



Energy
Solutions

Global Power Plant Analytics Report

Ezekiel Bernard Bassey

Introduction

The global demand for energy continues to rise, driven by population growth, urbanization, and industrialization. At the same time, the energy landscape is rapidly evolving with the growing adoption of renewable energy sources, policy shifts toward sustainability, and technological advancements in energy generation and distribution.

To support strategic decision making in this evolving context, this report presents a detailed analysis of global power plant data obtained from Kaggle. The dataset encompasses key information about power plants worldwide, including their geographic distribution, capacity, fuel types, generation output for 2021, ownership, and commissioning dates.

Using Power BI, I analysed this dataset with a focus on seven core objectives designed to uncover insights into energy generation trends, fuel composition, the expansion of renewables, and corporate ownership. The goal is to provide actionable intelligence for stakeholders such as energy planners, policymakers, sustainability advocates, and investors.

Dataset Overview

Dataset Summary

- **Source:** Kaggle Global Power Plant Dataset
- **Total Records:** 35,000 power plants
- **Columns Analysed:** 14
- **Primary Focus Year:** 2021
- **Coverage:** Global

Data Fields and Measurement

Column	Description	Data Type	Measurement Level
country code, country_long	ISO code & full country name	Categorical	Nominal
name of powerplant	Plant name	Categorical	Nominal
capacity in MW	Installed capacity	Numeric	Ratio
latitude, longitude	Plant location	Numeric	Interval
primary_fuel, secondary fuel, other_fuel2, other_fuel3	Fuel sources	Categorical	Nominal
start date	Year of commissioning	Date	Interval
owner of plant	Company owner	Categorical	Nominal
generation_gwh_2021	Actual generation	Numeric	Ratio
estimated_generation_gwh_2021	Estimated generation	Numeric	Ratio
geolocation_source	Source of geolocation data	Categorical	Nominal

Data Quality Assessment

Prior to analysis, several data quality issues were identified and addressed to ensure completeness, consistency, and reliability. The following steps were taken:

Cleaning and Transformation Steps:

1. **Secondary Fuel Columns**
 - Missing values in the secondary fuel and other_fuel2 columns were filled with "**Oil**", as it was a common and reasonable substitute fuel type.
 - The other_fuel3 column was **dropped** entirely due to containing only missing values.
2. **Start Date Column**
 - Missing values in the start date column were imputed with the **mode year**, which was **2016**, reflecting a recent and realistic commissioning date.
3. **Owner of Plant**
 - Null entries in the owner of plant column were replaced with the **most frequently occurring owner**, which was **Cypress Creek Renewables**.
4. **Generation Columns**
 - The generation_gwh_2021 column was retained for reference but not directly imputed.
 - The estimated_generation_gwh_2021 column had missing values replaced with the **mode value of 239.11 GWh**, ensuring consistency across estimates.
 - The other generation_gwh column was removed as it was completely blank.
5. **New Column Creation**
 - A derived column called Renew/nonRenew was created based on primary_fuel, allowing clear categorization of energy types into:
 - **Renewable**: Solar, Wind and Hydro.
 - **Non-Renewable**: Coal, Gas, Oil, etc.

Final Quality Summary Table:

Dimension	Evaluation
Completeness	Enhanced through strategic imputation and column drops
Consistency	Fuel types and ownership standardized
Validity	All numeric fields fall within logical boundaries
Accuracy	Geolocation and capacity values reflect original source; estimates used with caution
Integrity	Logical relationships (e.g., capacity vs generation) were maintained during cleaning

KPI Highlights

KPI	Value
Total Power Plants	35,000
Total Installed Capacity	5.7 million MW
Total Estimated Generation (2021)	8.35 million GWh

