Analysis of Power Generation Data in India

Introduction:

Electricity generation is a crucial aspect of India's infrastructure, directly impacting economic development and quality of life. The dataset under study comprises two CSV files: one containing daily power generation data of Indian states and union territories since 2013, and the other detailing power generation from different sources such as Hydro, Wind, Solar, etc.

Content:

There are two CSV files in this dataset. One has daily power generation of all states and union territories of India since 2013. The other has daily power generation from different sources like Hydro, Wind, Solar, Diesel, Naphtha, etc.

Data Dictionary:

Daily_Power_Gen_States

Region: India is divided into five different regions.

States: All states and union territories.

Max Demand Met during the Day: Maximum demand met during the day in Mega Watts. 1 MW = 1 million watts.

Shortage during Maximum Demand: Shortage during maximum demand In Mega Watts.

Energy Met (MU): Total power generated on that day in Million Units. 1 Unit = 1 Kilo Watt Hour.

***Energy Shortage (MU): Energy shortage in Mega Units.**

Daily_Power_Gen_Source

Source: Hydro, Coal, lignite, Nuclear, RES (Renewables: wind, solar, biomass.)

NR: Northern Region

SR: Southern Region

ER: Eastern Region

WR: Western Region

NER: North Eastern Region.

All India: Total power generation from all regions for the particular source.

date: date

Data Pre-processing:

Dataset 1: Daily_Power_Gen_Source

Loading Data: Initially, we imported the dataset and examined its structure.

Data Cleaning: Rows with 'Total' in the 'source' column were removed to ensure accurate analysis.

Handling Missing Values: We identified and addressed missing values in relevant columns.

Handling Duplicates: Checked for and removed any duplicate entries.

Feature Engineering: Extracted day, month, and year from the date column for further analysis.

Dataset 2: Daily_Power_Gen_States

Loading Data: Similar pre-processing steps were applied to the second dataset.

Handling Missing Values: Missing values in 'Shortage during maximum Demand (MW)' and 'Energy Met (MU)' columns were addressed.

Handling Duplicates: Checked for and removed any duplicate entries.

Feature Engineering: Extracted day, month, and year from the date column for further analysis.

Exploratory Data Analysis (EDA)

Dataset 1 Analysis:

Descriptive Statistics: Provided a summary of key statistics.

Trend Analysis: Visualized the trend of power generation by different sources over time.

Source Comparison: Compared power generation by different sources using bar plots.

Time Series Analysis: Analysed the time series data for Hydro power generation.

Renewable vs. Non-Renewable Sources: Compared the contribution of renewable and non-renewable sources to power generation.

