Kalyan Pothineni
06/30/2024
DSC680-T301 (2247-1)
Applied Data Science

Analysis of CO2 Emissions by Sector: Identifying Patterns and Proposing Solutions

Kalyan Pothineni

Applied Data Science

Course Number: DSC680

06/30/2024

Analysis of CO2 Emissions by Sector: Identifying Patterns and Proposing Solutions

This milestone evaluates emission values by tracking long-term trends in CO2 and contributes to the identification of globally largest emitters across sectors. Scope: Including quantitative sector-based emission assessment, comparing over the years, and checking data quality and completeness

Business Problem

Knowledge about sectoral CO2 emissions is essential in framing climate-related policies and strategies. Businesses and policymakers require a detailed understanding to limit environmental damage and meet international climate change commitments. In this milestone, we will work toward catalyzing the decarbonization of CO2 emissions at the sector level to ensure that interventions are sector-specific and impactful.

Background/History

CO2 emissions are a significant source of global warming and climate change. Moreover, each sector, such as transportation, industry (including power plants), and residential or commercial facilities, all release varying quantities of CO2. Agreements such as the International Protocol in Kyoto and the Paris Agreement hope to decrease global CO2 emissions. However, to meet that goal, their success depends upon understanding where different sectors stand now and how much they have changed over time.

Data Explanation

- **Data Sources**
    - ***Kaggle Dataset:*** This dataset contains a detailed record of CO2 emissions by sector with historical information and thus can be great for analyzing trends in CO2 emission overall. [1]
    - ***CSV Data File:*** intended as the main data file, simple dataset of country-wise sectoral CO2 emissions for 2019 later and utilized as a building block for analysis.

- **Data Preparation**

    - ***Data Cleaning:*** This step involves removing errors and inconsistencies from the dataset.
    - ***Data Transformation:*** The transformation of data which makes it usable for analysis, be it normalizing what comes out of emission boxes and grouping these by sector.

- **Data Dictionary**

    - ***Country:*** The country of the recorded emissions.
    - ***Date:*** The date of the recorded emissions.
    - ***Sector:*** The category of the industry (e.g., Power, Transportation, Industry).
    - ***Value:*** The amount of CO2 emitted is measured in metric tons.
    - ***Timestamp:*** Unix timestamp of the recorded data.

