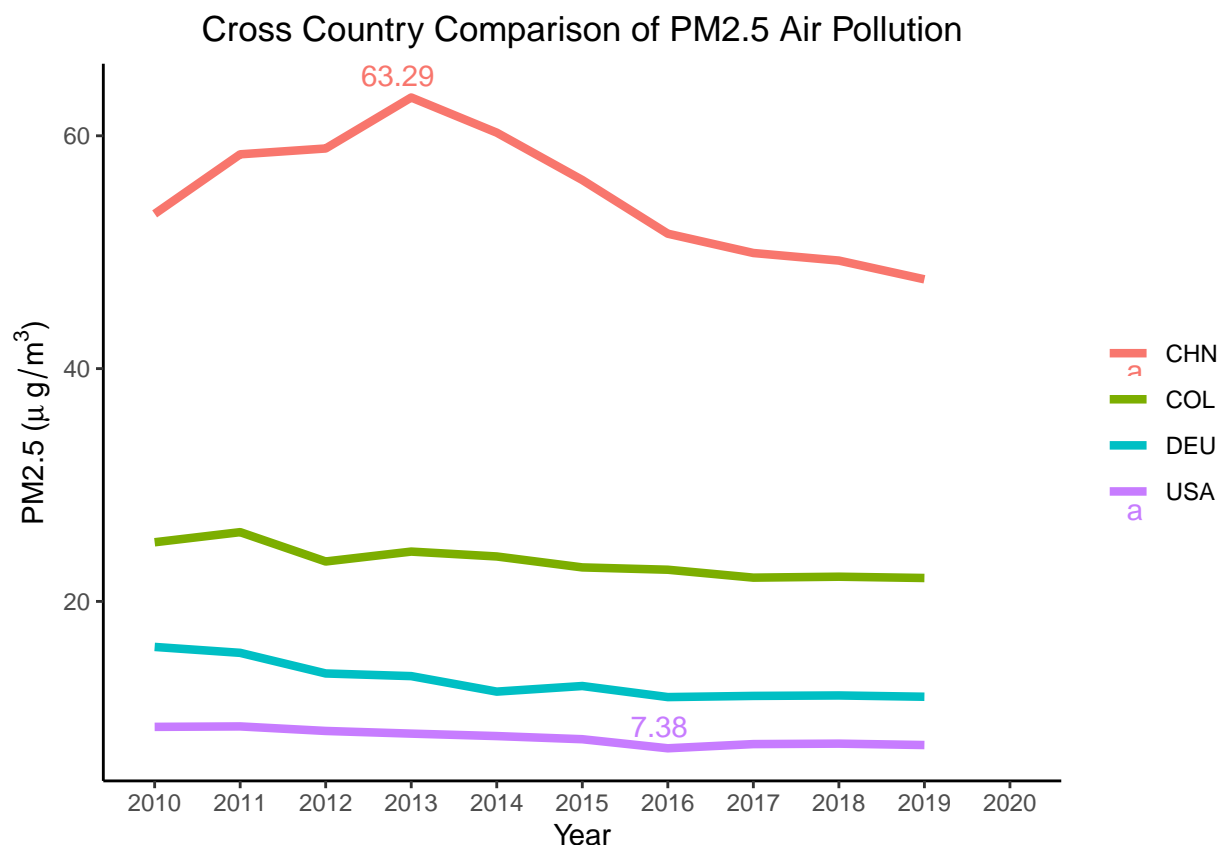
Overview of PM2.5 The proportion of a nation's population exposed to ambient PM2.5 concentrations above the WHO's recommended threshold of 10 micrograms per cubic meter serves as a critical indicator of air quality-related health risks. These fine particulate pollutants, primarily generated through combustion processes including vehicular emissions, industrial operations, and residential heating, demonstrate well-established dose-response relationships with adverse health outcomes. Epidemiological research confirms both acute effects such as increased hospitalizations for cardiorespiratory conditions and emergency room visits, as well as chronic consequences including reduced pulmonary function and elevated cardiovascular disease risk, with particular severity among pediatric and geriatric populations.

