CO_2 Emissions and Impact of Renewable Electricity Generation Final Report

By: Manasa Mangipudi and Vijay Koduri

16:954:597:01 DATA WRANGLING

Professor Stevenson Bolivar-Atuesta

Date: May 9th, 2025

Executive Summary

As the global community faces the escalating consequences of climate change, there is an urgent need to evaluate the effectiveness of renewable energy transitions. This project set out to explore an important question in our sustainability policy: Are countries that invest more in renewable electricity generation seeing a meaningful reduction in carbon dioxide (CO_2) emissions? By analyzing national data from 2000 to 2022, we aimed to understand both the relationship between renewable energy share and CO_2 emissions per capita, and how this relationship has evolved in major economies over the last two decades.

The analysis draws on comprehensive datasets from **Our World in Data (OWID)**, a globally trusted platform created by researchers at the University of Oxford. We combined multiple datasets that included electricity generation by source (such as coal, gas, solar, wind, and hydro), CO_2 emissions data, and country level demographic and economic indicators like population and GDP. These datasets were chosen for their high quality, longitudinal coverage and their relevance to the study of energy transitions and environmental impact.

Our objective was twofold. First, we wanted to quantify the connection between the share of electricity generated from renewable sources and per capita CO_2 emissions. Second, we aimed to trace how this connection has shifted over time in selected high impact economies—specifically the United States, China, India, Germany, and Brazil. These countries were selected because they are either major emitters, leaders in renewable energy, or both.

Through a rigorous data wrangling and cleaning process in R, we created a merged dataset that included newly engineered variables such as **`co2_per_capita`** and **`co2_per_kwh`**. Exploratory Data Analysis (EDA) revealed notable trends: countries with higher renewable electricity shares tended to emit less CO_2 per person. This was supported by a linear regression analysis, which showed a statistically significant negative relationship between renewable share and emissions, even after accounting for GDP (Check graphs Appendix F).

At the country level, Germany and Brazil emerged as success stories, having both expanded renewable electricity and reduced or stabilized their per capita emissions. On the other hand, countries like China and India have increased their renewable capacity but continue to experience rising CO_2 emissions, likely due to the scale of their industrial growth and ongoing reliance on fossil fuels.

This project demonstrates that while renewable electricity adoption is associated with improved emissions outcomes, its impact can vary greatly depending on national context. The findings provide useful insights for policymakers, researchers, and sustainability advocates aiming to understand the real world implications of clean energy transitions.

Introduction

Are countries that invest more in renewable electricity generation seeing a meaningful reduction in carbon dioxide (CO_2) emissions? This project explores that question by analyzing global trends in energy generation and emissions across more than 100 countries from 2000 to 2022.

The motivation behind this analysis is rooted in the urgent global need to address climate change. As energy production is one of the largest contributors to CO_2 emissions, the shift from fossil fuels to renewable sources such as solar, wind, and hydroelectricity is a key strategy adopted by governments and international climate agreements. While many countries are making significant investments in clean energy, it remains essential to evaluate whether these efforts are actually delivering results in terms of emission reduction.

To examine this, we use data compiled by **Our World in Data (OWID)**, a widely respected open source platform that aggregates and standardizes data from organizations like the International Energy Agency (IEA), Global Carbon Project, and World Bank. The datasets provide annual country level information on:

- Electricity generation, categorized by source (e.g., coal, oil, gas, hydro, solar, wind)
- CO_2 emissions from fossil fuel combustion and industry
- Population and GDP figures, allowing for per capita and normalized comparisons

(For more information on the data source - check Appendix A)

Using this data, we calculate key derived metrics such as:

- **CO_2 emissions per capita** (`co2 / population`)
- **Electricity per capita** (`electricity_generation / population`)
- **CO_2 per unit of electricity generated** (`co2 / electricity_generation`)
- **Renewables share of electricity generation**

(Other created metrics can be found in Appendix B)

The project centers around two main research objectives. First, we assess whether countries with a greater share of renewable electricity tend to have lower CO_2 emissions per capita, using regression analysis and cross country comparisons. Second, we study how this relationship has evolved over the last two decades in five major economies (United States, China, India, Germany, and Brazil) which collectively account for a significant portion of global energy use and emissions.