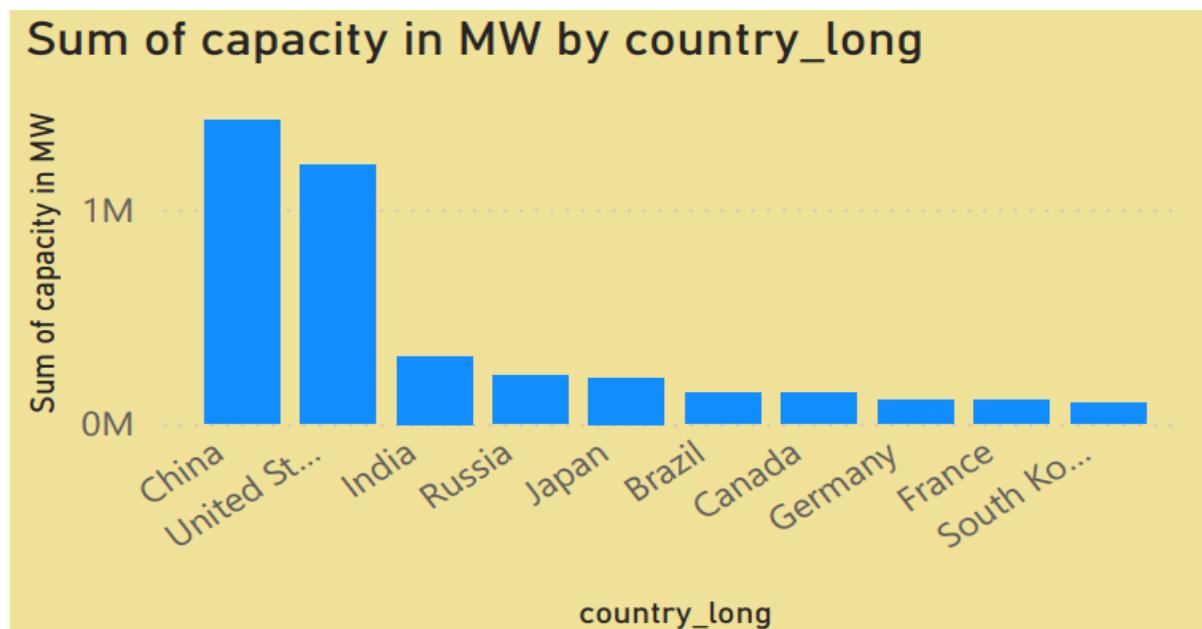
Objective 1: What is the Total Power Capacity and Generation by Country?

- Top Countries by Capacity (MW):

- CN China
- US United States
- IN India
- RU Russia
- JP Japan

Visuals:

- Column Chart: Total capacity by country

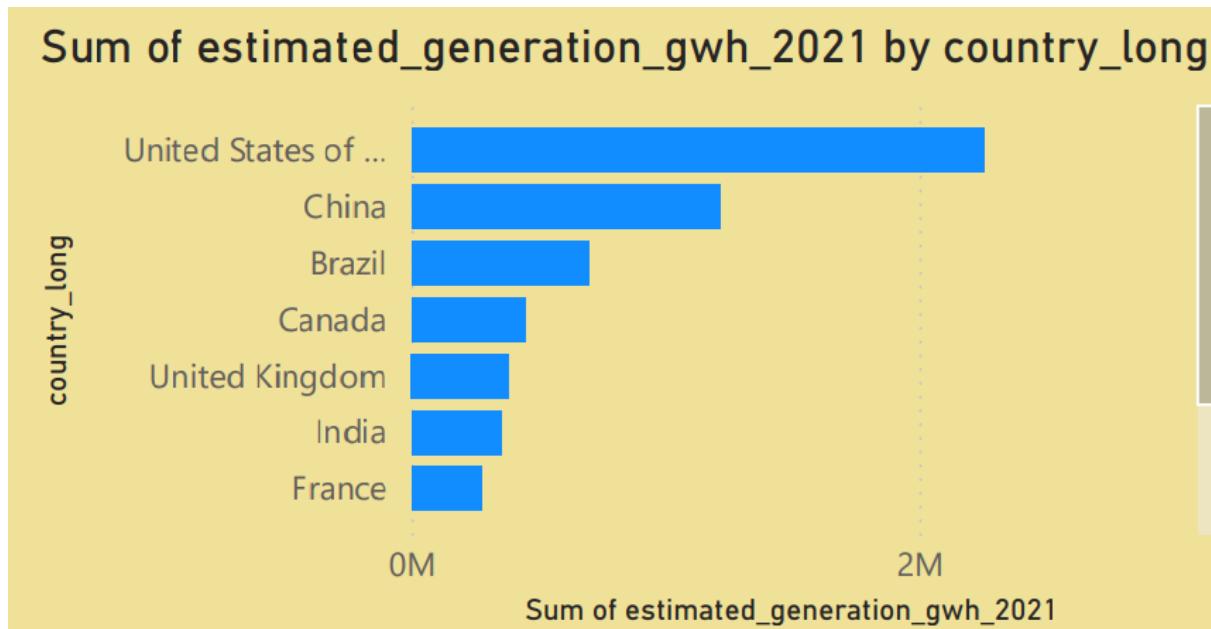


Objective 2: Which Countries Had the Highest Electricity Generation in 2021?

- United States generated the most electricity in 2021, followed by China, Brazil, and Canada.
- The Top 10 Countries collectively contributed the majority of global electricity output.

Visuals:

- Bar Chart: Estimated generation (GWh) by country

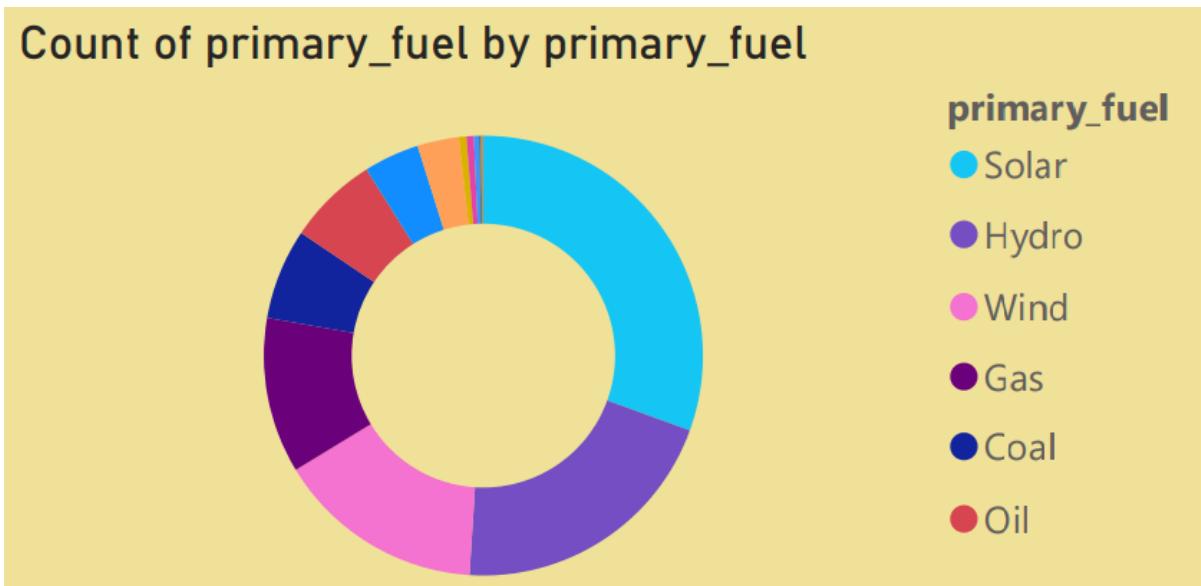


Objective 3: Which Fuel Types Are Used Most? Is Clean Energy Growing?