Our cross-national comparison reveals pronounced disparities in exposure levels, with China exhibiting systematically higher PM2.5 concentrations than the United States, Germany, and Colombia. This pattern likely reflects China's unique combination of coal-dependent energy infrastructure (accounting for approximately 60% of primary energy consumption), rapid industrial expansion during the 2010-2020 period averaging 8% annual GDP growth, and meteorological conditions prone to atmospheric stagnation. The United States demonstrates measurable improvements post-2010 attributable to Clean Air Act enforcement and fuel switching initiatives, while Germany maintains stable reductions despite complexities in its energy transition. Colombia's comparatively favorable profile correlates with its hydroelectric-dominated energy matrix, where renewables constitute over 70% of generation capacity.

These findings underscore the effectiveness of integrated policy approaches combining emission controls, energy infrastructure modernization, and urban planning interventions. The China-United States divergence particularly highlights the health benefits achievable through rigorous power plant emission standards, vehicle fleet renewal programs, and comprehensive air quality monitoring systems. Future research directions should prioritize quantitative health impact assessments using WHO-recommended methodologies, detailed source apportionment studies to identify sector-specific contributions, and longitudinal evaluations of intervention effectiveness, such as China's recent "Blue Sky" policy initiatives.

The Correlation Analysis

While observable correlations exist, their strength and direction vary significantly by national context. This underscores the need for country-specific air quality management strategies that account for local emission source profiles and development pathways. The strong US correlation (0.90) and moderate Germany correlation (0.63) suggests co-control benefits from decarbonization, whereas China/Colombia's results (-0.21 and -0.33, respectively) indicate PM2.5 reduction may require additional targeted measures beyond CO2 mitigation.

*Linear Regression Analysis

```
##
## Call:
## lm(formula = PM_COL$Value ~ CO2_COL$Value)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.61055 -1.09015 -0.07536  0.86313  2.18527
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     30.529      7.094   4.303   0.0026 **
## CO2_COL$Value   -4.395      4.391  -1.001   0.3462
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.35 on 8 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.1113, Adjusted R-squared:  0.0002086
## F-statistic: 1.002 on 1 and 8 DF,  p-value: 0.3462


##
## Call:
## lm(formula = PM_CHN$Value ~ CO2_CHN$Value)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -5.690 -4.363 -1.083  3.442  8.914
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     77.102     35.981   2.143   0.0645 .
## CO2_CHN$Value   -3.105      5.022  -0.618   0.5536
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.476 on 8 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.0456, Adjusted R-squared:  -0.0737
## F-statistic: 0.3823 on 1 and 8 DF,  p-value: 0.5536


##
## Call:
## lm(formula = PM_USA$Value ~ CO2_USA$Value)
##
## Residuals:
```

```
##      Min      1Q  Median      3Q      Max
## -0.5150 -0.1522  0.0194  0.1033  0.5194
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -2.9171     1.9563  -1.491 0.174258
## CO2_USA$Value   0.7136     0.1241   5.749 0.000429 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3156 on 8 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.8051, Adjusted R-squared:  0.7808
## F-statistic: 33.06 on 1 and 8 DF,  p-value: 0.0004293


##
## Call:
## lm(formula = PM_DEU$Value ~ CO2_DEU$Value)
##
## Residuals:
##     Min      1Q  Median      3Q      Max
## -1.4283 -0.8557 -0.3562  0.6415  2.1134
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -5.1006     7.8341  -0.651    0.533
## CO2_DEU$Value   2.0180     0.8654   2.332    0.048 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.306 on 8 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.4046, Adjusted R-squared:  0.3302
## F-statistic: 5.437 on 1 and 8 DF,  p-value: 0.04803


##
## Call:
## lm(formula = merged_pm_data$Value_PM ~ merged_pm_data$Value_CO2)
##
## Residuals:
##     Min      1Q  Median      3Q      Max
## -13.834 -11.702  -7.594   3.898  36.725
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)               37.6725     5.2151   7.224 1.23e-08 ***
## merged_pm_data$Value_CO2  -1.5174     0.5321  -2.852  0.00699 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17.07 on 38 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.1763, Adjusted R-squared:  0.1546
## F-statistic: 8.134 on 1 and 38 DF,  p-value: 0.00699
```

```
##                          Country Coefficient Direction  p_value Significant
## CO2_COL$Value            Colombia     -4.395        ↓ 0.346000          No
## CO2_CHN$Value               China     -3.105        ↓ 0.554000          No
## CO2_USA$Value                 USA      0.714        ↑ 0.000429         Yes
## CO2_DEU$Value             Germany      2.018        ↑ 0.048000         Yes
## merged_pm_data$Value_CO2  Overall     -1.517        ↓ 0.006990         Yes
##                          R_squared
## CO2_COL$Value                 0.11
## CO2_CHN$Value                 0.05
## CO2_USA$Value                 0.81
## CO2_DEU$Value                 0.40
## merged_pm_data$Value_CO2      0.18
```

