Electric Disturbance Events Analysis

Report

Prepared by

INSIGHT INNOVATORS

- Abdulrahman Haider.
- A'laa Taha.
- Amany Mohamed.
- Nashwa Ahmed Mohamed.

I. Content:

1. **Project excerpt**
2. **Cleaning & Assumptions**
3. **Modeling**
4. **KPI**
5. **Design**
6. **Dashboard**
7. **Analysis**

II. Project excerpt

1. Introduction

- **Project Overview**:
  This project aims to analyze Electric Disturbance Events in different NERC (North American Electric Reliability Corporation) regions across the USA, focusing on the impact of these events in terms of demand loss (MW) and the number of customers affected. The analysis utilizes historical data on power outages, categorized by NERC region, event type, and other key factors. The goal is to identify patterns and provide insights into the frequency, duration, and magnitude of electricity disruptions.
- **Objective of the Cleaning Phase**:
  The data cleaning phase is a critical step in preparing the dataset for analysis. Due to the complexity of the data, which includes multiple regions, inconsistent formats, and missing values, this phase aims to ensure accuracy, completeness, and consistency. By addressing these issues, the data will be better suited for generating reliable insights using Power BI. Key tasks include handling missing data, correcting regional inconsistencies, standardizing formats, and resolving duplicate or irrelevant records.

2. Data Sources

The dataset used in this project consists of a fact table and several dimension tables. These tables capture various attributes related to electricity outages across different NERC regions.

- **Fact Table (fData)**:
  This table contains essential columns such as:
  - Event Began
  - Restoration
  - NERC Region
  - Event Type
  - Demand Loss (MW)
  - Number of Customers Affected

- **Dimension Tables**:
  - **dEvent**: A table that provides details about different types of events and a conditional column for event categorization.
  - **dNERC**: A dimension table listing NERC regions.
  - **dCalender**: A dimension table listing dates.

III. Cleaning & Assumptions

3. Cleaning Methodology

- **Data Validation**:
  Steps were taken to validate the data and identify errors, including:
  - Ensuring all dates in "Event Began" and "Restoration" were valid and in chronological order.
  - Checking for missing or incorrect values in critical columns like Demand Loss (MW) and Number of Customers Affected.
  - Checking NERC region for each event

- **Replacement Rules:**

  - **General Replacements**:
    - Ongoing, Unknown, UNK, NA, None, -, --, Interruptible Tariff 1-6 customers, Major Industrial Customer Load Reduction → **Null**
    - GRE (1,900) Total 11,175 → **11,175**
    - PGE → **5.5 Million**
    - Noon, Evening → **PM**
    - Midnight → **AM**
    - (Number) Approximately, (Number) peak, Greater than, (Number) Over, (Number) Under → **Number**
    - (Num1 - Num2) → **Maximum**
    - 5:78 PM → **5:58 PM**
    - 2077 → **2011**
  - **Event Type Categorization**: Events were grouped into the following categories:
    1. Weather.
    2. Technical.
    3. Security.
    4. Supply.
    5. Public.
    6. Misc.
    7. Special Event.
  - **NERC Region Replacements**:
    - Regions with multiple areas → **Multiple**
    - ECAR, MAAC, WeEnergies, MAIN, Midwest ISO (RFC) → **RFC**
    - NPPC → **NPCC**
    - MAPP, MR0 → **MRO**
    - WSCC, HI, HECO, MECO → **WECC**
    - REC → **SERC**
    - N/A, Null → **PR**
    - NP → **NPCC**
    - TE → **TRE**
    - SPP RE → **SPP**

- **Data Transformation:**

  - **Standardizing Formats**:
    Formats across the tables were standardized, such as:

    - Converting date formats to MM/DD/YYYY.
    - Ensuring consistency in "MW" formatting.