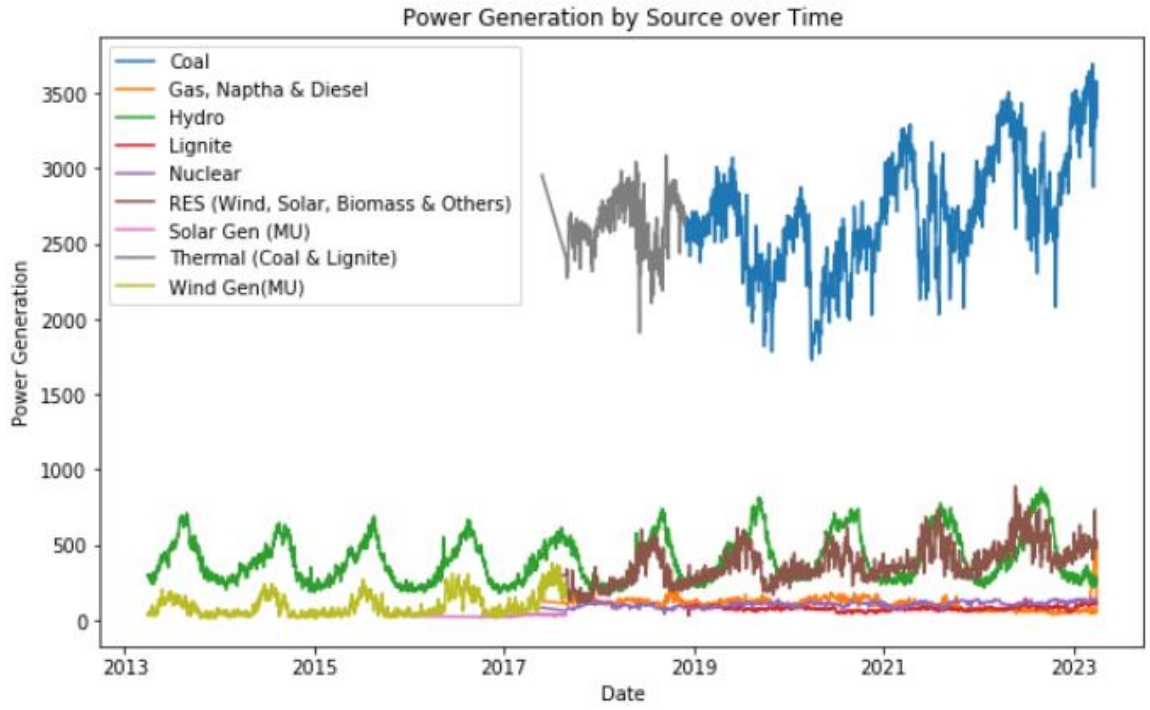
Total Energy Production: Presented the distribution of total energy production by source using a pie chart.

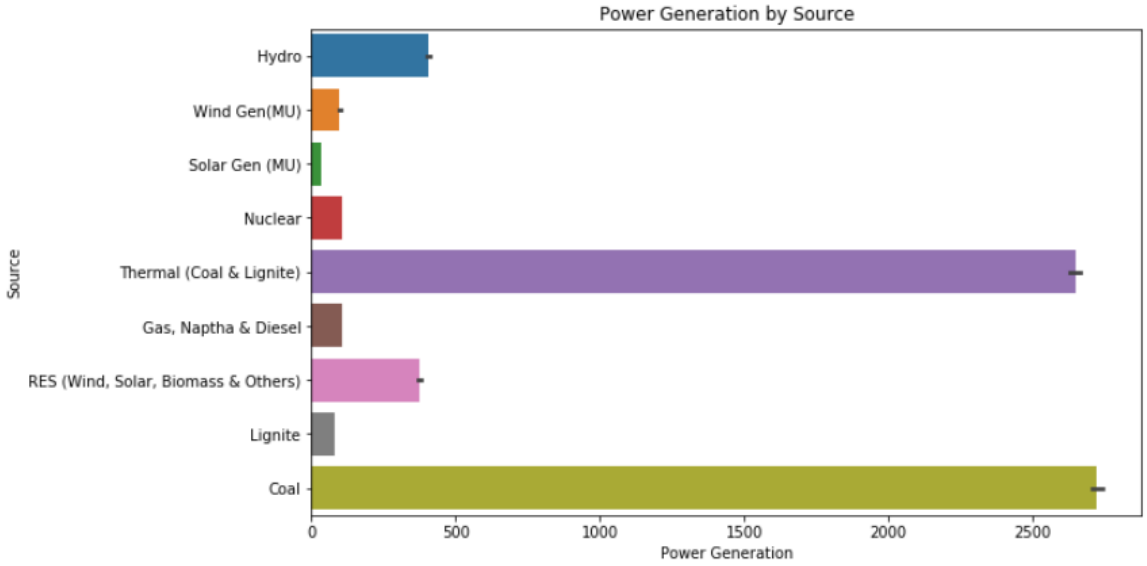
Dataset 2 Analysis:

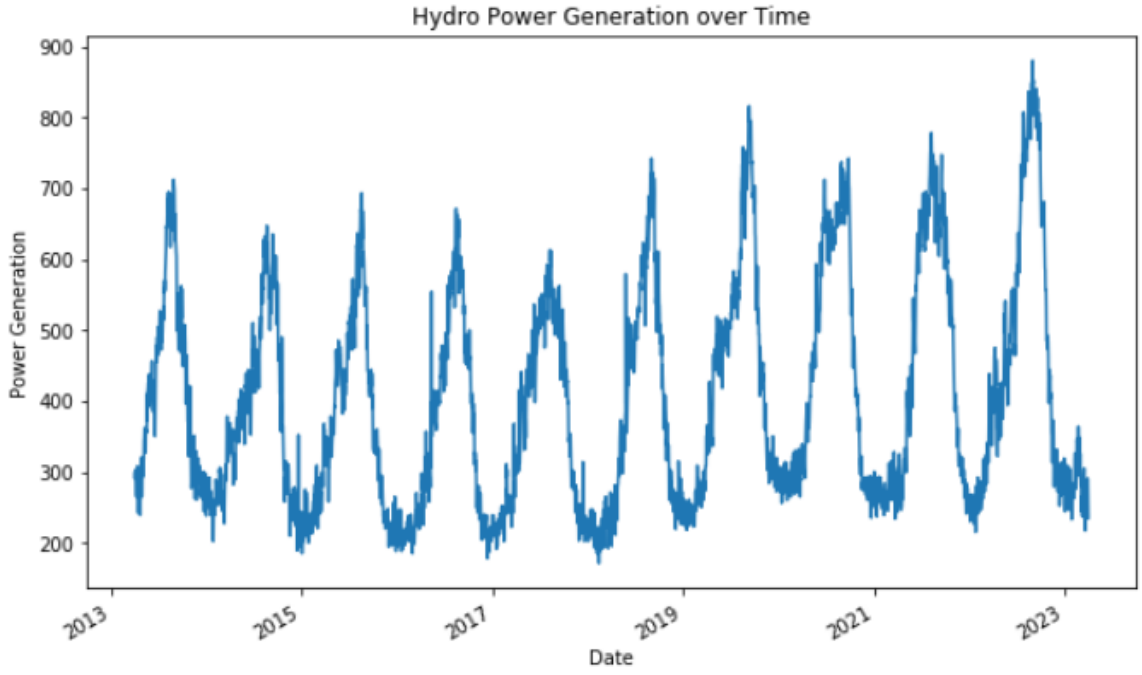
Regional Distribution: Examined the distribution of power generation by region using pie charts.

