

# Energy Consumption Analysis and Prediction

## Introduction-

The "Energy Consumption Analysis and Prediction" project comprehensively explores global and individual country-specific energy landscapes, employing advanced data analytics and machine learning techniques. Through a holistic examination, it uncovers trends in energy mix, per capita consumption, renewable energy growth, and carbon intensity. The project provides a global perspective by scrutinizing the intricate relationship between a country's GDP and primary energy consumption, offering insights into worldwide trends. It uses sophisticated linear regression models to predict future energy consumption, evaluated through Root Mean Squared Error, ensuring robust modeling. The project's interactive visualization, facilitated by Matplotlib and Seaborn, culminates in a unified dashboard, offering stakeholders a coherent narrative. Beyond data analysis, the project holds policy implications, providing strategic insights for policymakers and highlighting renewable energy potential. Comprehensive documentation, including detailed reports and user guides, ensures transparency and accessibility. This project at the intersection of data science, environmental sustainability, and strategic decision-making aims to pave the way for a more sustainable and informed energy future.

## Problem Statement-

To build a Machine Learning model to predict future energy consumption, which is crucial for effective resource planning and sustainable energy management.

## Objectives –

1. **Comprehensive Data Analysis:** Conduct an in-depth analysis of historical energy data, exploring key indicators such as energy mix, per capita consumption, and renewable energy growth to gain insights into a country's energy dynamics.
2. **Machine Learning Predictions:** Develop and implement machine learning models, particularly linear regression, to predict future energy consumption trends both at the country and global levels, providing a basis for informed decision-making.

## Data Source –

1. **Energy consumption (primary energy, energy mix, and energy intensity):** this data is sourced from a combination of two sources—the BP Statistical Review of World Energy and the SHIFT Data Portal.
2. **Electricity consumption (electricity consumption, and electricity mix):** this data is sourced from a combination of two sources—the BP Statistical Review of World Energy and EMBER – Global Electricity Dashboard.

3. **Other variables:** this data is collected from a variety of sources (United Nations, World Bank, Gapminder, Maddison Project Database, etc.).
4. <https://www.kaggle.com/datasets/pralabhpoudel/world-energy-consumption>

### Data Analytics Software used –

1. Python & Visual Studio Code
2. Libraries used:
  1. **Pandas:** Employed for efficient data manipulation and analysis, facilitating seamless handling of the dataset's temporal and structural components.
  2. **Matplotlib:** Integrated for creating visually informative plots and charts, enabling a clear representation of energy consumption trends and patterns.
  3. **Sklearn:** Implementing Machine Learning Algorithms, especially linear regression enhancing the project's predictive modeling capabilities.
  4. **NumPy:** Applied for its fundamental role in numerical operations and array manipulations.

### Methodology –

1. **Data Acquisition and Preprocessing:**
  - **Dataset:** Obtained a comprehensive energy consumption dataset, structured in CSV format.
  - **Preprocessing:** Utilized the Pandas library to load, clean, and preprocess the dataset. This involved converting the 'year' column to DateTime format, setting it as the index, and sorting the data chronologically.
2. **Exploratory Data Analysis (EDA):**
  - **Energy Mix Trends:** Visualized the share of different energy sources (coal, gas, renewables, nuclear) over the years using Matplotlib, providing insights into the evolving energy landscape.
  - **Per Capita Consumption:** Plotted the per capita electricity consumption trends, offering a perspective on individual energy usage patterns.
  - **Renewable Energy Growth:** Examined the growth of renewable sources (solar, wind, hydro) to assess the shift towards sustainable energy.
3. **Statistical Analysis:**
  - **Correlation Analysis:** Investigated the relationship between energy consumption and GDP using Seaborn, exploring potential correlations between economic indicators and energy usage.
4. **Machine Learning Model Building:**
  - **Linear Regression:** Employed Scikit-learn to build a linear regression model for predicting future energy consumption based on historical data. Conducted training and testing, evaluating the model's performance using the Root Mean Squared Error (RMSE).
5. **Worldwide Energy Consumption Analysis:**

- **Aggregated Data:** Aggregated global energy consumption data and repeated the machine learning process to predict future energy consumption for the entire world.
6. **Visualization and Interpretation:**
    - **Matplotlib:** Leveraged Matplotlib to create a unified and visually compelling representation of all analysis results, including historical trends, model predictions, and global comparisons.
  7. **Documentation and Reporting:**
    - **Jupyter Notebooks:** Incorporated Jupyter Notebooks for a well-documented and interactive presentation of the analysis. The entire codebase is structured to promote readability and ease of understanding.