By combining data wrangling, visualization, and modeling, we aim to provide a data driven perspective on the effectiveness of renewable energy transitions, and the degree to which clean electricity generation correlates with reduced national emissions. This analysis offers valuable insights for policy makers, climate researchers, and the broader public interested in the outcomes of global decarbonization efforts.

Data Wrangling & Cleaning

Our project combined two main datasets from **Our World in Data (OWID)**: one focusing on electricity generation by source, and the other on CO_2 emissions, GDP, and population figures. Together, these datasets spanned over 130 columns and more than 21,000 rows, offering broad global coverage from 1800 to 2022. However, to keep the analysis relevant and consistent across countries, we focused on data from **2000 to 2022**.

Initial Cleaning: Handling Missing and Irrelevant Data

The raw electricity dataset had several columns with either outdated or sparse data (many with over 60–90% missing values). To address this, we:

- Dropped all data before the year 2000.
- Removed columns with **more than 60% missing values**.
- Verified and found no duplicate rows.
- Retained the most informative energy related columns such as `electricity_generation`, `fossil_electricity`, `renewables_electricity`, `solar`, `wind`, and `hydro` generation.

On the emissions dataset, we removed irrelevant fields (e.g., cement related emissions) and retained columns like `co2`, `co2_growth_abs`, `co2_per_capita`, and fuel specific CO_2 breakdowns (coal, oil, gas). (Specifics on cleaning the data can be found in Appendix B)

Creating Derived Variables

To enable comparative and per capita analysis across countries, we created several new columns:

- **Normalized Metrics**:
  - `electricity_per_capita`, `renewables_per_capita`, `fossil_per_capita`
  - `electricity_per_gdp`, `gdp_per_electricity`
- **Log Transformed Metrics**:
  - `log_gdp`, `log_population`, `log_electricity` (to handle skewed distributions)
- **Energy Splits and Ratios**:
  - `fossil_to_renewable_ratio`, `fossil_share_elec`, `solar_share`, `wind_share`, `hydro_share`
- **Binary Classifications**:

- `high_renewable`: =1 if renewable share > 50%
- `transitioning`: =1 if renewable share > fossil shareDataWrangling_Group8

CO_2 Data set:

Data Wrangling Step	Rows	Columns	Missing Values
Raw Dataset	50191	79	53.11%
N/A Columns Drop	3,772	65	12.34%
New Columns Addition	3,772	69	12.25%

Energy Data set:

Data Wrangling Step	Rows	Columns	Missing Values
Raw Dataset	21812	130	67.32%
N/A Columns Drop	3795	66	36.68%
New Columns Addition	3795	31	0.18%

Merging the Datasets

The cleaned data sets were merged using a **left join on `country` and `year`**, with the electricity dataset (`energy_transformed`) as the left table and emissions as the right. This preserved all countries with valid energy data and joined only corresponding rows with available CO_2 and population information.

After merging the dataset, we created new variable : co2_per_kwh - which gives an insight over how much CO2 is getting emitted per Kilowatt Hour of energy, an important metric about how renewable energy is impacting the CO2 emissions

After the merge:

- **Final Row Count**: ~3,795 rows
- **Final Column Count**: 95 Columns. We added a `continent` column to support continent wise plots and summaries.

Summary of Cleaned Dataset

The final dataset enabled year over year analysis of:

- CO_2 emissions (total and per capita)
- Electricity generation volumes and types
- Shares of fossil vs renewable electricity
- Emissions per unit of energy
- Economic and demographic normalization

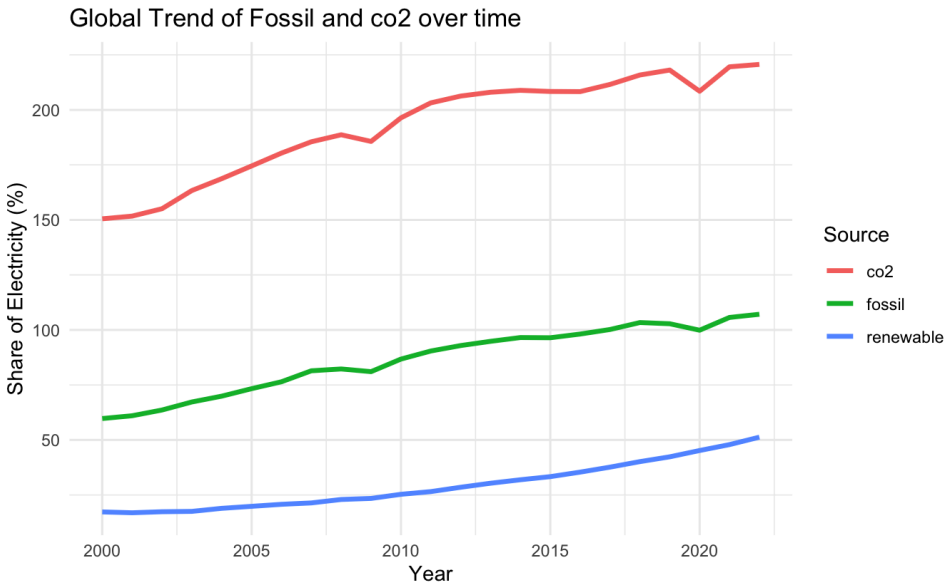
This cleaned and merged dataset provided the foundation for our visual exploration and regression based analysis of global emissions trends.