  - **Categorization of Multiple NERC Regions**:
    Rows with multiple NERC regions were handled as follows:

    - Regions like "SERC, MECO" or "WECC/NPCC/RF" were categorized as "Multiple" for consistency.
    - Assumption: Events impacting multiple regions could be analyzed as one for high-level insights.

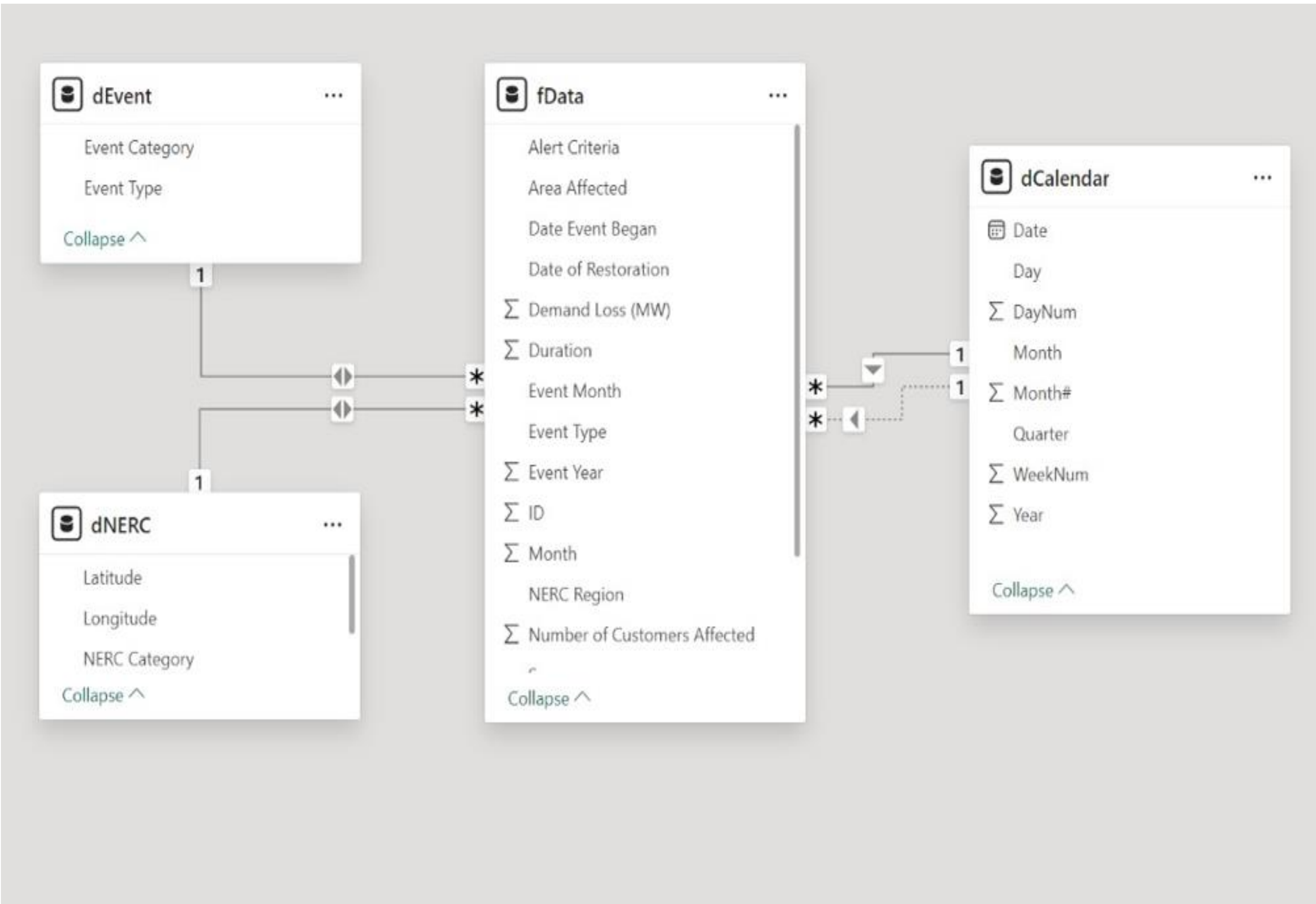
- **Data Integration:**

  - **Append Tables**:
    Append was initially avoided due to regular change of data collection system, so we started with cleaning for each year individually to have a similar data frame across all tables and then append was done starting 2002 to 2023.