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 135408 entries, 0 to 135407
Data columns (total 5 columns):
 #   Column     Non-Null Count   Dtype
---  ------     --------------   -----
 0   country    135408 non-null  object
 1   date       135408 non-null  datetime64[ns]
 2   sector     135408 non-null  object
 3   value      135408 non-null  float64
 4   timestamp  135408 non-null  int64
dtypes: datetime64[ns](1), float64(1), int64(1), object(2)
memory usage: 5.2+ MB
```

Methods

- **Statistical Analysis Tools:**

    - **Trend Analysis:** To identify temporal patterns in CO2 emissions.
    - **Correlation Analysis:** To determine the relationship between sector-specific emissions and global economic/industrial activities.
    - **Regression Models:** To predict future emission trends based on historical data.

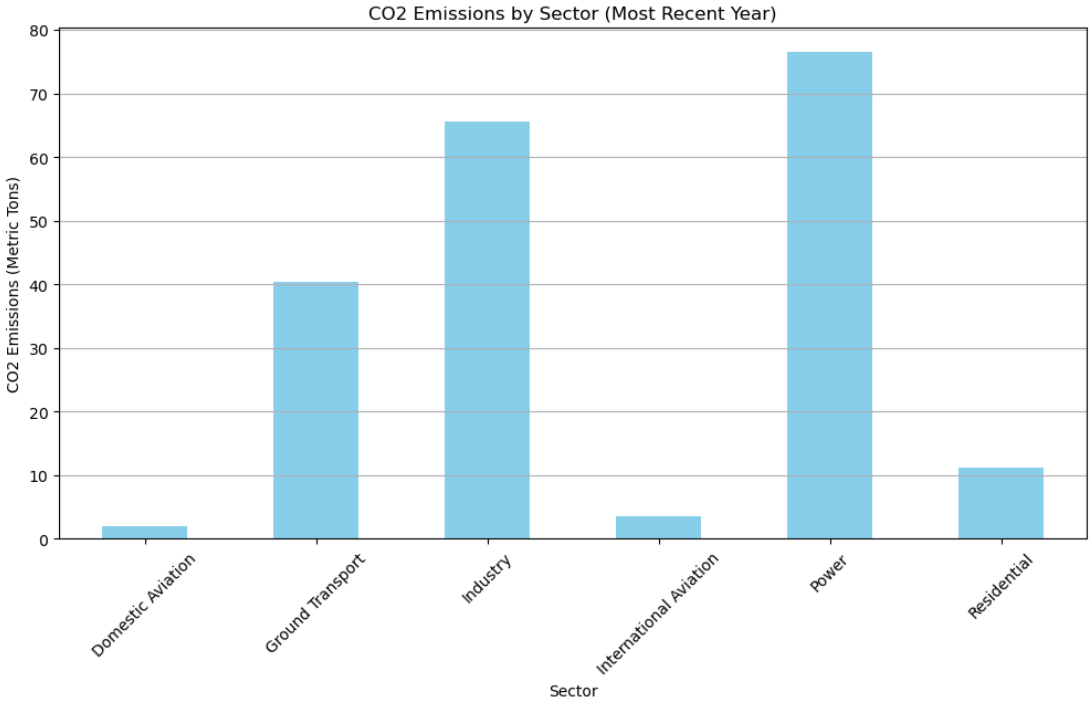
- **Machine Learning Models:**

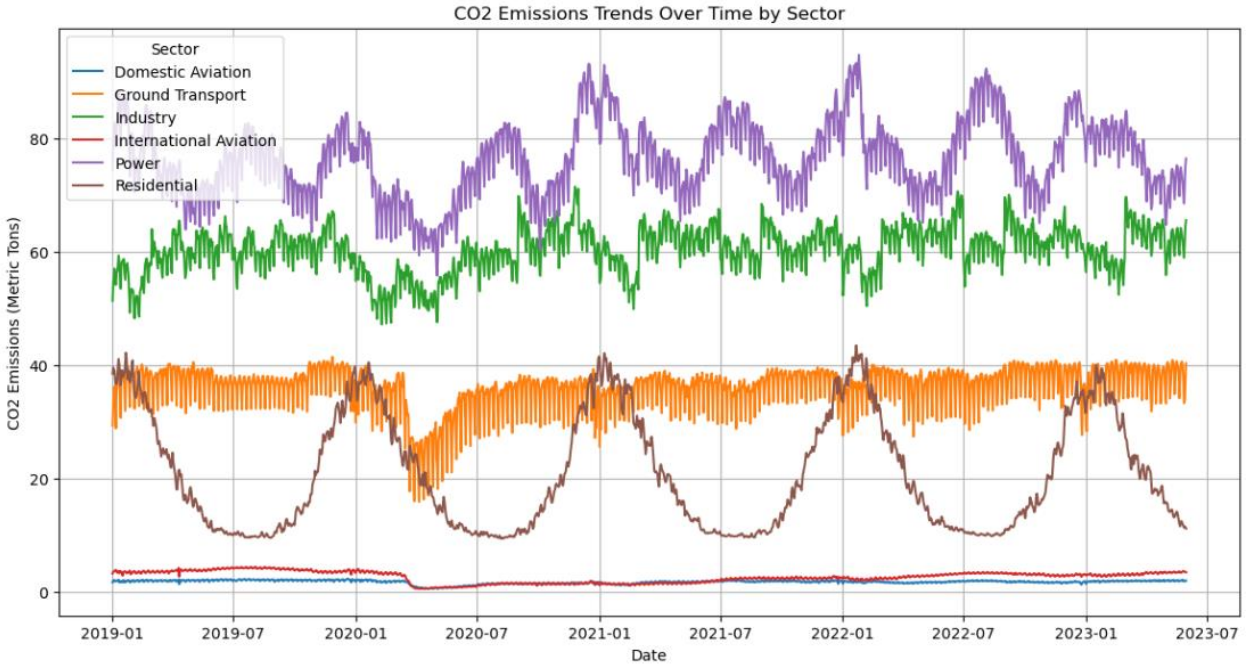
- **Time Series Forecasting:** ARIMA models learn regularities in CO2 emissions data (such as seasonal variations) and use those regularities to forecast future CO2 per capita emissions for future years.
- **Classification Models:** Using decision trees or random forests to identify critical factors driving CO2 emissions in different sectors.

- **Visualization Techniques:**
  - **Line Graphs:** To show trends over time.
  - **Bar Charts:** To compare emissions across sectors.

Analysis

1. **CO2 Emissions by Sector:**
   - These charts illustrate the total CO2 emissions by sector, providing insights into which sectors are the most significant contributors to emissions.

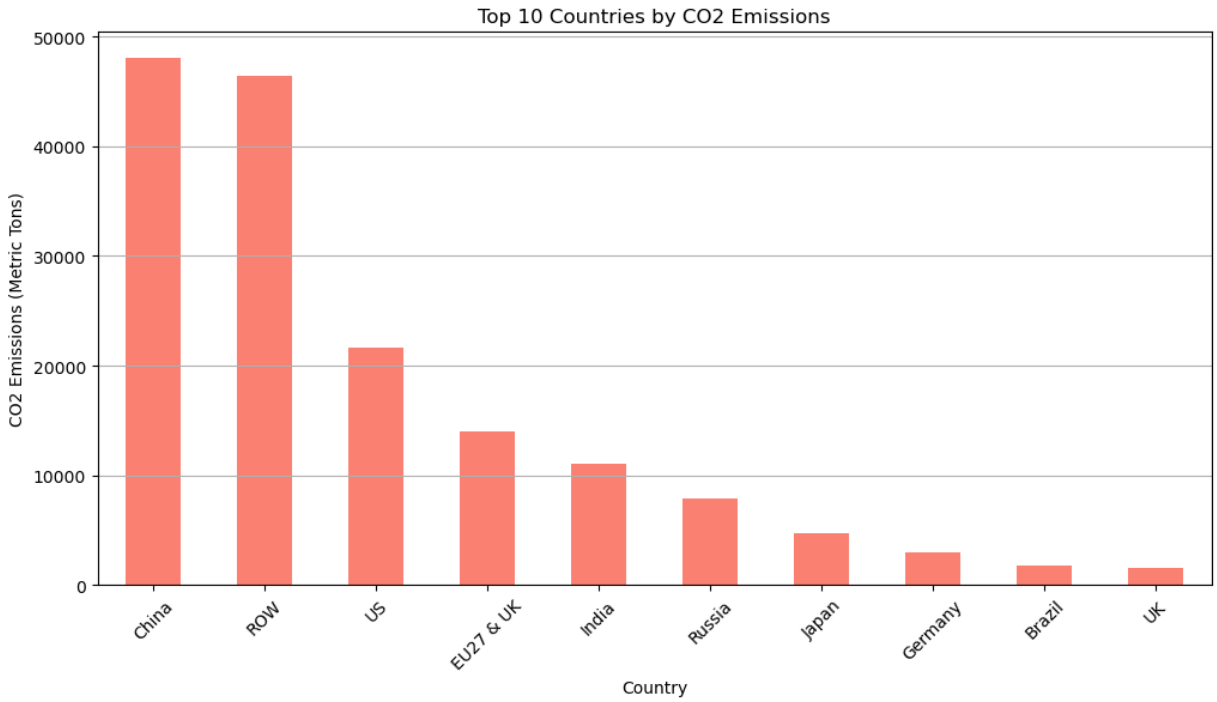




- ***Key Observations:***
  - The power sector is the most significant contributor to CO2 emissions, followed by the transportation and industry sectors.
  - Efforts to reduce emissions in these sectors should be prioritized.
  - Transitioning to renewable energy sources and improving energy efficiency in these sectors could significantly reduce overall emissions.

2. Top 10 Countries by CO2 Emissions

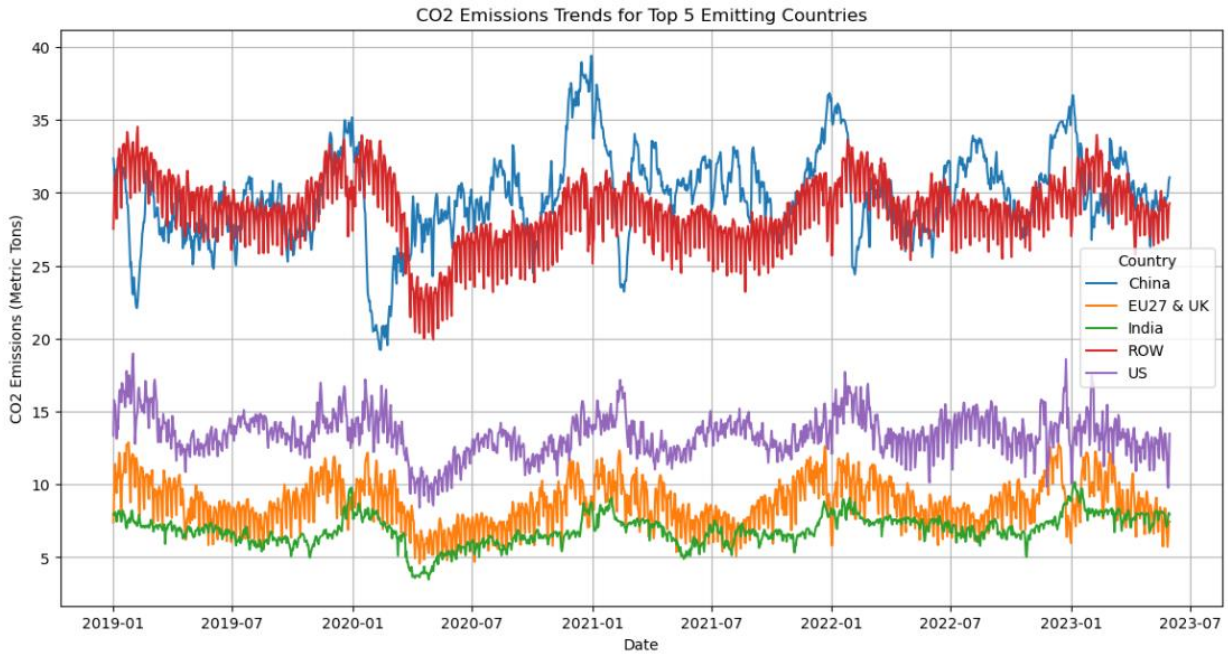
- The bar chart identifies the top 10 countries contributing the most to CO2 emissions.



- This visualization reveals that a few countries dominate global CO2 emissions, highlighting the need for targeted interventions in these critical regions.
- ***Key Observations:***
  - The top emitters include China, the United States, and Russia.
  - These countries have significantly higher emissions compared to others in the top 10 list.
  - Policy efforts in these countries could have a substantial impact on global CO2 reduction.