Exploratory Data Analysis:

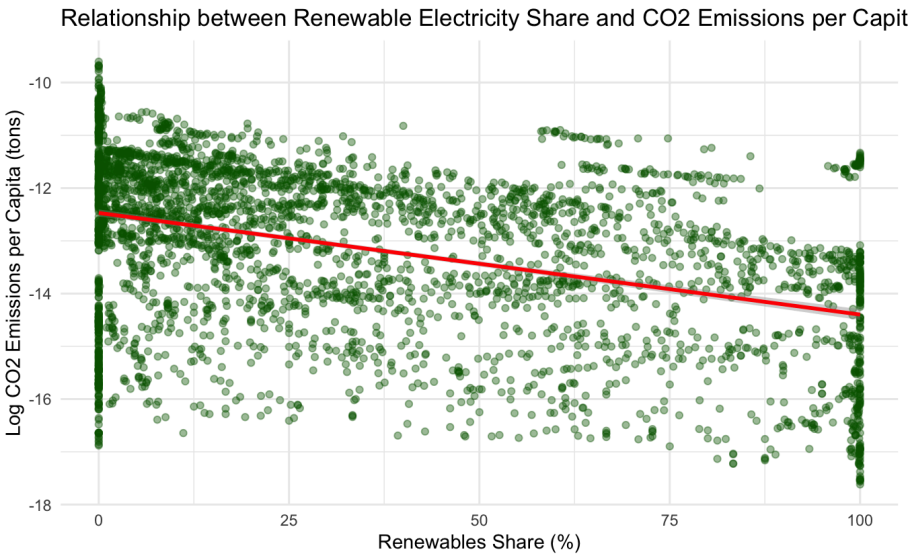
Are countries that invest more in renewable electricity generation seeing a meaningful reduction in carbon dioxide (CO_2) emissions?

To explore this question, we began by examining broad global and continental trends, then zoomed into key country level patterns, and finally evaluated emissions efficiency metrics like CO_2 per kWh. Each step in this analysis was designed to build insight into whether renewable electricity adoption leads to meaningful emission reductions.

To evaluate whether the world as a whole is transitioning to cleaner electricity, we first analyzed trends in renewable and fossil electricity generation alongside global CO_2 emissions over the past two decades.