- **Challenges and Solutions**

  - **Handling Complex NERC Regions**:
    Complex NERC regions were addressed by creating the "Multiple" category to simplify the analysis.
  - **Data Inconsistencies**:
    Data quality issues were resolved using the provided PDFs, external validation for events with unreasonable/unknown numbers.

IV. Modeling


dEvent
Event Category
Event Type
Collapse
1
fData
Alert Criteria
Area Affected
Date Event Began
Date of Restoration
Demand Loss (MW)
Duration
Event Month
Event Type
Event Year
ID
Month
NERC Region
Number of Customers Affected
Collapse
dCalendar
Date
Day
DayNum
Month
Month#
Quarter
WeekNum
Year
Collapse
1
1
1
dNERC
Latitude
Longitude
NERC Category
Collapse

V. KPI

- **Outage frequency and trends: -**
  - How has the number of power outage events changed over time (from 2002 to
  - 2023)? (# Events over time)
  - What are the peak months or seasons for power outages? (Seasonality)
  - Which years had the highest number of outages?
  - Is there a trend in increasing or decreasing outages over time?

- **Regional and area-specific insights: -**
  - Which areas or NERC regions experience the most outages?
  - Are there regions where outages are more frequent during certain months?
  - Which event is the most repeated per region?
  - Which region has the highest number of affected customers?

- **Outage duration and restoration efficiency: -**
  - What is the average duration of outages (difference between Date/Time Event Began and Date/Time of Restoration)?
  - Which regions have the longest outage durations?
  - Has restoration time improved over the years?

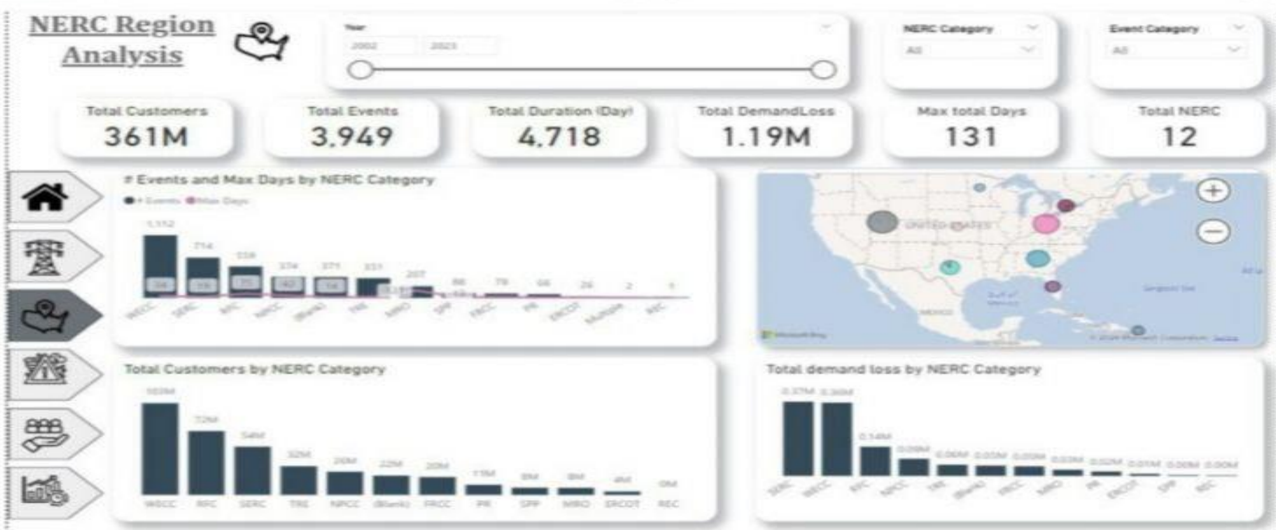
- **Impact on customers and demand: -**
  - What is the total number of customers affected by outages over the years?
  - What is the average number of customers affected per outage event?
  - What is the total demand loss (MW) during outages?