3. CO2 Emissions Trends for Top 5 Emitting Countries:

- The line graph shows the trends in CO2 emissions over time for the top 5 emitting countries, allowing us to analyze how emissions have changed over the selected period.



- ***Key Observations:***
  - Emissions trends vary significantly among the top emitters.
  - Some countries show a decreasing trend in emissions, while others have increasing or fluctuating patterns.
  - Understanding these trends helps assess existing policies' effectiveness and identify areas for improvement.

Conclusion

This study examines carbon dioxide emissions across sectors, focusing on major sources and patterns. The results suggest the importance of implementing strategies for reducing emissions in sectors with pollution levels, such as transportation and industry.

Assumptions

- The data from Kaggle and Carbon Monitor is accurate and comprehensive.
- The historical trends will continue without significant policy changes or technological advancements.

Limitations

- Real-time data may not capture all emissions due to reporting delays.

- The analysis may not account for all variables influencing CO2 emissions, such as technological changes or unreported activities.

Challenges

- Ensuring data accuracy and completeness.
- Addressing potential biases in data sources and analysis methods.
- Communicating complex data insights to a non-technical audience.

Future Uses/Additional Applications

- Extending the analysis to include more granular data, such as city-level emissions.
- Integrating additional data sources, like satellite imagery, for more accurate real-time monitoring.
- Applying machine learning models to predict the impact of specific policies on CO2 emissions.

Recommendations

- Prioritize Major Polluters: Concentrate on nations with emissions for policy measures.
- Tailored Approaches for Different Sectors: Give importance to lowering emissions in the Energy, Transport, and Manufacturing industries.
- Track Developments: Regularly observe emission patterns to assess policy efficacy and adjust.

Implementation Plan

- **Data Collection and Validation:**
    - Continuous updating and validation of data sources.
- **Policy Development:**
    - Collaborate with stakeholders to develop targeted policies.
- **Monitoring and Reporting:**
    - Establish a framework for regular monitoring and reporting of emissions.

Ethical Assessment

- Ensure that the analysis and interpretation of data is unbiased, without any institutional biases.
- Consider the economic implications of proposed policies, for all segments of society.