Results: - Colombia: It is predicted 4.395 metric tons per capita CO_2 emissions will decrease for every 1 micrograms per cubic meter mean annual PM2.5 pollution exposure increase (df = 8, F = 1.002., p = 0.3462). The p-value is 0.3462, suggesting this result is statistically insignificant. The adjusted R-squared is 0.0002086, it suggests 0.02% of the variability in CO_2 emissions is explained by changes in PM2.5 pollution exposure. - China: It is predicted 3.105 metric tons per capita CO_2 emissions will decrease for every 1 micrograms per cubic meter mean annual PM2.5 pollution exposure increase (df = 8, F = 0.3823., p = 0.5536). The p-value is 0.5536, suggesting this result is statistically insignificant. The adjusted R-squared is -0.0737. - The U.S.: It is predicted 0.7136 metric tons per capita CO_2 emissions will increase for every 1 micrograms per cubic meter mean annual PM2.5 pollution exposure increase (df = 8, F = 33.06., p = 0.0004293). The p-value is 0.0004293 and less than 0.05, suggesting this result is statistically significant. The adjusted R-squared is 0.7808, it suggests 78.08% of the variability in CO_2 emissions is explained by changes in PM2.5 pollution exposure. - Germany: It is predicted 2.0180 metric tons per capita CO_2 emissions will increase for every 1 micrograms per cubic meter mean annual PM2.5 pollution exposure increase (df = 8, F = 5.437., p = 0.04803). The p-value is 0.04803 and less than 0.05, suggesting this result is statistically significant. The adjusted R-squared is 0.3302, it suggests 33.02% of the variability in CO_2 emissions is explained by changes in PM2.5 pollution exposure. - All four countries: It is predicted 1.5174 metric tons per capita CO_2 emissions will decrease for every 1 micrograms per cubic meter mean annual PM2.5 pollution exposure increase (df = 38, F = 8.134, p = 0.00699). The p-value is 0.00699 and less than 0.05, suggesting this result is statistically significant. The adjusted R-squared is 0.1546, it suggests 15.46% of the variability in CO_2 emissions is explained by changes in PM2.5 pollution exposure.