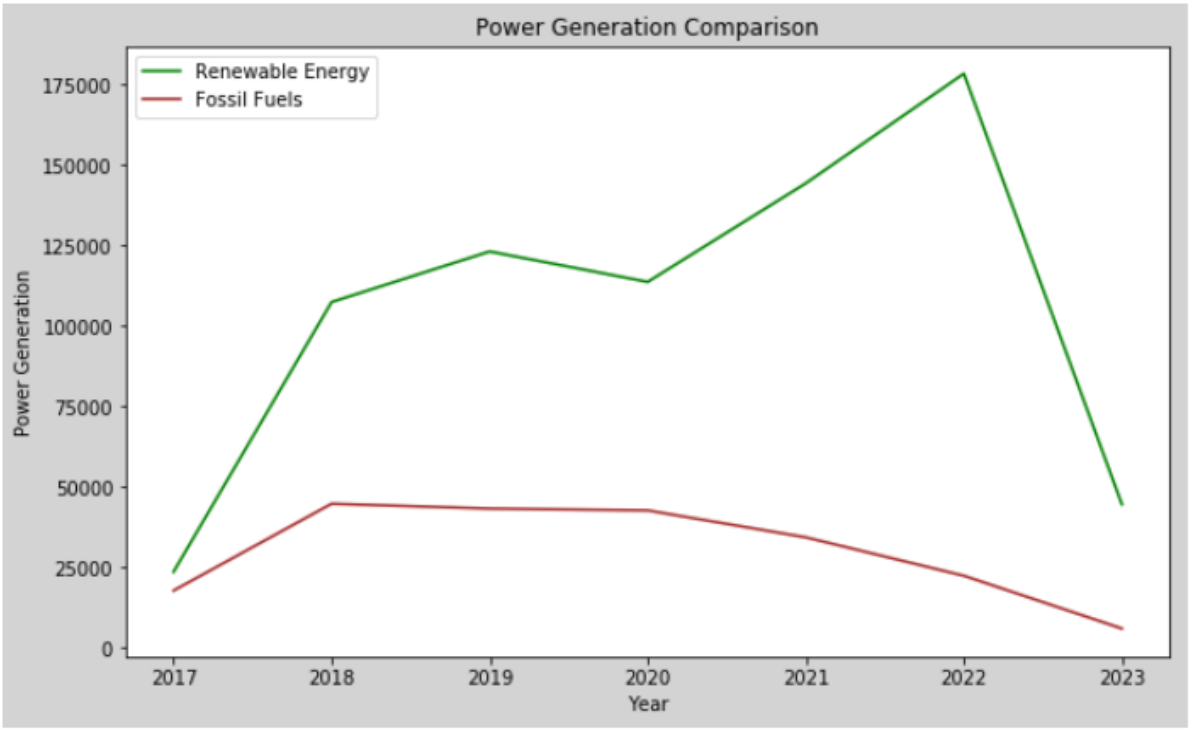
State-wise Analysis: Analysed power generation at the state level, identifying top states with the highest and lowest power generation.

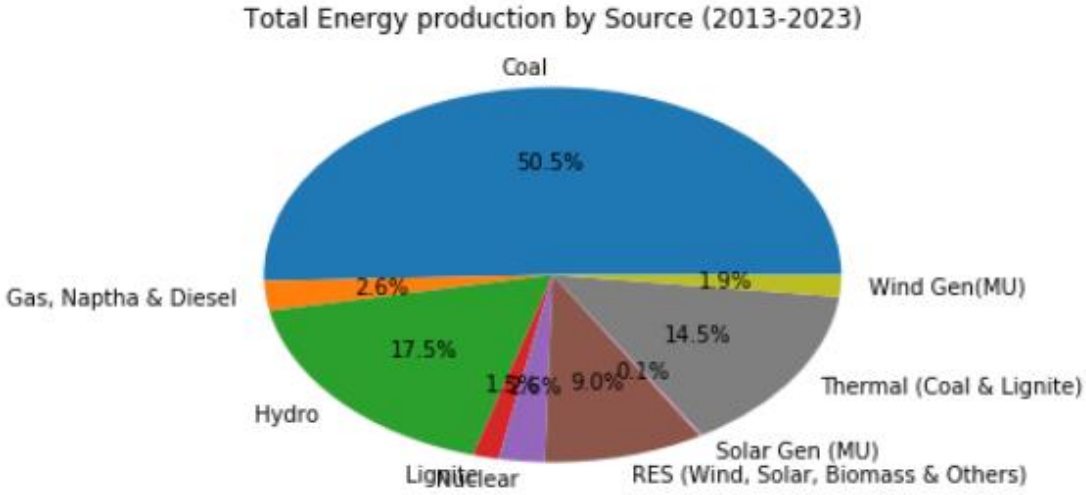
Charts and Visualizations:

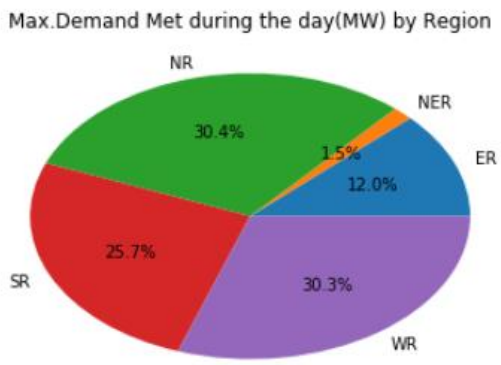


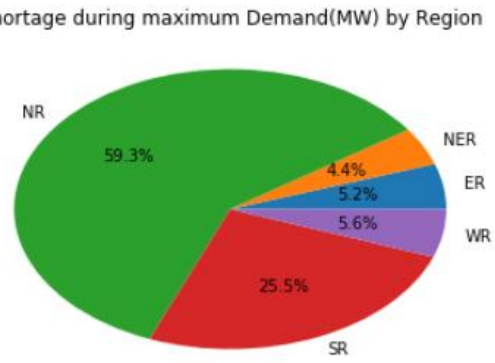

Power Generation by Source
Hydro
Wind Gen(MU)
Solar Gen (MU)
Nuclear
Thermal (Coal & Lignite)
Gas, Naptha & Diesel
RES (Wind, Solar, Biomass & Others)
Lignite
Coal
Source
0
500
1000
1500
2000
2500
Power Generation

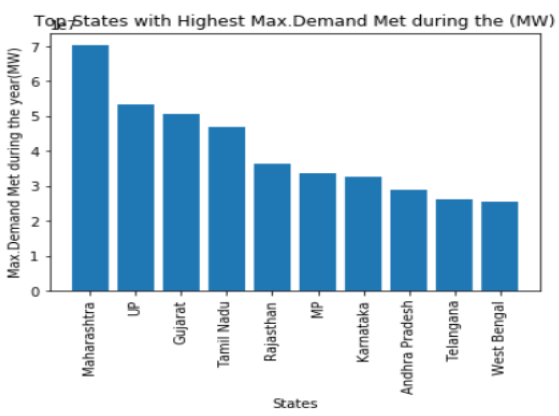

Hydro Power Generation over Time
900
800
700
600
500
400
300
200
Power Generation
2013
2015
2017
2019
2021
2023
Date

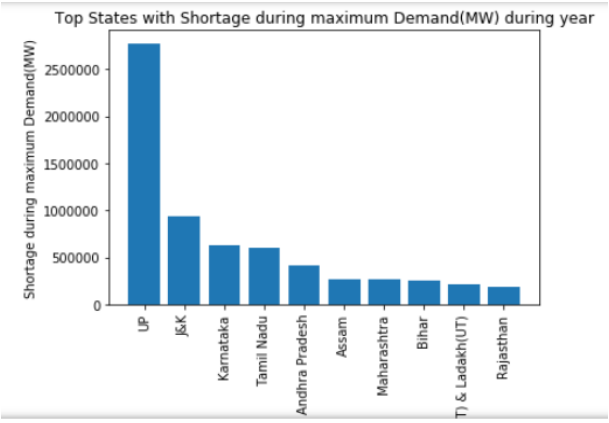

Power Generation Comparison
Renewable Energy
Fossil Fuels
Power Generation
175000
150000
125000
100000
75000
50000
25000
0
2017
2018
2019
2020
2021
2022
2023
Year


Total Energy production by Source (2013-2023)
Coal
50.5%
Gas, Naptha & Diesel
2.6%
Wind Gen(MU)
1.9%
14.5%
17.5%
0.1%
Thermal (Coal & Lignite)
Hydro
9.0%
Solar Gen (MU)
Lignite
Nuclear
RES (Wind, Solar, Biomass & Others)









Conclusion:

In conclusion, the analysis of electricity generation data in India provides valuable insights into regional and source-wise trends. Key findings highlight the significance of renewable energy sources and regional disparities in power generation capacity. These insights can inform policy-making and investment decisions to enhance energy security and sustainability in India.

THANK YOU

NITESHKUMAR G