This comprehensive methodology integrates data science techniques, statistical analysis, and machine learning to derive actionable insights into energy consumption trends and forecast future patterns. The combination of powerful Python libraries ensures efficient data processing, analysis, and visualization throughout the project lifecycle.

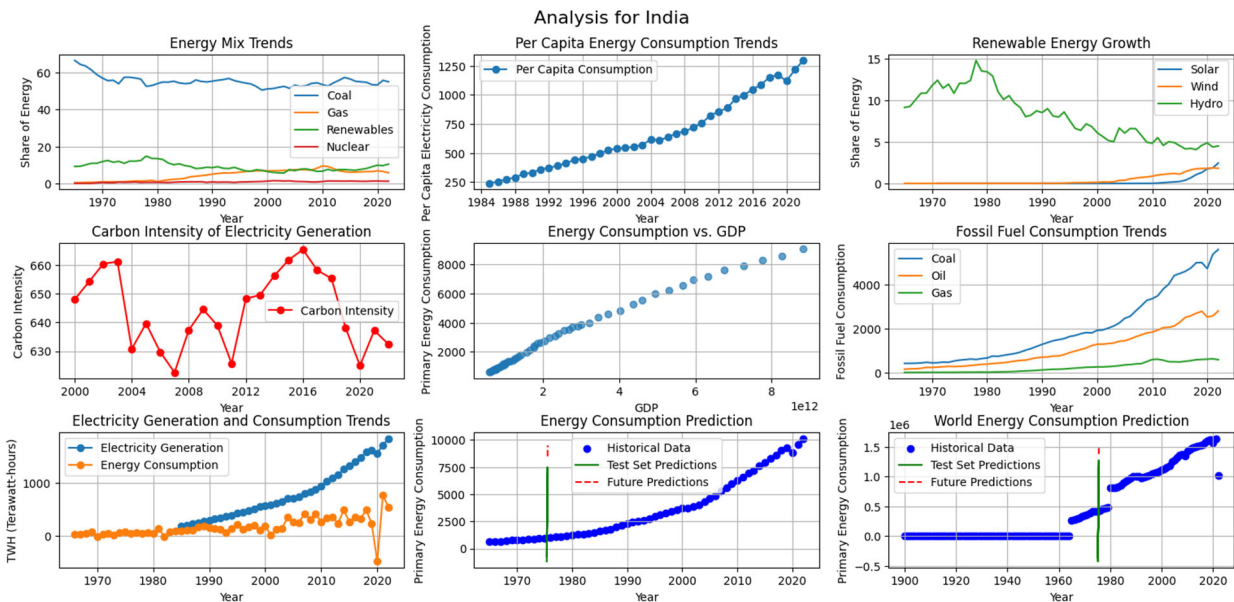
### Expected Outcomes –

1. **Insights into Energy Consumption Patterns:**
  - Understand how different countries contribute to the global energy landscape.
  - Identify trends and shifts in energy consumption behavior over time
2. **Predictive Modeling:**
  - Provide accurate predictions for future energy consumption based on historical data.
  - Assess the reliability of predictions through RMSE evaluation.
3. **Policy Implications:**
  - Offer insights for policymakers to develop strategies for sustainable energy consumption.
  - Highlight the impact of economic factors on energy use and the potential for renewable energy growth.

### Conclusion –

The Energy Consumption Analysis project aims to provide valuable insights into energy consumption patterns, supporting informed decision-making for sustainable energy policies. By leveraging data analytics and machine learning, the project delivers a comprehensive analysis that addresses both country-specific and global perspectives on energy usage.

## Sample Analysis of India –



The aforementioned output illustrates the functionality of my project, capable of analyzing and forecasting future energy consumption for a particular country and the global context. This is achieved through the application of a conventional linear regression algorithm, leveraging historical data for predictive modeling.