The second set of linear regression models investigates the relationship between PM2.5 concentrations and CO_2 emissions in each country and all four countries. Across all four countries, the linear regression analysis reveals a statistically significant negative relationship between the PM2.5 pollution exposure and CO_2 emissions, suggesting that higher CO_2 emissions are associated with lower PM2.5 concentrations, which might reflect the influence of decarbonization technologies or cleaner combustion. For the USA and Germany, both coefficients are positive and statistically significant, indicating that higher CO_2 emissions are associated with increased PM2.5 levels. However, the results for Colombia and China are not statistically significant, with p-values above 0.3 and very low R-squared values, implying little explanatory power of CO_2 emissions on PM2.5 in these contexts. Overall, the findings underscore heterogeneous national dynamics, possibly driven by differing energy sources, pollution control technologies, and industrial policies.
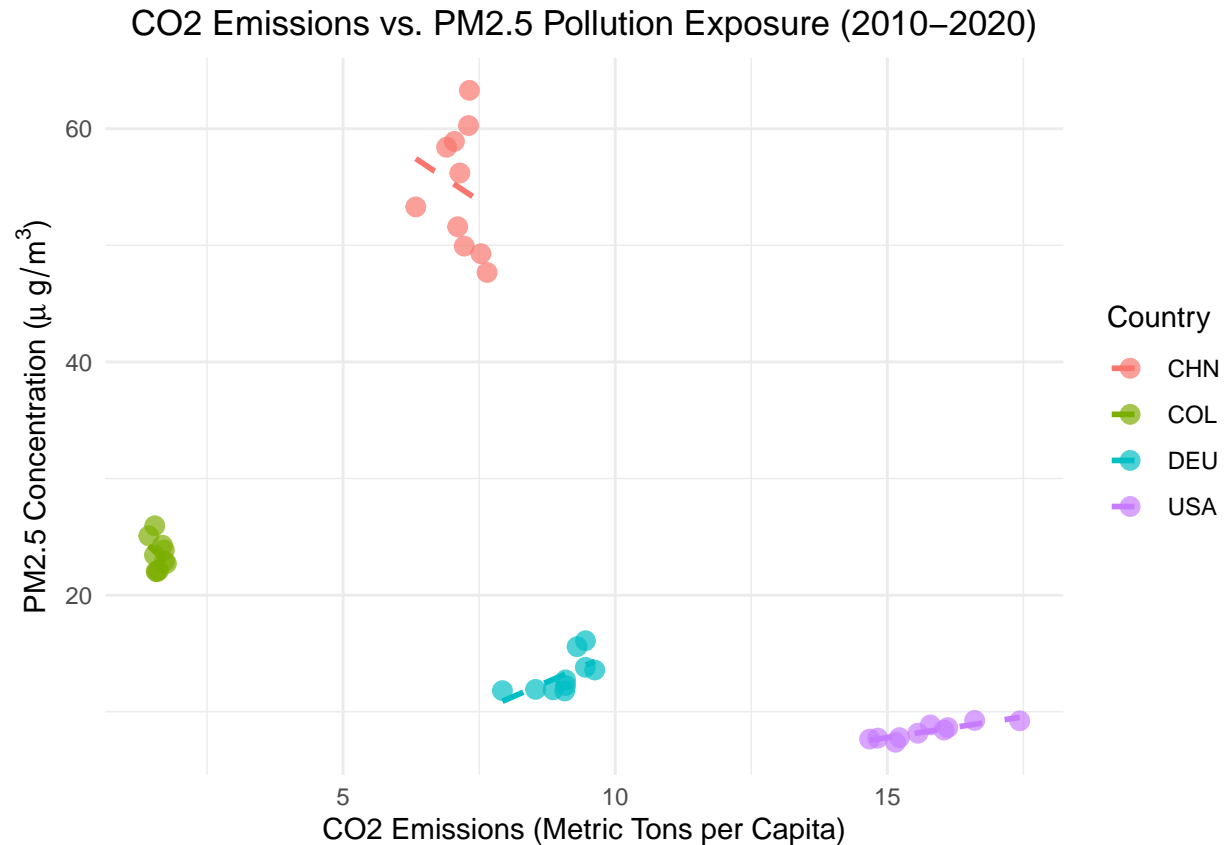
*Scatterplot with Regression Line