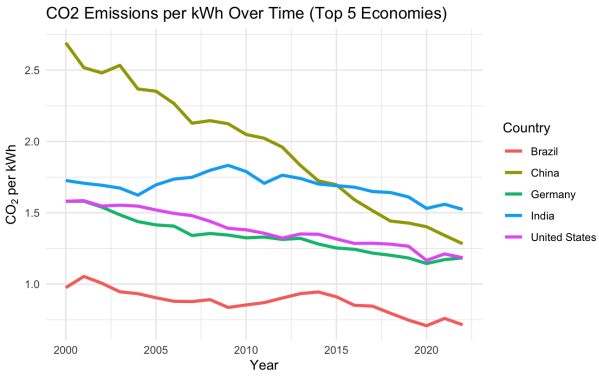


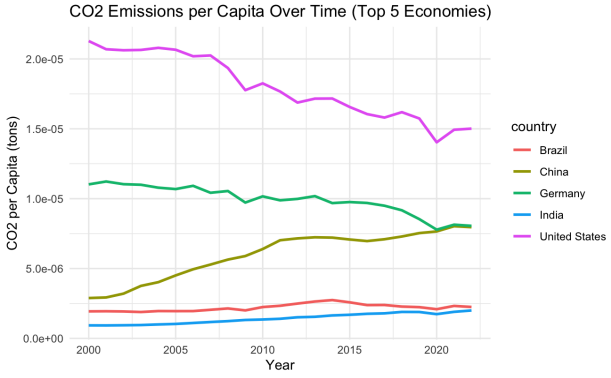
From this plot, we can see that, over the past two decades, renewable electricity generation has increased faster while fossil based generation and CO_2 emissions have begun to slow down. This suggests that the global energy system is slowly producing less and less CO_2.
Seeing this global trend, we next asked: **Are countries that use more renewable electricity also emitting less CO_2 per person?** To answer this, we visualized each country's renewable share and their CO_2 emissions per capita.



This reveals a negative relationship: countries with higher renewable electricity shares tend to emit less CO_2 per person. While not perfectly linear, the trend indicates that increasing your renewable electricity may contribute to lower emissions.

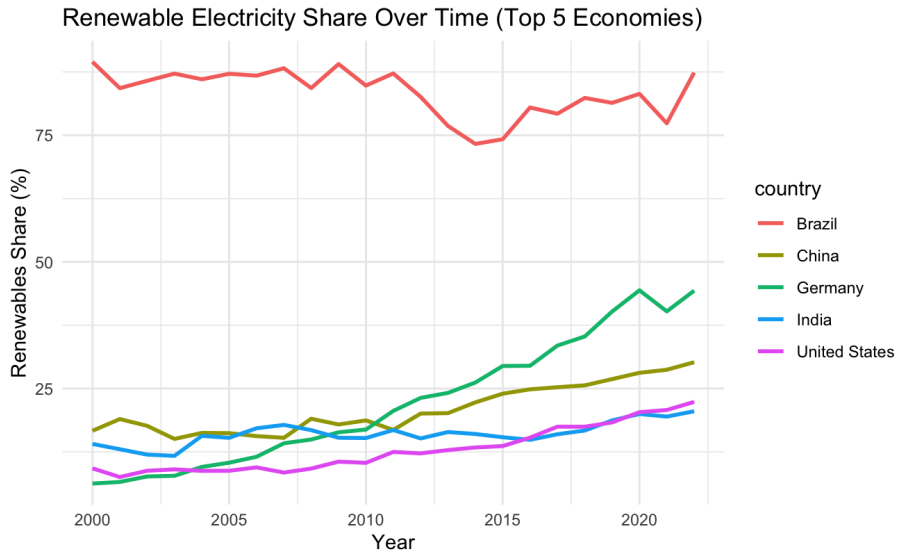
To evaluate how renewable growth and emissions intensity evolve over time, we focused on five major economies: the United States, China, India, Germany, and Brazil. These countries were chosen for their large populations, energy demands, and differing amounts of renewable energy generation.





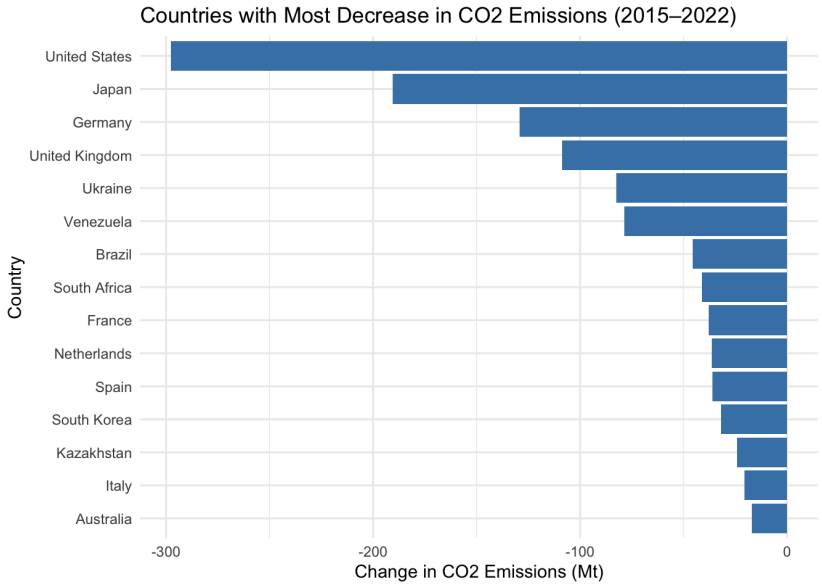
These plots track two dimensions of emissions in the United States, China, India, Germany, and Brazil from 2000 to 2022. The first shows trends in total CO_2 emissions per person, reflecting national carbon footprints. The second focuses on CO_2 emissions per unit of electricity (per kWh), which measures the carbon efficiency of power generation. Together, they reveal that Germany, Brazil and the United States have reduced both per capita emissions and emissions intensity, aligning with their renewable energy expansion. In contrast, China and India show continued growth in emissions per capita and per kWh, underscoring the challenges of balancing development with decarbonization

To support the trends seen above, we compared how each of these countries' renewable shares changed from 2000 to 2022.