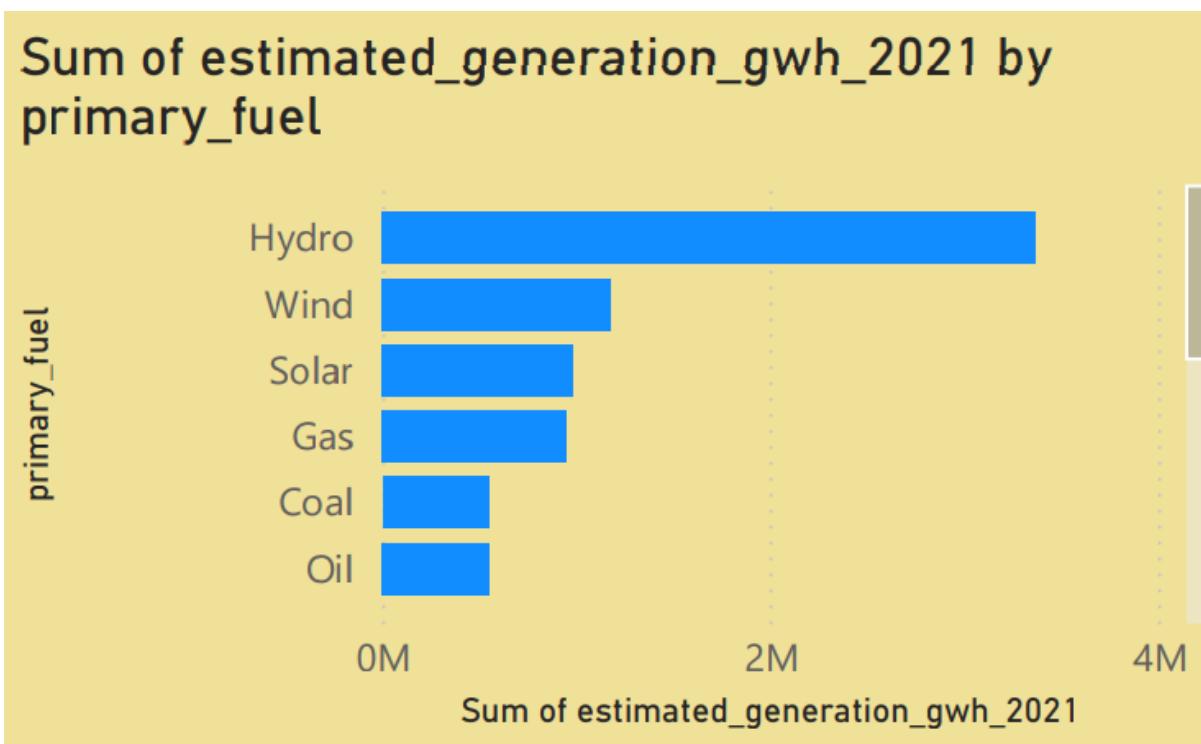
- **Most Common Fuel Types by Plant Count:**
 - **Solar** (highest)
 - **Hydro**
 - **Wind**
 - **Gas**
 - **Coal**
 - **Oil**
- **Top Fuel Types by Estimated Generation (2021):**
 - **Hydro** leads, followed by **Wind, Solar, Gas, Coal, and Oil**.
- Clean energy (Hydro, Wind, Solar) shows a **rising trend**, both in plant count and generation output.

Visuals:

- Donut Chart: Distribution of primary fuel types



- Stacked Bar Chart: Generation by fuel type



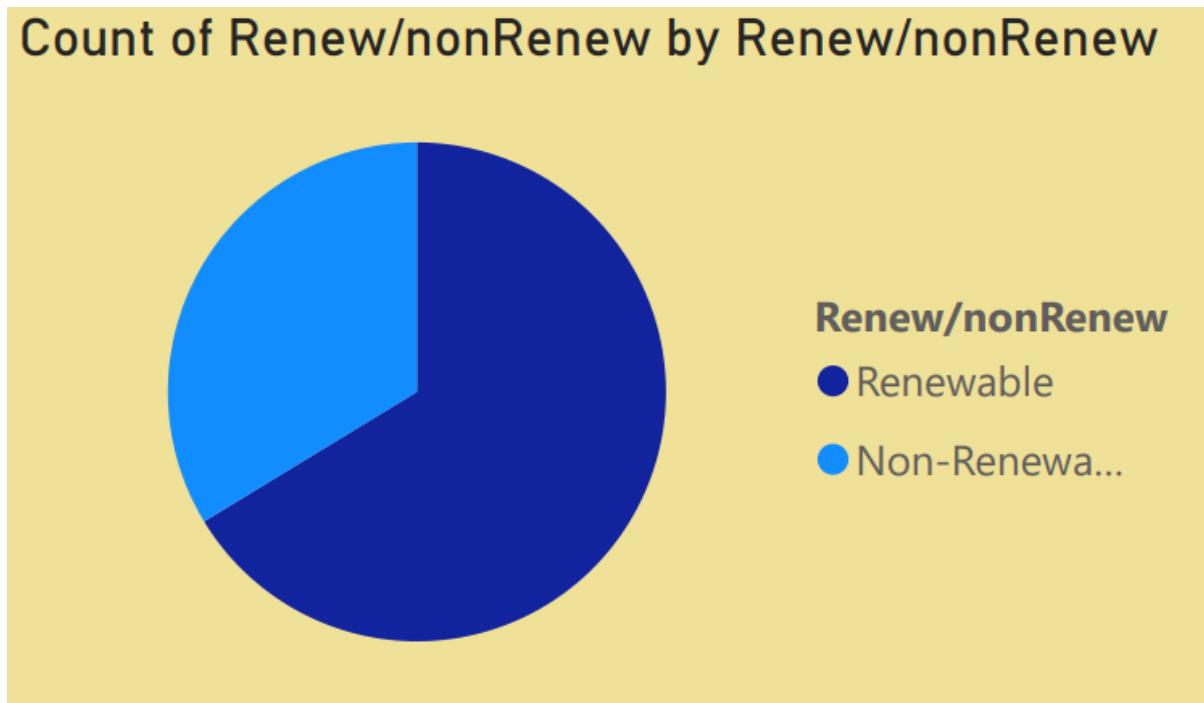
Objective 4: Where Are Renewable Energy Sources Located?

Key Insights:

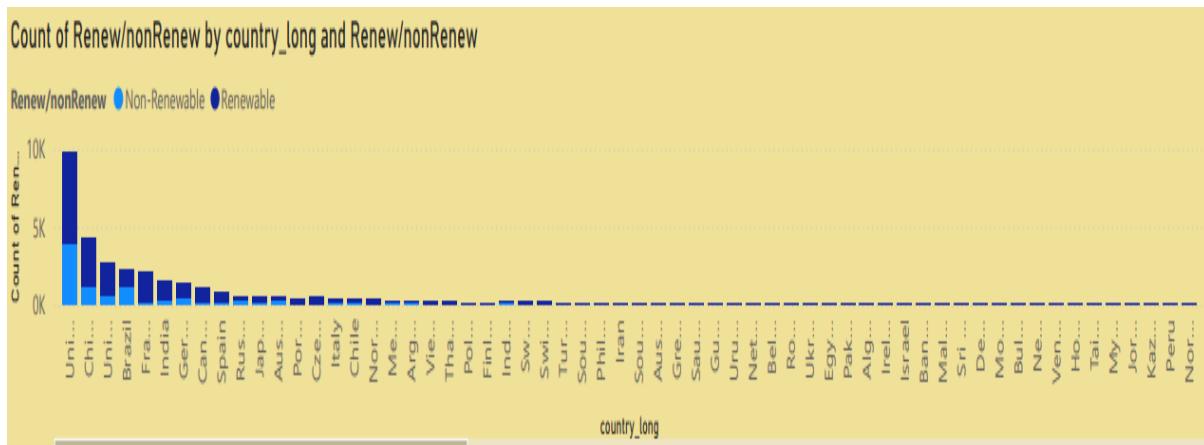
- **Solar Plants:** Concentrated in India, USA, and Spain.
- **Wind Farms:** Prominent in USA, Germany, and China.
- **Hydro Plants:** Dominant in Brazil, Canada, and Scandinavian countries.
- Countries such as **India, France, and Germany** show a **balanced energy mix** between renewables and non-renewables.

Visuals:

- Pie Chart: Count of Renewable Vs Non-renewable



- Stacked Column Chart: Renewable vs Non-renewable by country



Objective 5: Which Companies Own the Most Capacity?