VI. Design


DOE Analysis: (2002 - 2023)
Slicer with Date
Slicer with events type / NERC
Total Affected Customers
Total # Events
Total Duration
Demand Loss
Max Days
Regions
Regions Donut
Events Types Donut
Duration (categories) Over Time
# of events over time
DOE Analysis: (2002 - 2023)
Slicer with Date
Slicer with events type / NERC
Total Affected Customers
Total # Events
Total Duration
Demand Loss
Max Days
Regions
Events per NERC
Map for NERC
Total Customers / NERC
DM Loss Over NERC
DOE Analysis: (2002 - 2023)
Slicer with Date
Slicer with events type / NERC
Total Affected Customers
Total # Events
Total Duration
Demand Loss
Max Days
Regions
Seasonality of events
Map for NERC
Avg DM Loss by event
Total DM Loss by event
DOE Analysis: (2002 - 2023)
Slicer with Date
Slicer with events type / NERC
Total Affected Customers
Total # Events
Total Duration
Demand Loss
Max Days
Regions
Seasonality of # Affe. Customers
Total Affected Customers by year
Affected Customer By NERC
Total Customer by DM loss

I. Dashboard


General
Total Customers
361M
Total Events
3,949
Total Duration (Day)
4,718
Demand Loss (MW)
1.19M
Max total Days
131
Total NERC
12
Top 4 NERC region by # Events
Events Categories Distribution
Events over Years
Total Duration for Top 5 Events by year


NERC Region Analysis
Total Customers
361M
Total Events
3,949
Total Duration (Day)
4,718
Total DemandLoss
1.19M
Max total Days
131
Total NERC
12
Events and Max Days by NERC Category
Total Customers by NERC Category
Total demand loss by NERC Category


Event Type Analysis
Total Customers
361M
Total Events
3,949
Total Duration (Day)
4,718
Demand Loss (MW)
1.19M
Max total Days
131
Total NERC
12
Total of Events by Season
Seasonal Event Type Analysis
Impact of Event Types on Customers
Demand Loss by Season
Demand Loss by Event Type
Avg. Duration by Event Type


Customers Analysis
Year
2002
2023
NERC Category
All
Event Category
All
Total Customers
361M
Total Events
3,949
Total Duration (Day)
4,718
Demand Loss (MW)
1.19M
Max total Days
131
Total NERC
12
Total Customers Affected by NERC
NERC Category
WECC
RFC
SERC
TRE
NPCC
(Blank)
FRCC
PR
SPP
MRO
ERCOT
REC
103M
72M
54M
26M
22M
20M
32M
11M
Affected customer Seasonal Analysis
54.71M (15.15%)
133.65M (37%)
77.81M (21.54%)
95.05M (26.31%)
Season
summer
fall
winter
spring
Total Customers by Demand Loss (MW)
0M
5M
10M
15M


Analysis Over Time
Year
2002
2023
NERC Category
All
Event Category
All
Total Number of Customers Affected by Year
2005
2010
2015
2020
Total Demand Loss (MW) by Year
2005
2010
2015
2020
Events and Count of NERC Category by Quarter
Events
Count of NERC Category
Qtr 1
Qtr 2
Qtr 3
Qtr 4
Average of Duration by Year
Average of Duration
2
0
2005
2010
2015
2020
Events and Count of NERC Category by Month
Events
Count of NERC Category
January
February
March
April
May
June
July
August
September
October
November
December

II. Analysis

General analysis:

What questions can be answered using the analysis? KPIs to be calculated

- **Outage frequency and trends: -**

How has the number of power outage events changed over time (from 2002 to 2023)? (# Events over time)

1. 2023 has significant improvement in # of events compared to the last 4 years.
2. Despite the reported data counts 3949 events, almost 1300 events were reported without reason (Blank).

