Exploratory Data Analysis Report

**Carbon Emissions across United States**

**Project Title:** **Carbon Emissions across America: Unveiling Regional Disparities, Key Contributors**

**Introduction**

This report presents a Comprehensive Analysis of CO2 Emissions in the United States, focusing on trends across various sectors and individual states. The objective is to understand the patterns, variations, and potential impacts of CO2 emissions

The motivation for the project comes from the need to tackle the Environmental challenges and the leading causes for these issues.

**Data Description**

The Dataset comprises of detailed Co2 emissions across different sectors including Transportation, Electric Power, Industrial Sectors, and others. The data spans a timeline from 2010 to 2021 covering both national and state-level analysis.

1. US EPA (Environmental Protection Agency – Data on the Greenhouse gas emissions in the United States.
2. US EIA (Energy Information Administration) – Data on State level emissions, Energy production and Electricity Generation Source.

**Data Cleaning and Preparation-** The steps involved

1. Import Libraries – Import the necessary libraries, like Pandas, Matplotlib.
2. Load Data into the Data Frame – loading the US and Statewise datasets

Converted our data from Excel and CSV files into Pandas Data Frames using the pandas .read\_excel() and .read\_csv()

1. Preliminary Data Examination –Used various Pandas function to get an overview of the data. Looked at the structure of the data frame using .head(), the .info() to check the datatypes and values, and the .describe() for Summary Statistics
2. Data Cleaning

* Handling missing values – Checked for missing values in the data frame using the df.is.null().sum()
* Checking for duplicates – using the df.duplicates()
* Removal on unnecessary Data –drop() function
* Quality checks on the data – looked to see the final structure of the cleaned-up data frame and made sure that the information and format is consistent across the Data Frames used for our analysis

1. Data Visualizations – Bar charts and Plots for showing the Trends across the years for different states, sectors.

**Data Analysis – Visualizing datasets, plotting, and finding answers to the research questions**

Research Questions:

Time Period for our analysis is from 2010-2021

**1. What part of the overall Greenhouse Gas emissions are Co2 emissions**

1. **Greenhouse Gas Emissions Stacked Chart for Overall emissions by Gas**

A graph with different colored lines

Description automatically generated with medium confidence

**2. What are the Overall trends in Carbon Dioxide (CO2) Emissions in US over the past decade**

1. **Greenhouse Gas Emissions Trendline**

A trendline was plotted to visualize the overall trend of greenhouse gas emissions for the period of 2010-2021. The line chart revealed the general trajectory of CO2 Emissions over the specified period

A graph with different colored lines

Description automatically generated

* Analysis –
  + Carbon dioxide consistently sits nearly **five times higher** than methane, nitrous oxide, or fluorinated gases

**3.How do the emissions vary across regions /states in US**

1. **CO2 Emissions for States in the US (Average Bar Chart)**

A bar chart was created to display the average CO2 Emissions for each state. This visualization allowed for a comparison of emission levels across different states.

A white screen with colorful bars

Description automatically generated with medium confidence

**Analysis –**

* + Texas and California, two of the largest states in the US, are the two highest producers of Co2 Emissions, followed by Florida, Pennsylvania, Illinois, and Ohio. This could be because they are both rank in the top 2 states for population and economic size.
  + Texas’ emissions are **significantly higher than any other state. (Approximately 650 million metric tons)**
  + States with Significantly lower emissions- According to EPA, these states often exhibit a combination of lower population density, less energy-intensive industries, cleaner energy sources, and proactive environmental policies contributing to their reduced carbon footprint.

**4. What are the major sectors or activities contributing to the emissions in US and how is it different regionally?**

1. Trendline showing emissions across Sectors

Sectoral emissions were explored through a trendline chart, illustration the variations in CO2 emissions across sectors- Transportation, Electric Power, Industrial and others over the specified period.