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 4 rows containing non-finite outside the scale range
## (`stat_smooth()`).
```

```
## Warning: Removed 4 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

CO2 Emissions vs. PM2.5 Pollution Exposure (2010–2020)

Looking at the scatter plot of CO2 emissions versus PM2.5 pollution exposure from 2010-2020 across China, Colombia, Germany, and the USA reveals complex relationships between these environmental indicators. While the overall cross-country comparison shows a counterintuitive pattern—with higher-emission countries like the USA having lower PM2.5 levels than lower-emission countries like China and Colombia—there are notable within-country correlations for developed economies. As observed in the data points for Germany and the USA, there appears to be a positive correlation where years with higher carbon emissions tend to correspond with higher particulate pollution levels. This suggests that while economic development stage and environmental regulation may disconnect these measures globally, within similar economic systems, both types of emissions may share common drivers or respond to similar policy interventions. The USA exhibits the clearest positive correlation, with PM2.5 concentrations increasing alongside rising per capita CO2 emissions, indicating that domestic pollution control efforts need to address both greenhouse gases and particulate matter simultaneously. This nuanced relationship highlights the importance of considering both global patterns and country-specific trends when developing environmental policies.

Question 3:Which factor better predicts changes in CO2 emissions: renewable energy consumption, exposure to PM2.5 pollution, or a combination of both?

```
##                         AIC
## All                  30.722
## Exclude Value_PM      78.688
## Exclude Value_Renew 126.144
```

Since the smaller the AIC value, the better. Across all four countries, the model that includes both renewable energy consumption and the exposure to PM2.5 pollution can better explain the changes in CO2 emissions. By running a multiple regression, it is predicted 0.11720 metric tons per capita CO2 emissions will decrease

for every 1 micrograms per cubic meter mean annual PM2.5 pollution exposure increase, and it is predicted 0.52076 metric tons per capita CO_2 emissions will decrease for every 1% increase in the percentage of renewable energy consumption of total final energy consumption (df = 37, F = 238, p = < 2.2e-16). The p-value is < 2.2e-16 and less than 0.05, suggesting this model is statistically significant. The adjusted R-squared is 0.924, it suggests 92.4% of the variability in CO_2 emissions is explained by changes in PM2.5 pollution exposure and renewable energy consumption.

```
##                       AIC
## All                -44.336
## Exclude Value_PM    -46.282
## Exclude Value_Renew -43.798
```

In Colombia, the model that includes only renewable energy consumption yields the lowest AIC result. Therefore, a multiple regression analysis is not conducted.

```
##                       AIC
## All                 26.221
## Exclude Value_PM    -20.483
## Exclude Value_Renew -17.762
```

In China, the model that includes both renewable energy consumption and the exposure to PM2.5 pollution can better explain the changes in CO_2 emissions. By running a multiple regression, it is predicted 0.10314 metric tons per capita CO_2 emissions will increase for every 1 micrograms per cubic meter mean annual PM2.5 pollution exposure increase, and it is predicted 0.69779 metric tons per capita CO_2 emissions will increase for every 1% increase in the percentage of renewable energy consumption of total final energy consumption (df = 7, F = 6.937, p = 0.02184). The p-value is 0.02184 and less than 0.05, suggesting this model is statistically significant. The adjusted R-squared is 0.5689, it suggests 56.89% of the variability in CO_2 emissions is explained by changes in PM2.5 pollution exposure and renewable energy consumption.

```
##                       AIC
## All                -22.954
## Exclude Value_PM    -22.077
## Exclude Value_Renew -16.715
```