What are the peak months or seasons for power outages? (Seasonality)

1. July, August, September shows seasonality over the years with the highest number of events in most of the years

Which years had the highest number of outages?

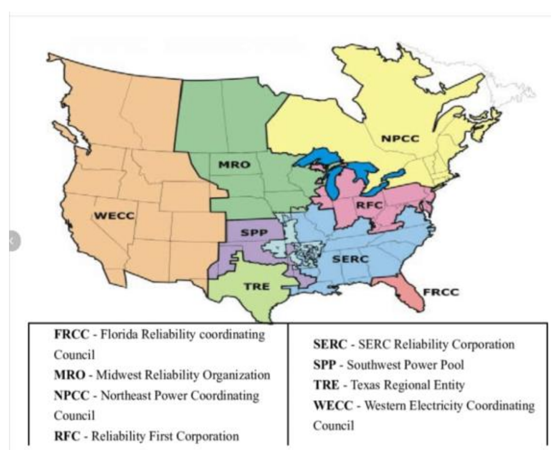
1. (2020 – 2022)

Is there a trend in increasing or decreasing outages over time?

1. 2011 marked as peak of events over 17 years (2002 – 2019) and due to high number of Weather events occurred in this year.
2. 2013 passed safely without any weather events.

Regions Analysis:

Over the years, **WECC has recorded the highest number of events and affected customers** among NERC regions, which comes as no surprise given that it covers the **largest geographic area** within NERC. Currently, **WECC spans approximately 1.8 million square kilometers (700,000 square miles)** across the western United States, Canada, and northern Mexico. This accounts for **about 48% of the U.S. land area**, making it the most extensive region managed by NERC in terms of size.



WECC oversees the largest geographic area of any reliability council in North America, spanning across:

- **Western U.S.**: California, Arizona, Nevada, Utah, Colorado, Washington, Oregon, and others
- **Western Canada**: British Columbia, Alberta
- **Mexico**: Northern Baja California

This coverage includes major metropolitan areas like **Los Angeles, San Francisco, Seattle, Phoenix, Denver**, and **Las Vegas**. These cities rely heavily on WECC's infrastructure to support their populations and industries.

Frequent and Significant Events: Security Attacks and Outages:

From the project's data, it is clear that **security-related incidents** have emerged as the most frequent disruptions in recent years. These security attacks range from **cyberattacks targeting the grid's digital infrastructure** to **physical threats on substations and transmission lines**.

1. **Southwest Blackout (2011)**
   - Although primarily caused by human error, it revealed the vulnerabilities of interconnected systems. This incident led to heightened awareness of how intentional attacks could exploit these weaknesses.
2. **Cyberattacks on Utilities (2015-2023)**
   - Multiple utilities in WECC's region have faced cyber intrusions aimed at disrupting operations or stealing sensitive data. These attacks highlight the increasing reliance on digital infrastructure and the need for stronger cybersecurity defenses.
3. **Metcalf Substation Attack (2013)**
   - A well-coordinated physical attack on a power substation near San Jose, California, caused significant damage and raised concerns about the security of critical infrastructure.
4. **California Blackouts (2020-2023)**
   - During rolling blackouts caused by wildfires and heatwaves, utilities also faced phishing attacks and attempts to breach operational networks, complicating recovery efforts.