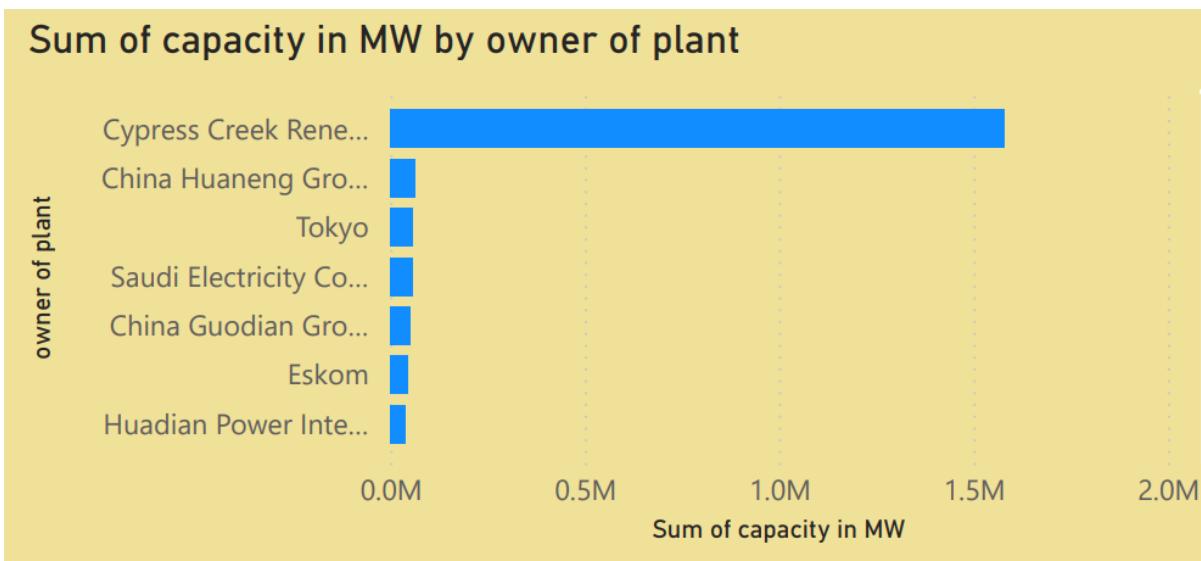
Top Plant Owners:

1. **Cypress Creek Renewables**
2. **China Huaneng Group**
3. **Tokyo Electric Power**
4. **Saudi Electricity Co.**
5. **China Guodian Group**

These companies each control significant installed capacity some exceeding **1.5 million MW**.

Visuals:

- Bar Chart: Capacity by owner



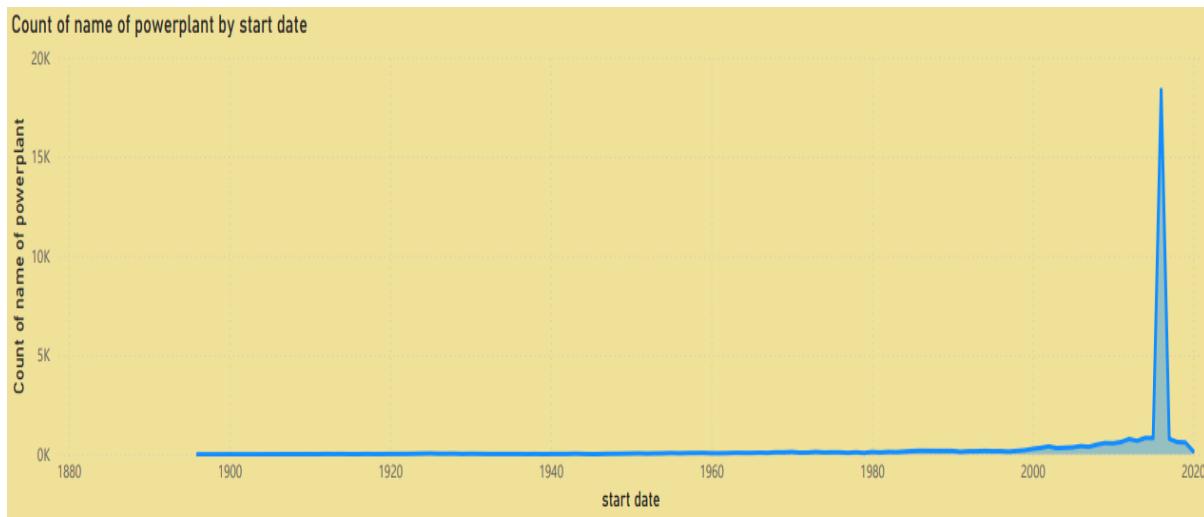
Objective 6: Power Plant Growth Over the Years

Key Insights:

- The **number of new power plants** spiked between **1980 and 2010**.
- A slight decline is observed after 2015, indicating a strategic shift toward **larger-capacity plants** (especially solar and wind).