Germany and Brazil show strong, steady growth in renewable share, consistent with their reduced emissions. India and China show improvement, but at a slower pace, highlighting the challenge of balancing energy access and decarbonization.

After analyzing long term emission trends, we zoom in on the years following the **2015 Paris Agreement**, when over 150 countries formally committed to reducing greenhouse gas emissions. The plot below shows the countries that achieved the largest reductions in absolute CO_2 emissions between 2015 and 2022.



The United States, Japan, and Germany recorded the largest declines in CO_2 emissions since 2015, with Germany and Brazil also appearing in our earlier analysis as countries that simultaneously increased renewable electricity. This reinforces the idea that renewable adoption and strong climate policy can lead to meaningful reductions in emissions.

Key Insights and Summary :

Summary of Findings

Our analysis, spanning over 100 countries from 2000 to 2022, revealed key relationships between renewable electricity generation and CO_2 emissions:

1. **Higher renewable electricity share correlates with lower CO_2 emissions**
    - Countries with a greater percentage of electricity generated from renewable sources tend to have lower emissions, both **per capita** and **per kilowatt hour**.
    - This correlation holds even when controlling for population and GDP by using normalized variables like **`co2_per_capita`** and **`co2_per_kwh`**.
2. **Some countries show clear decarbonization success stories**
    - **Germany and Brazil** stood out with simultaneous increases in renewable electricity and sustained reductions in emissions, suggesting that renewables, when deployed at scale, can drive meaningful progress.
    - The **United States** showed a modest but consistent decline in emissions after 2015, likely driven by both renewable investment and a shift from coal to natural gas.
3. **Rapidly developing economies still face challenges**

    - **China and India**, despite expanding their renewable capacity, have rising emissions, driven by increased industrial activity and total electricity demand.
    - This suggests that renewable growth alone is not enough unless accompanied by fossil fuel displacement and efficiency improvements.
4. **Post Paris Agreement reductions are real for some**
    - The majority of emission reductions between 2015 and 2022 came from countries with stronger climate commitments.
    - The **United States, Japan, and Germany** led absolute emission cuts, reinforcing that **policy commitment + renewable investment = results**.
5. **CO_2 per kWh is a high resolution diagnostic metric**
    - This metric captures how efficiently countries generate electricity, accounting for both generation mix and grid carbon intensity.
    - It offers a better understanding of decarbonization progress than raw CO_2 figures alone.

Business or Practical Implications

1. **Invest in renewable infrastructure, but also target fossil fuel phase out**
    - Simply adding renewables does not guarantee emission cuts unless fossil generation is actively displaced.
    - Countries must create **transition pathways** that retire coal and reduce natural gas dependence, particularly in energy intensive sectors.
2. **Make CO_2 per kWh a reporting standard**
    - Governments and utilities should adopt this as a KPI to track energy system efficiency.
    - Transparency on this metric can drive competition, encourage innovation, and support ESG (Environmental, Social, Governance) efforts.
3. **Support clean energy in high growth economies**
    - China and India's emission trajectories show the need for international financing, technology transfer, and policy support to accelerate clean energy scaling.
    - **Climate funds, green bonds, and multilateral climate financing mechanisms** should prioritize these economies.
4. **Pair renewable investment with grid modernization**
    - Without upgraded grids, high renewable input can lead to inefficiencies or curtailment.
    - Investment in **smart grids, storage, and flexible demand systems** is necessary to fully leverage renewable capacity.
5. **Use emissions trends post policy shifts (e.g., Paris Agreement) as benchmarks**
    - Countries that signed on to international climate accords and followed up with national action (e.g., Germany's coal exit plan) saw sharper CO_2 declines.
    - Emission progress should be measured in **multi year policy cycles**, not single year metrics.
6. **Incorporate decarbonization into national development strategies**

- For developing nations, emissions reductions need to align with economic growth.
- Programs that combine **job creation, energy access, and clean infrastructure** (e.g., solar microgrids, renewable manufacturing hubs) are vital for inclusive transitions.