Significant weather Events in WECC's History:

1. **California Energy Crisis (2000-2001)**
   - Though predating WECC's official formation, the energy crisis had lasting effects on policies. Market manipulation by companies like Enron led to rolling blackouts, exposing weaknesses in grid management and market structures.
2. **Southwest Blackout (2011)**
   - A transmission line failure in Arizona triggered a cascading outage that affected **2.7 million people** across parts of Arizona, California, and Baja California. This event highlighted the challenges of managing interconnected systems.
3. **California Wildfires and Rolling Blackouts (2020)**
   - High temperatures and wildfires caused massive strain on the grid, leading to **rolling blackouts** affecting millions. Wildfires also damaged critical infrastructure, raising concerns about grid resilience under extreme conditions.
4. **Heatwaves and Grid Instability (2022-2023)**

- Record-breaking heat across the western U.S. increased energy demand, forcing utilities to issue emergency alerts and implement energy-saving measures. The events exposed the region's vulnerability to climate change.

NERC Region and Restoration Efficiency:

Regional Variance in Outage Impact:

The WECC region continues to show the highest frequency of events and customer impact, which highlights its vulnerability to outages.

This region should be prioritized for infrastructure improvements and grid strengthening.

Disparity in Outage Restoration:

Some regions, such as RFC, WECC and SERC (Top 3) experience longer restoration times. Reducing restoration times in these areas can significantly minimize customer inconvenience and economic loss.

Insight:

WECC and SERC regions should be prioritized for infrastructure upgrades and faster restoration plans. SERC and RFC regions, in particular, require focused efforts to reduce restoration times to minimize customer impacts.

Event Type Analysis:

Weather Events Dominate:

Weather-related events remain the most frequent outage cause, contributing to both a higher number of total events and having the longest duration. This highlights the vulnerability of the grid to weather-related disruptions.

High Demand Loss in summer:

Summer still accounts for the largest portion of demand loss (34.4%), largely due to higher cooling demands, leading to potential grid overloads.

Winter also shows significant demand loss (33.07%), driven by heating demand and weather-related events.

Event Frequency by Season:

Both summer and winter show the highest number of outage events, corresponding to the periods of peak energy consumption and extreme weather.

Insight:

Seasonality remains a critical factor in outages, with the highest demand loss and event counts in summer and winter.

The prolonged duration of weather-related outages points to the need for infrastructure improvements to withstand extreme weather conditions more effectively.

Customers Impact Analysis:

Large Customer Base Affected:

Regions like WECC and RFC still have the largest number of affected customers, indicating that outages in these areas have broader impacts, likely due to the larger population centers or critical infrastructure.

Summer Outages Affect the Most Customers:

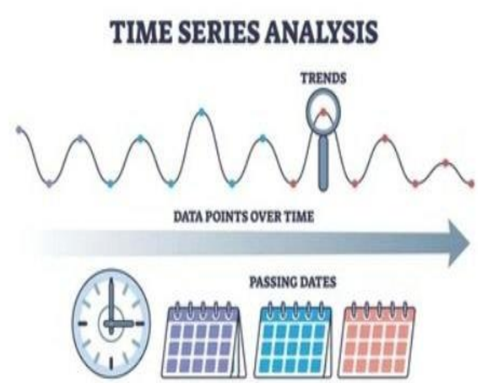
Summer continues to affect most customers, with 37% of the customer impact occurring during this season.

Customer Demand Loss Trend:

The scatter plot shows a positive correlation (direct relationship) between the number of affected customers and the demand loss in MW. Larger outages tend to result in higher demand loss.

Insight:

WECC and RFC regions are significantly affected, representing a large portion of impacted customers. Summer outages continue to cause the most widespread customer disruptions and demand loss, emphasizing the need for stronger grid stability and preparedness during high-demand periods.



Analysis Over Time:

Increasing Trend in Affected Customers in Certain Years:

Some years show a sharp increase in the number of affected customers, which could indicate more severe outages, possibly due to extreme weather events or increasing infrastructure strain.