A graph with different colored lines

Description automatically generated

**Analysis**

* Electric power industry, transportation, and industry sectors have remained the **three highest producers of Co2 emissions**
* In recent years, **the top three producers have decreased emissions** to be between 1500 to approximately 1800 mmt (million metric tons)
* **A significant gap has remained** between the top three producers and the bottom three

**According to Sources (EPA and EIA) this is because of**

* Heavy reliance on fossil fuels for operations especially in Transportation and Industrial, and Electrical
* Higher energy intensity demand for powering equipment’s and machinery
* Electricity sector has seen a significant decrease possibly due to the use of renewable sources of energy
* Agriculture produces more of methane and Nitrous Oxide and emissions are significantly less
* Residential and commercial contribute mainly through energy consumption for heating, electricity uses etc. there have been advancements in energy efficient technologies, so this explains why it is lower compared to the top 3

1. Top 5 States in each of the 3 Top Sector

An analysis was conducted to identify the top 5 states with the highest emissions in each sector. This analysis provided insights into the regional variations and emphasized the significant contributors in each sector

**Transportation sector CO2 Emissions**

A graph with different colored lines

Description automatically generated

**Analysis Transportation Sector**

* The top five Co2 producers change for the transportation sector, with Texas being the only to remain
* Texas and California are the highest Co2 contributors in the transportation sector and both follow similar trends over the past 10 years
* All five states experience a sharp rise in emissions after 2020 (this could be due to Post COVID)

**Information from other sources for reasons specific as to why these states are the top contributors of CO2 Emissions in the Transportation Industry:**

According to U.S Environmental Protection Agency (EPA), and U.S Energy Information Administration (EIA), these states rank high due to factors such as traffic congestion, reliance on personal vehicles, limited public transportation alternatives and in the case of California the significant port and trade activities.

These states continue to work on various initiatives to reduce transportation emissions such as promoting Electric Vehicles (EV) and enhancing fuel efficiency standards, infrastructure development to support cleaner and sustainable transport options. However, the scale of their populations and economic activities makes it a challenge to reduce the emissions significantly

**Electric Sector CO2 Emissions**

A graph with different colored lines

Description automatically generated

**Analysis Electric Sector**

* Texas’ electric industry Co2 emissions is around 100 mmt **more** than four out of the top five Co2 producing states
* There has been a gradual **downward trend** in all five of the top five Co2 producing states in the last 10 years

**Information from other sources for reasons specific as to why these states are the top contributors of CO2 Emissions in the Electric Power Industry**

* According to U.S Environmental Protection Agency (EPA), and U.S Energy Information Administration (EIA), these states rank high because
  + Heavy reliance on fossil fuels for electricity generation.
* These states are taking efforts to reduce the CO2 emissions in this sector by increasing the use of renewable sources of energy (Wind and Solar Power), improvements in energy efficiency. However, these remain a challenge due to various factors (economic, policy related and infrastructure)
* Indiana has seen a downward trend possibly due to Environmental Regulations, transitions away from coal, increased renewable energy integration, they have been implementing energy efficiency programs aimed at reducing energy consumption and using Energy efficient technologies
* Texas also sees a decrease due to some of the reasons like closure of Coal plants, Transition to Natural Gas and renewable sources of energy, Air quality regulations.

**Industrial Sector CO2 Emissions**

A graph of the state of texas

Description automatically generated

**Analysis Industrial Sector**

* Texas’ industrial sector is also the highest contributor among the top five states
* Texas also displays a **gradual upward trend** in industrial Co2 emissions, whereas other states remain steady
* **Texas currently sits above 225 mmt**, whereas the other four states are all **below 125 mmt**

Information from other sources for reasons specific as to why these states are the top contributors of CO2 Emissions in the Industrial Sector (Source EIA and EPA)

Heavy industrial presence is related to the higher emissions in these states that encompass oil and gas plants, chemical manufacturing, and other production activities.

Presence of Energy Intensive industries (Oil and Gas refining, chemical manufacturing, steel production), Fossil fuel extraction and processing (Texas,Lousiana,Pennsylvania) , Chemical and Petro chemical manufacturing(Texas and Louisiana), Steel and Metal production(Pennsylvania and Indiana), Large Power Plants, the inherent nature of these industries in these specific states contribute to the Co2 emissions

1. Trendlines for state and Sector Emissions

Detailed trendlines were plotted to analyze the total CO2 emissions and sectoral emissions for the top 5 states and national average. These trendlines illustrated the fluctuations in emissions and sectoral contributions over time.