Visual:

- Area Chart: Count of plants by year of start date



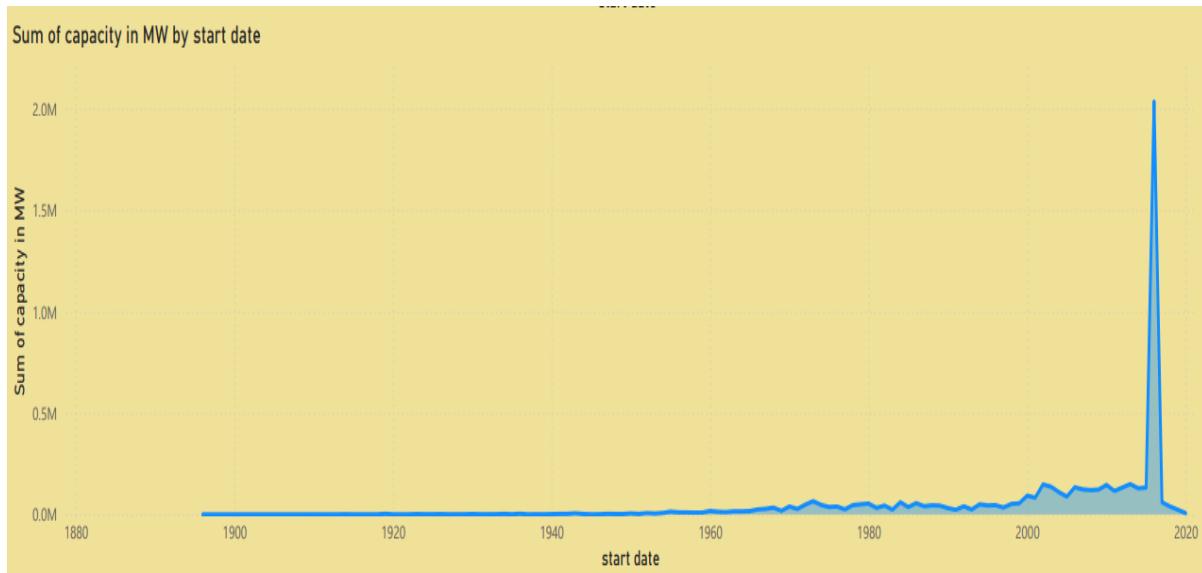
Objective 7: Power Capacity Trends Over the Years

Key Insights:

- **Total installed capacity** has grown rapidly, with major jumps seen during **2000–2015**.
- The capacity continues to rise, though fewer plants are being built — an indicator of a shift to **higher-efficiency and larger plants**.

Visual:

- Area Chart: Total capacity in MW by year



Conclusion

This analysis of global power plant data highlights significant progress in clean energy adoption and capacity expansion. Fossil fuels remain a dominant force in energy generation, but the transition to renewables is clearly underway, especially in high-capacity countries like the USA, China, and Brazil.

Notably, renewable energy plants outnumber fossil fuel plants by count, though not yet in output. Plant development peaked around 2010, and since then, investment seems to be shifting toward fewer but more powerful and sustainable installations.

Recommendations

- **Accelerate Renewable Energy Projects:** Encourage development of wind and solar plants, particularly in underrepresented regions.
- **Modernize Legacy Infrastructure:** Focus on retrofitting or retiring inefficient coal and oil plants.
- **Incentivize Green Investments:** Governments and investors should offer incentives for renewable technologies and grid expansion.
- **Prioritize Energy Mix Diversification:** Aim for balanced portfolios (hydro, solar, wind, gas) to ensure reliability and sustainability.
- **Leverage Predictive Analytics:** Use historical capacity trends to forecast future energy demands and plan investments accordingly.