Challenges Faced

1. Many countries had missing or sparse data, particularly for renewable energy sources in earlier years. We addressed this by restricting the analysis to 2000–2022 and dropping columns with excessive missingness.
2. Merging the electricity and CO_2 datasets required resolving inconsistencies in country names and handling rows with unmatched years.
3. Comparing countries of vastly different sizes introduced scale issues, which we mitigated by creating normalized metrics such as CO_2 per capita and CO_2 per kWh.
4. Log transformed plots were affected by zero or near zero values, which we handled by filtering out invalid data points or applying small adjustments.
5. A few countries showed extreme outliers in emissions or energy generation, which had to be visually checked and filtered to avoid skewing the plots.

Limitations and Next Steps

1. The analysis focused only on electricity related emissions and did not include other major sectors like transportation or industry.
2. The relationships we observed are correlational; we cannot conclusively say that renewable adoption directly caused emissions to fall.
3. Annual data limited our ability to detect shorter term changes or policy impacts; using monthly or quarterly data would improve resolution.
4. The dataset lacked variables on climate policies, carbon pricing, or renewable subsidies, which could provide better context for national trends.
5. In future work, we aim to integrate policy data, explore causal modeling, and develop a clean growth index to better assess equitable decarbonization.

Conclusion

This project set out to examine whether countries that invest more in renewable electricity generation are seeing a meaningful reduction in carbon dioxide (CO_2) emissions. We began by wrangling and cleaning two complementary datasets from Our World in Data, one covering electricity generation by source and the other focused on CO_2 emissions, GDP, and population. After filtering for the years 2000–2022 and engineering new variables such as **`co2_per_capita`**, and **`renewables_share_elec`**, we merged the datasets and created few other metrics based on both datasets like **`co2_per_kwh`**, which needs co2 details from 1st dataset and electricity generation from 2nd dataset and conducted exploratory analysis through time series trends, scatter plots, and regional comparisons.

The analysis led to several important insights. Countries with a higher share of renewable electricity generally emit less CO_2 per capita and per unit of electricity, with Germany and Brazil serving as strong examples of effective energy transition. The United States also showed steady post 2015 progress, while China and India revealed the complexities of decarbonizing under rapid economic growth. The Paris Agreement year (2015) provided a useful reference point to assess post policy emission reductions, and countries with clear policy action, like Germany, the U.S., and Japan, showed the largest absolute drops in emissions since then.

Looking forward, future work could expand the analysis to other emission sectors like transport and industry, integrate policy and investment data, and explore causal relationships between renewable adoption and emissions reduction. Creating a “clean growth” index that reflects countries achieving both economic expansion and emission control could also offer valuable insight for global climate strategy.

Ultimately, the findings reinforce that renewable electricity plays a crucial role in emissions reduction, but its full impact depends on national energy policy, fossil fuel phase out, and equitable infrastructure development.

References

1. Ritchie, H., Roser, M., & Rosado, P**.** (2023). *Energy*. Our World in Data. Retrieved from https://ourworldindata.org/energy
2. Ritchie, H., Roser, M., & Rosado, P. (2023). *CO_2 and Greenhouse Gas Emissions*. Our World in Data. Retrieved from https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions
3. United Nations Framework Convention on Climate Change (UNFCCC). (2015). *The Paris Agreement*. Retrieved from https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement
4. Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
5. Wickham, H. et al. (2023). *tidyverse: Easily Install and Load the Tidyverse*. R package version 2.0.0. URL: https://CRAN.R-project.org/package=tidyverse
6. Janitor package: Firke, B. (2023). *janitor: Simple Tools for Examining and Cleaning Dirty Data*. R package version 2.2.0.
7. naniar package: Tierney, N., Cook, D., McBain, M., & Fay, C. (2021). *naniar: Data Structures, Summaries, and Visualisations for Missing Data*. R package version 0.6-1.
8. skimr package: Waring, E., & Tierney, N. (2020). *skimr: Compact and Flexible Summaries of Data*. R package version 2.1.2.
9. Plotly Technologies Inc. (2015). *Collaborative data science*. Montreal, QC. https://plot.ly