**Industrial Sector emission comparison to total emission for Top 5 States**

A Comparison chart was generated to showcase CO2 emissions for the top5 states in the Industrial sector against their total Co2 emissions. This comparison highlighted the dominance of Industrial sector in contributing to the overall emissions for these states

**A graph with different colored lines

Description automatically generated**

**Analysis**

* Texas is the leading state in Industrial CO2 emissions.
* While California’s average CO2 emissions is ranked 2nd, the same is not true for their Industrial emissions.
* This is related to California’s goal to go electric over the years. **Industrial tasks that usually require gas and oil California is working to make it electric.**

**Electric Sector emission comparison to total emission for Top 5 States**

A Comparison chart was generated to showcase CO2 emissions for the top5 states in the Electric sector against their total Co2 emissions. This comparison highlighted the dominance of Electric sector in contributing to the overall emissions for these states

A graph with different colored lines

Description automatically generated

**Transportation Sector emission comparison to total emission for Top 5 States**

A Comparison chart was generated to showcase CO2 emissions for the top5 states in the Transportation sector against their total Co2 emissions. This comparison highlighted the dominance of Transportation sector in contributing to the overall emissions for these states

A graph of different colored lines

Description automatically generated

**5.How do total CO2 emissions for the Top 5 states compare against the National Average of the total CO2 Emissions?**

1. Bar Chart Comparison for the Top 5 States Vs National Average

A bar chart was presented to compare the average CO2 emissions for the top 5 states with the national average. This comparison shed light on the states that surpassed or the national average in CO2 emissions

A graph of blue and green bars

Description automatically generated

**Analysis**

* The national average of CO2 emissions is minimal compared to the top 5 producing states.
* Texas, California, Florida, Pennsylvania, and Illinois are the leading states in CO2 emissions.
* The gap between the leading state, Texas, and California in CO2 emissions shows the difference between California's push to quadruple clean energy production while Texas has no current goals.

**Information from other sources for reasons specific to the trend (Source EIA and EPA)**

Though there have been measures to reduce emissions the energy demand and the historical reliance on fossil fuels make the transition to lower Co2 emissions a challenge.

The higher presence of industries, population, and economic factors (some of the populous states in the country have high levels of economic activity compared to the smaller states), Energy production and consumption, transportation and infrastructure, Electricity generation mix, Climate and Geographical factors, Industrial composition all contribute to the varied difference with the National Average.

1. Deeper Dive into Texas CO2 Emissions by Sector

A graph with different colored lines

Description automatically generated

**Analysis:** Texas has been the Top State in all the sectors we looked at (Industrial, Electric Power and Transportation).

Texas the electric industry has seen a decline in Co2. This could be due to use of Renewable sources of energy for Electricity Generation.

Information for Texas from other Sources (EPA and EIA)

* Texas leads as a top state in Crude oil and natural gas production. It also has the largest number of crude oil refineries, it has 1/3rd of the nation’s total refining capacity
* Texas produced more electricity than any other state
* Texas leads in the energy consumption for the entire of US across all the sectors (Largest Energy consumption by any state in the US)
* Texas has a lot of State refineries and Petro-chemical plants in the state.

**6. What are the limitations of the analysis, and what additional data or research might be needed to address these limitations?**

1. We used the dataset for the last decade, we can further our research to identify the trend across the last 3 decades to identify what has been resulting in the decrease in Co2 emissions over the years.
2. It would also be a good idea to explore datasets from this year and last year to see how the emissions across states have changed post Covid.
3. We can also further our research to see how the policies and initiatives (like renewable energy adoption) have an impact on the Co2 emissions.
4. We have seen a dip in Co2 emissions across states and sectors, this observed dip could be due the effect of Covid and lockdown. This impact could be verified using data from the period and other datasets

**Impact Analysis and Conclusions -**

The analysis revealed several key insights

* The steady increase in CO2 emissions at the national level, indicates a pressing need for mitigation strategies.
* Variance in Sectoral contributions across states emphasizing specific sectors demands targeted interventions
* Disparities between states and the national average signaling the need for tailored policies at the state level.

**Conclusion –** The analysis of CO2 emissions from 2010-2021 showcased diverse trends, sectoral contributions, and state-level variations. The findings show the urgency for proactive measures and polices to address emissions with a particular focus on high -contributing sectors and regions. This data-driven analysis can serve as a valuable guide for formulating targeted strategies to mitigate environmental impacts and foster sustainable practices

**References:**

1. <https://www.eia.gov/environment/emissions/state/>
2. <https://cfpub.epa.gov/ghgdata/inventoryexplorer/#allsectors/allsectors/allgas/gas/all>
3. <https://www.statista.com/statistics/220174/total-us-electricity-net-generation-by-fuel/>
4. <https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions>
5. <https://calmatters.org/environment/2023/01/california-electric-cars>