# Project Documentation: Analysis of Nuclear Energy Dataset

## 1. Project Overview

This project aims to analyze data from the Kaggle dataset named 'Nuclear Energy Datasets' to derive meaningful insights. The process involves using Excel for initial data cleaning, MySQL for querying and data transformation, and Power BI for visualization. The project showcases the relationship between deaths during electricity production by different energy sources and other key metrics.

## 2. Dataset Details

Source: Kaggle dataset - 'Nuclear Energy Datasets'

## Analysis

### 3.1 ****Region with the Highest Nuclear Electricity Production****

* **Objective**: Identify regions with the highest contribution to nuclear electricity production.
* **Interpretation:**  
  The visualization highlights how nuclear electricity production is distributed across regions, emphasizing regions with the highest contribution. The data reveals that North America and Europe are the leading contributors, owing to their advanced nuclear infrastructure and energy policies favoring nuclear energy. These regions' dominance indicates their reliance on nuclear as a clean energy source to meet electricity demands while reducing carbon emissions. The graph also shows a stark contrast with regions like Africa and parts of Asia, where nuclear production is significantly lower, likely due to infrastructure or policy constraints.

### 3.2 ****Top Companies Owning the Most Reactors****

* **Objective**: Determine which companies dominate nuclear reactor ownership.
* **Interpretation:**  
  The column chart highlights the distribution of reactor ownership among major companies. While industry leaders like EDF (Électricité de France) and Rosatom are well-established in the sector, the visualization also reveals emerging players. Companies like Duke Energy,NextEra Energy, Inc., Dominion Generation, and Energy Harbor Corp. are noteworthy due to their growing investments in reactor technology. These emerging players have the potential to become key contributors to the nuclear energy market as they expand their operations and reactor portfolios.

### 3.3 ****Safety Metrics: Deaths per TWh of Electricity Production****

* **Objective**: Analyze safety by comparing death rates across primary energy sources.
* **Interpretation:**  
  The column graph highlights a striking disparity in safety metrics among energy sources. Nuclear energy stands out as one of the safest, with a death rate per TWh significantly lower than fossil fuels like coal, oil, and gas. This underscores nuclear energy's advantage in terms of public health and environmental safety. Renewable sources like solar and wind also exhibit extremely low death rates, further solidifying their role in the transition to sustainable energy. The graph provides an essential perspective for policymakers prioritizing energy safety while planning future investments in energy infrastructure.

### 3.4 ****Year-wise Electricity Generation & Nuclear Share****

* **Objective**: Examine year-wise trends in total electricity generation and nuclear’s share.
* **Interpretation:**  
  The upper column chart offers a clear temporal analysis of electricity generation trends. It shows steady growth in global electricity production, reflecting rising energy demands due to population and economic growth. Simultaneously, the share of nuclear energy demonstrates fluctuations, likely influenced by policy changes, technological advancements, and public opinion on nuclear energy safety. The graph provides a useful overview of how nuclear energy's role in the global energy mix evolves over time, offering insights into its future potential.

### 3.5 ****Global Capacity of Primary Fuels****

* **Objective**: Analyze the total capacity of various primary fuels globally.
* **Interpretation:**  
  The bar graph offers a comparative view of the capacity of primary fuels used for electricity generation worldwide. Fossil fuels like coal and gas dominate global capacity, while nuclear and renewables like solar and wind occupy smaller portions. This visualization reinforces the ongoing global dependency on fossil fuels despite efforts to transition toward cleaner energy sources. It also highlights the opportunity for renewables and nuclear energy to grow, especially in regions looking to diversify their energy mix and achieve sustainability goals.

### 3.6 ****Nuclear-Dependent Countries****

* **Objective**: Identify countries where nuclear energy contributes more than 30% of total capacity.
* **Interpretation:**  
  The top-left table highlights countries heavily reliant on nuclear energy. Nations like France, Slovakia, and Ukraine stand out as leaders in nuclear dependency, with nuclear energy contributing significantly to their total electricity capacity. These countries likely benefit from stable and sustainable energy systems supported by nuclear infrastructure. The visualization underscores the critical role nuclear energy plays in these nations’ energy security and their efforts to minimize reliance on fossil fuels.

## 4. Project Workflow

* **Step 1:** Data Cleaning  
  Null values were resolved in Excel by filtering relevant columns and logically entering missing data based on the datatype.  
  Column names were formatted using logical naming conventions with underscores (\_).
* **Step 2:** MySQL Querying  
  Script 1: Created multiple tables to organize the dataset.  
   Script 2: Executed queries to analyze patterns and generate insights. Challenges included finding nuclear-dependent countries where nuclear energy contributes more than 30% to their total capacity.
* **Step 3:** Visualization in Power BI  
  Created six visualizations to present insights interactively and intuitively.

## 5. Key Visualizations

5.1 Regions with the highest nuclear electricity production *(heatmap , table)*.  
5.2 Companies owning the most reactors ( *column graph*).  
5.3 Safety analysis: Deaths by primary fuels per TWh of electricity production *(column graph)*.  
5.4 Year-wise total electricity generation vs. nuclear fuel share *( column chart).*  
5.5 Global total capacity by primary fuels *(bar graph).*  
5.6 Nuclear-dependent countries where nuclear contributes >30% to total capacity *( matrix visualization ).*

## 6. Challenges Faced

* Dealing with null values: Resolved by filtering and entering missing data logically in Excel.
* Formatting column names: Used consistent naming conventions for clarity.
* Writing complex SQL queries: Found patterns through trial and error and referenced notes to solve issues with nuclear dependency analysis.

## 7. Project Highlights

* Combined Excel, MySQL, and Power BI for effective data processing, querying, and visualization.
* Interactive dashboard features include filters, drill-throughs, and dynamic visualizations to enhance user engagement.

## 8. Future Scope

* Work on more advanced datasets with complex queries to derive deeper insights.
* Explore advanced Power BI features and integrate Python or R for data analysis.

## 9. Conclusion

This project demonstrates the ability to manage a complete data pipeline from cleaning to visualization, highlighting a strong relationship between tools like Excel, MySQL, and Power BI. It effectively derives insights and patterns from data.

The insights derived from these visualizations effectively demonstrate the critical role of nuclear energy in the global energy mix. From its safety metrics to its share in total electricity generation and the regions or companies dominating the sector, the data provides a comprehensive overview of nuclear energy's impact. These analyses are valuable for policymakers, energy companies, and researchers focusing on sustainable energy transitions.