- Maintain transparency in reporting. Ensure accountability in implementing policies.

Appendix

- **Supporting Documentation:**
    - **Data Dictionary:** Detailed explanation of each variable in the dataset.
    - **Data Cleaning Steps:** Steps and methods used for cleaning the data.
    - **Additional Visualizations:** Above Charts and graphs that supplement main findings.

Questions and Answers

- **Which industries have the highest contribution to the emission of CO2?**

  The significant areas of emission are Power Generation, Transportation, and Industries. These sectors contribute to the emission more than any other sector, and the sector emitting the most is the power sector.

- **What trends characterized these sectors' emissions in the past?**

  Analyzing the trends for emissions, emissions have increased in the power and transportation sectors. The industry sector also reveals oscillations connected to economic activities. Hence, sectors would exhibit a declining factor due to efficiency and policy change.

- **What have been the 'sources of change' for emissions?**

  Other factors that have led to changes in emissions comprise growth of the economy, industrial developments, technological innovation, energy usage, and changes in policies. Fluctuations in these factors directly impact the level of CO2 emissions from different industries.

- **Open The credibility of data used in the analysis to what extent are the data used in the analysis valid?**

  The data applied in the analysis is obtained from authenticated sources such as Kaggle and Carbon Monitor. In general, the data gathered is extensive and sufficient for analysis. Despite this, the real-time data might only account for some emissions since it is based on reporting that it might experience delays. In conclusion, the data is reliable for completing the analysis.

- **What are the implications of the formulated policies on the suggested reform?**

If implemented, such policies could help drastically cut the emissions of CO2 from the major emission sectors. Measures related to the power sector include changing power sources and increasing demand for efficiency, which can reduce emissions. Likewise, enhancing sustainable transportation and industries can reduce emissions within those industries.

- **What decisions can organizations make to help decrease the amount of CO2 emissions?**

  Companies can do this by integrating advanced products, converting to green power, practicing eco-friendliness, and buying carbon offsets. Also, companies can demand and obey environmental standards and measures to decrease emissions.