Demand Loss Peaks in Specific Years:

There are spikes in demand loss in certain years, suggesting that some events were more catastrophic, possibly due to major weather events or systemic failures.

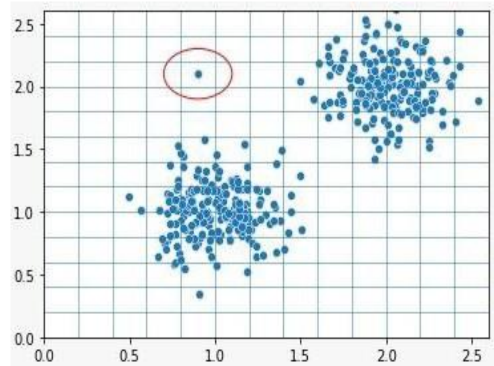
Quarterly and Monthly Patterns:

Outage is more frequent in Q3 (summer months) and Q4 (winter months), with noticeable peaks in July, August, and December. This confirms the cyclical and seasonal nature of outages.

Insight:

Outages exhibit a seasonal and cyclical trend, with peaks in summer and winter. Specific years stand out with more severe outages, likely driven by extreme weather or infrastructure challenges.

Over time, with the advancement of technology and improved response efforts, the average outage restoration time has decreased. This indicates that restoration processes have become more efficient, leading to quicker recovery from outages.



Outliers:

- **Total Demand Loss by Year:**

  The line chart for demand loss shows sharp spikes in certain years. These spikes represent years when demand loss was much higher than the typical levels, likely due to major events like extreme weather or large-scale power grid failures.

- **Total Number of Customers Affected by Year:**

  In some years, there is a noticeable sharp increase in the number of affected customers. These years can be outliers, where certain large-scale events impacted a disproportionately high number of customers.

Recommendations:

□ **Weather-Related Events (Wildfires, Heatwaves, Storms)**

- **Underground Power Lines:** Bury lines in fire-prone areas to prevent wildfire damage.
- **Vegetation Management Programs:** Regularly clear vegetation around power lines to reduce fire risks.
- **Smart Grid Technology:** Use automated systems to isolate faults and restore power faster during extreme weather events.
- **Demand Response Programs:** Encourage energy conservation during heatwaves through incentives for customers.

□ **Infrastructure Failures (Equipment Malfunctions, Aging Assets)**

- **Proactive Maintenance:** Implement predictive maintenance with IoT sensors to detect equipment failures before they happen.
- **Grid Modernization:** Upgrade aging infrastructure to improve efficiency and reduce downtime.
- **Redundancy Planning:** Develop backup systems for critical assets to prevent cascading failures.

Seasonal Preparedness is Key:

- The consistent seasonal patterns of outages highlight that summer and winter months are critical periods. Utilities should prioritize reinforcing infrastructure, weatherproofing, and boosting restoration capabilities before these high-demand seasons.

 Cybersecurity Threats

- **Zero-Trust Architecture**: Limit access to systems based on user roles to minimize the impact of breaches.
- **Employee Training Programs:** Regularly train staff on recognizing phishing attempts and following cybersecurity protocols.
- **Penetration Testing:** Conduct simulated cyberattacks to find and fix vulnerabilities before they are exploited.

 Physical Security Threats (Vandalism, Sabotage)

- **Surveillance and Monitoring:** Deploy cameras, motion sensors, and drones to monitor critical infrastructure.
- **Collaborate with Law Enforcement**: Establish strong partnerships with security agencies to respond swiftly to attacks.
- **Harden Substations:** Reinforce physical barriers around substations and key equipment to deter sabotage**.**

 Energy Demand Surges

- **Energy Storage Solutions:** Install large-scale battery systems to store excess energy for peak demand periods.
- **Interregional Energy Exchange:** Establish agreements with neighboring regions to share power during shortages.
- **Public Awareness Campaigns:** Promote energy-saving behaviors to reduce unnecessary consumption during peak times.