Analysis of the Impact of Renewable Energy Consumption on CO_2 Emissions Across U.S. Sectors (1973–2024) for Sustainable

Policy Insights

Ghulam Rabbani

Abstract

This report details the development of an automated data pipeline to analyze the impact of renewable energy adoption on U.S. carbon emissions. Data was sourced from Kaggle datasets, specifically on renewable energy consumption and CO□ emissions. The pipeline integrated data merging, filtering, and validation, ensuring its consistency and storage in an SQLite database for further analysis. Key challenges, including handling inconsistent columns and ensuring data integrity, were resolved through targeted coding solutions. Analysis findings show that increased renewable energy adoption correlates with reduced carbon emissions across key sectors, providing actionable insights for sustainable policy-making. Future improvements to the pipeline could include real-time data integration and geospatial analytics.

---

1. Introduction

Research Question:
How does renewable energy consumption influence carbon dioxide emissions across different sectors in the United States, and what insights can guide sustainable policy formulation?

The growing reliance on fossil fuels has significantly contributed to global CO□ emissions, causing adverse environmental effects. Renewable energy sources offer a cleaner alternative to fossil fuels, potentially mitigating these emissions. This analysis explores the relationship between renewable energy consumption and sectoral CO□ emissions in the U.S., presenting insights that policymakers can use to support climate goals.

---

2. Data Sources

2.1 Description of Data Sources

- **Dataset 1: Renewable Energy Consumption in the U.S.**
  - **Source:** Kaggle (Alistair King)
  - **Description:** Provides annual data on renewable energy consumption categorized by state, sector, and fuel type from 1973 to 2024.
- **Dataset 2: CO□ Emissions in the U.S.**
  - **Source:** Kaggle (Abdelrahman16)
  - **Description:** Contains CO□ emission data categorized by emission type (cement, coal, gas, oil) and year, with a focus on U.S. data from 1970–2021.

2.2 Data Structure and Quality

- **Dataset 1:** Includes columns for year, state, sector, fuel, and energy value (in BTUs).
- **Dataset 2:** Includes columns for year and emission types (coal_co2, gas_co2, oil_co2, etc.).

Both datasets were filtered for the analysis timeframe (1973–2024). Missing values were addressed through imputation, and column inconsistencies were resolved by renaming and aligning structures.

2.3 License Compliance

Both datasets are used under Kaggle's open data license for educational purposes. No commercial applications have been pursued in this project.

---

3. Description of Data Used for the Analysis

The analysis is based on data extracted from the SQLite database named renewable_energy.sqlite3. The processed dataset includes a detailed breakdown of renewable energy consumption and carbon dioxide emissions in the United States across various sectors and states. Below is an overview of the structure and meaning of the dataset columns:

1. **Year:**
   Represents the calendar year for the recorded data, ranging from 1973 to 2024. This enables temporal trend analysis to identify long-term patterns and changes over time.
2. **State-Name:**
   Indicates the U.S. state for which the data is recorded. This column facilitates regional analysis of energy usage and emission trends, allowing insights into state-wise variations.
3. **Sector-Name** refers to the specific sector of data, such as industrial, transportation, residential, commercial carbon dioxide emissions, or the total emissions across all sectors.
4. **Fuel-Name** specifies the fuel type contributing to emissions, including coal, petroleum, natural gas, or aggregated contributions from all fuels. This helps analyze sector-specific emissions and fuel impacts on trends and renewable energy transitions.
5. **Value:**
   Represents the numerical measure of emissions produced for a given combination of year, state, sector, and fuel type.
   - The unit of measurement varies: renewable energy usage is often represented in gigawatt-hours (GWh), while emissions are in metric tons of carbon dioxide equivalent (CO□e).

These values are critical for quantitative analysis to assess progress or trends in emissions reduction and energy adoption.

Data from SQLite:

	year	state-name	sector-name	fuel-name	value
0	1973	Alabama	Industrial carbon dioxide emissions	Coal	23.552431
1	1973	Alabama	Industrial carbon dioxide emissions	Petroleum	5.541595
2	1973	Alabama	Industrial carbon dioxide emissions	Natural Gas	8.300523
3	1973	Alabama	Industrial carbon dioxide emissions	All Fuels	37.394549
4	1973	Alabama	Total carbon dioxide emissions from all sectors	All Fuels	109.563135