- **What are those technologies that can prompt the reduction of emissions?**

  Technological solutions that may be implemented to lower emissions include:

  - Renewable energy technologies (solar, wind).
  - Energy storage technologies.
  - Electric and hybrid vehicles.
  - Carbon capture and storage (CCS) technologies.
  - Efficiency improvement in industries.

- **In what way does the government impact CO2 emissions through global economic activities?**

  Looking at it from an economic perspective, activities in the world economy are linked to the rate of CO2 emissions. Many economists have proposed that economic growth would promote industrialization, hence energy consumption and transport activities leading to higher CO2 emissions. At the same time, emissions may decline due to an economic crisis and a decrease in manufacturing and vehicular movements.

- **What are the weaknesses of the approach used in the current analysis?**

  The reductionist approach to the analysis of significant quantitative aspects can omit several factors that may affect the level of emitted CO2, including activities that are not disclosed to the public or technological shifts. Also, the data used in real-time may be delayed, and the analysis of

the trends in the past few years may not be continued without more severe changes in the political decisions or development of modern technologies.

- **What future directions could be taken in this analysis, or how could it be applied for improvement?**

  Further research into this type of work can be conducted by collecting and analyzing the data on CO2 emissions at a more refined geographical level, like city level, as well as including other sources of data, like high-resolution satellite images for the real-time monitoring of the emissions, and using more sophisticated machine learning techniques to analyze the performed indicators depend on some policy changes. Further, many variables and scenarios should be included in the analysis to understand emissions and their factors better.

By addressing these points, this white paper aims to provide a detailed and actionable insight into CO2 emissions by sector, supporting informed decision-making for environmental policy and strategy development.

References

Carbon Monitor. (n.d.). Real-time estimates of CO2 emissions. Retrieved June 9, 2024, from https://carbonmonitor.org/

Saloni1712. (n.d.). CO2 emissions by sectors [Data set]. Kaggle. Retrieved June 9, 2024, from https://www.kaggle.com/datasets/saloni1712/co2-emissions