In the United States, the model that includes both renewable energy consumption and the exposure to PM2.5 pollution can better explain the changes in CO_2 emissions. By running a multiple regression, it is predicted 0.4195 metric tons per capita CO_2 emissions will increase for every 1 micrograms per cubic meter mean annual PM2.5 pollution exposure increase, and it is predicted 0.6289 metric tons per capita CO_2 emissions will decrease for every 1% increase in the percentage of renewable energy consumption of total final energy consumption (df = 7, F = 37.44, p = 0.0001826). The p-value is 0.0001826 and less than 0.05, suggesting this model is statistically significant. The adjusted R-squared is 0.8901, it suggests 89.01% of the variability in CO_2 emissions is explained by changes in PM2.5 pollution exposure and renewable energy consumption.

```
##                       AIC
## All                -26.587
## Exclude Value_PM    -25.725
## Exclude Value_Renew -15.977
```

In Germany, the model that includes both renewable energy consumption and the exposure to PM2.5 pollution can better explain the changes in CO_2 emissions. By running a multiple regression, it is predicted 0.14587 metric tons per capita CO_2 emissions will decrease for every 1 micrograms per cubic meter mean annual PM2.5 pollution exposure increase, and it is predicted 0.40331 metric tons per capita CO_2 emissions

will decrease for every 1% increase in the percentage of renewable energy consumption of total final energy consumption (df = 7, F = 17.25, p = 0.001972). The p-value is 0.001972 and less than 0.05, suggesting this model is statistically significant. The adjusted R-squared is 0.7831, it suggests 78.31% of the variability in CO_2 emissions is explained by changes in PM2.5 pollution exposure and renewable energy consumption.

Question 4: Which countries demonstrate best practices or policy approaches that may serve as models for others in addressing these environmental challenges?

The analysis reveals distinct yet complementary policy approaches that countries can emulate to address interconnected environmental challenges. Colombia stands out as a renewable energy leader, achieving 32% renewable energy consumption through strategic harnessing of its natural resources. The country's success stems from innovative policies like the Energy Transition Law, which created tax incentives for renewable projects, and robust public-private partnerships exemplified by Ecopetrol's investments. Colombia's National Energy Plan demonstrates how developing nations can leverage geographic advantages - in this case, hydropower and bioenergy resources - to build sustainable energy systems. The planned expansion into wind and solar energy through competitive auctions in the Guajira region offers a replicable model for resource-rich developing countries.

The United States provides critical lessons in air quality management, maintaining the lowest PM2.5 levels among the studied nations. This achievement primarily results from decades of stringent regulation under the Clean Air Act and its amendments, which imposed strict emissions caps on industries and vehicles. While the correlation between CO_2 and PM2.5 reductions appears statistically weak, the U.S. experience shows how fuel switching and efficiency improvements can create valuable co-benefits for both climate and public health. America's regulatory framework offers a template for urbanizing nations struggling with particulate pollution, proving that emissions reductions aren't strictly prerequisite for air quality gains.

Germany exemplifies comprehensive decarbonization, successfully reducing CO_2 emissions by 14% while cutting PM2.5 exposure by nearly 30%. This dual achievement stems from Germany's integrated policy approach, combining aggressive renewable energy subsidies with a structured coal phaseout plan and meaningful carbon pricing (€25-30/ton). The German model demonstrates how developed economies can orchestrate just transitions through policy sequencing - first building renewable capacity, then retiring fossil assets - while maintaining energy security. This approach has created measurable climate and health benefits, making it particularly relevant for industrialized nations.

China's experience, while highlighting the challenges of coal dependence, also shows promising policy innovations. The "Blue Sky" initiative reduced PM2.5 concentrations by 11% through stringent industrial emission standards and pilot carbon trading schemes. For coal-reliant developing economies, China's approach suggests that targeted industrial policies and gradual market mechanisms can achieve meaningful improvements even before full energy transitions are feasible. Each country's experience contributes unique insights to a global toolkit for sustainable development, emphasizing that effective solutions must be tailored to national circumstances while drawing on cross-border lessons. The collective experience underscores that environmental progress requires policy packages combining regulation, incentives, and innovation, adapted to each nation's resources and development stage.

Summary and Conclusions

References