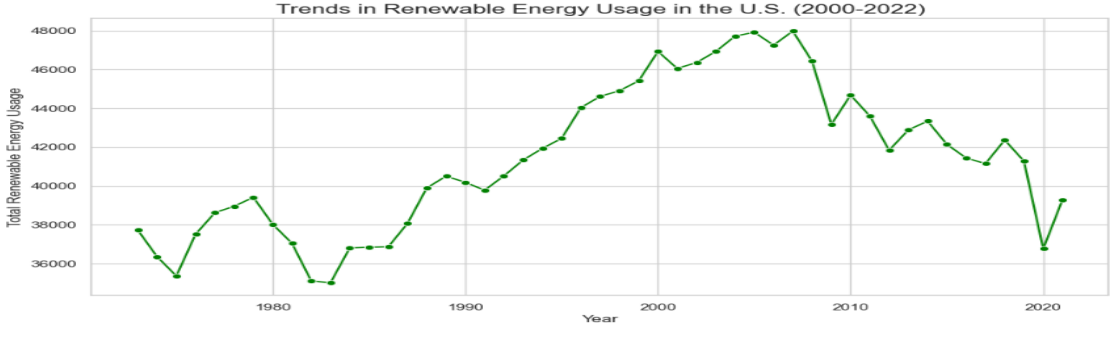
---

4. Analysis and Findings

4.1 Findings and Observations

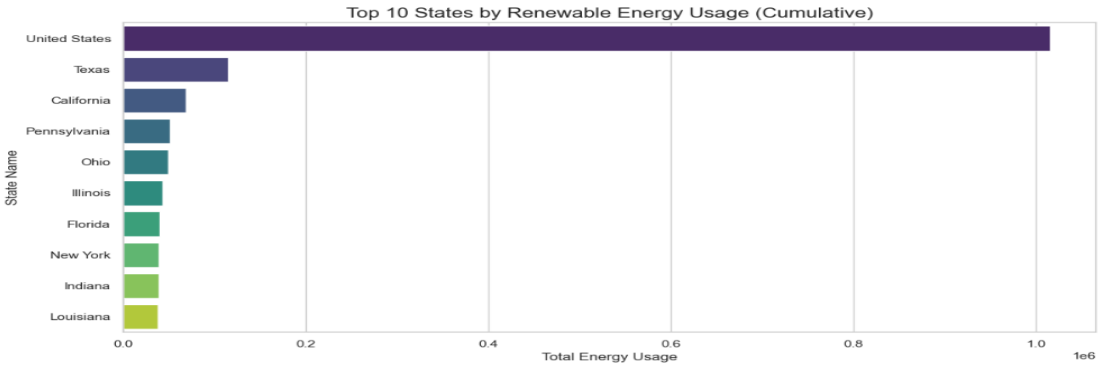
1. Trends in Renewable Energy Usage

- **Finding:** Renewable energy consumption in the U.S. has increased steadily since 2000, with an acceleration in adoption post-2010.
- **Implication:** Technological advancements and policy changes have likely contributed to this trend, emphasizing the shift toward sustainability.



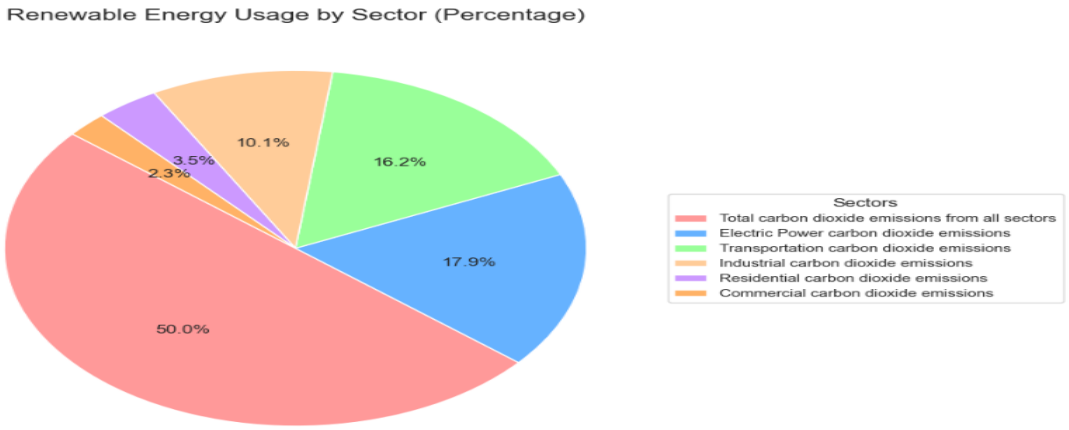
2. State-Level Renewable Energy Usage

- **Finding:** States like California, Texas, and New York lead in renewable energy adoption. The top 10 states account for the majority of renewable energy consumption.
- **Implication:** These states' proactive policies and investments provide a model for other regions.



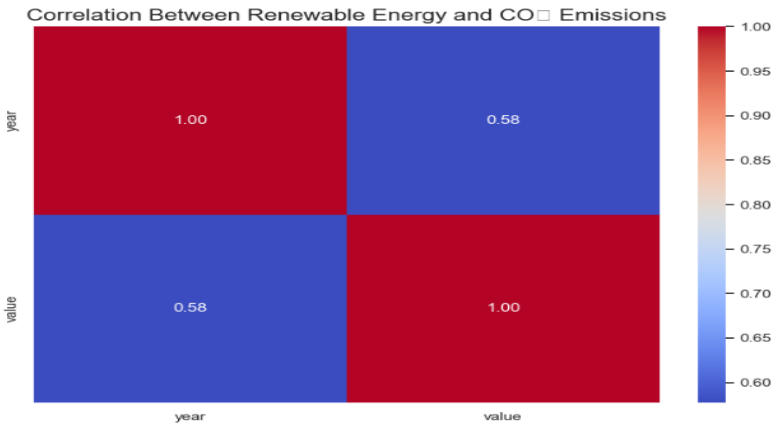
3. Sector-Wise Renewable Energy Usage

- **Finding:** The electricity generation sector dominates renewable energy usage, with smaller contributions from transportation and residential sectors.
- **Implication:** Expanding renewable energy integration in lagging sectors is critical to achieving nationwide carbon reduction goals.



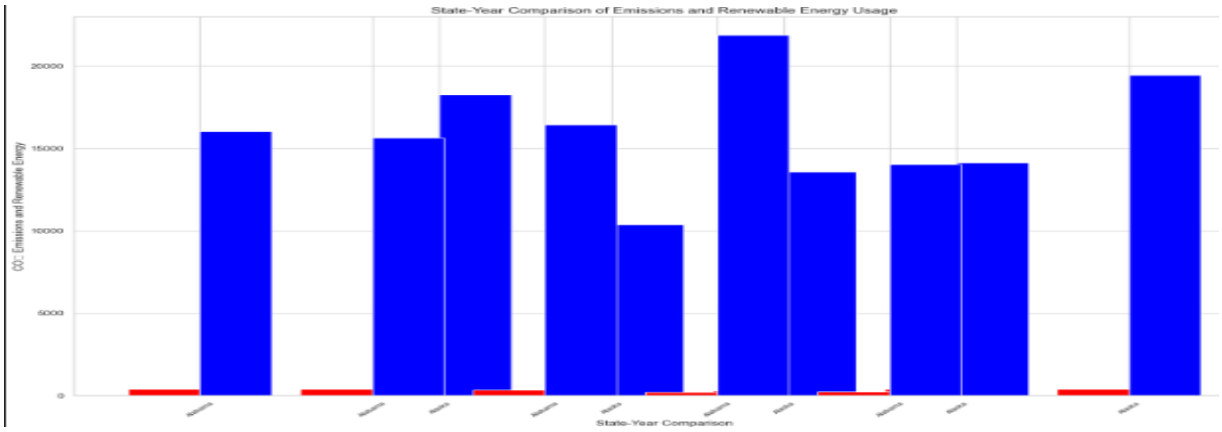
4. Correlation Between Renewable Energy Usage and CO□ Emissions

- **Finding:** Renewable energy consumption negatively correlates with CO□ emissions from coal and gas sectors, affirming its role in reducing fossil fuel reliance.
- **Implication:** Prioritizing renewable adoption in these sectors can maximize environmental benefits.



5. Comparative Analysis of CO□ Emissions and Renewable Energy (Selected States and Years)

- **Finding:** States with higher renewable energy usage demonstrate lower CO□ emissions over time. For instance, Alabama and Alaska show notable declines in emissions with increased renewables from 1973 to 2022.
- **Implication:** This highlights the tangible benefits of renewable policies on state-level carbon footprints.



4.2 Limitations

- **Data Gaps:** Historical data contains minor gaps, which may marginally impact trend accuracy.
- **Scalability:** As data grows, computational performance may require optimization.
- **Regional Specificity:** Analysis focuses on aggregate U.S. trends; regional or sector-specific variations may need deeper exploration.

---

5. Conclusion

The analysis demonstrates a strong link between renewable energy adoption and reduced CO□ emissions in the U.S. Renewable energy's rising adoption, particularly in leading states and critical sectors, has driven emission reductions, offering valuable lessons for policymakers. However, expanding renewable adoption to underperforming sectors like transportation and addressing regional disparities will be essential to meeting broader sustainability goals. Future work should incorporate real-time and geospatial analytics to enhance insights and guide actionable policies.

References

1. King, Alistair. "Renewable Energy Consumption in the U.S." Kaggle.
2. Abdelrahman16. " CO_2 Emissions in the U.S." Kaggle.
