Power Outage Trends in California, Florida, Pennsylvania, and Texas from 2015 to 2023

https://github.com/sarahj-hall/Antonucci_Hall_Huang_Kuehn_ENV872_FinalProject.git

December 12th, 2025

Contents

List of Tables

List of Figures

1 Rationale and Research Questions

Electric power outages are costly for both utilities and their customers. Outages disrupt economic activity, and impact critical facilities such as hospitals (Wing et al., 2025). The United States electric grid is increasingly vulnerable to outages due to increasing extreme weather events, such as hurricanes, heatwaves, wildfires, and winter storms, which are increasing in frequency and intensity due to climate change (Wendel, 2024). Because many of these events occur during specific times of the year, understanding weather power outages follow a seasonal pattern is important for planning and resilience. If outages tend to occur in specific months, utilities and communities can better allocate resources, resilience investments, and emergency preparedness.

At the same time, electric utilities have made advancements in system planning, grid hardening, and outage detection (IBM, 2025). Together this raises key questions about whether outages are becoming more or less frequent, whether they show predictable seasonal cycles, and whether their impacts are increasing over time.

To address these questions, this study focuses on Texas, California, Florida, and Pennsylvania. These states represent diverse geographic regions, climates, and electric power systems. For each state, the study examines trends in frequency of power outages, seasonality, and severity of outages. Customer weighted hours of outages will be used to quantify the impact of outages, as it combines the number of customers affected with the outage duration.

The research questions of this study are:

- Question 1: How has the frequency of power outages changed over time?
- Question 2: Is there a seasonal trend? Are certain months more prone to outages?
- Question 3: How has the severity of power outages changed over time?

2 Dataset Information

The Event-correlated Outage Dataset in America by the Pacific Northwest National Laboratory was downloaded from the Open Energy Data Initiative (OEDI) (https://data.openei.org/submissions/6458). The dataset includes an aggregated and event-correlated analysis of power outages in the United States. The specific dataset selected for this analysis is the Aggregated Outage Data which integrates data from the Environment for the Analysis of Geo-Located Energy Information (EAGLE-I), and Annual Estimates of the Resident Population for Counties 2024 (CO-EST2024-POP). The EAGLE-I dataset, provides county-level electricity outage estimates at 15-minute intervals from 2014 to 2023. It encompasses over 146 million customers, but this coverage has increased over time from 137 million in 2018. EAGLE-I only started providing data quality estimates starting in 2018. The Aggregated Outage Dataset provides monthly outage data at the state level, including total number of outages, the maximum duration of outages, and the customer weighted average of outages.

The data was wrangled by combining the yearly data from 2015 to 2023 into one dataframe. The year 2014 was removed from the analysis because it did not have monthly data, only the yearly summary. From this file, four datasets were created by filtering for each state (CA, FL, PA, and TX). For each state, the monthly value equal to 0 was filtered out, which represented the yearly summary. Additionally, a date column was added that combined the monthly and yearly columns into a date object.

This data structure shown in Table 1 applies to all four state datasets. Table 2 summarizes the key statistical characteristics of each datasets.

Table 1: Summary of Outage Dataset Structure

Variable	Description	Units
state	Two-letter state abbreviation	N/A
year	Year of outage	N/A
month	Month of outage	N/A
outage_count	Number of outages per month	N/A
max_outage_duration	Longest outage duration in a month	Hours
customer_weighted_hours	Customer-weighted outage hours	N/A
date	Date of outage	N/A

Table 2: Summary Statisitcs of CA, FL, PA, and TX Datasets

State	Outage Count Range	Outage Count Mean	Max Duration Range	Max Duration Mean	Customer Weighted Hours Range	Customer Weighted Hours Mean
CA	140–2599	900.55	27.25–609.75	131.82	491034.75–88045909.75	6,151,189
FL	402–1515	970.73	11.5–532.75	56.80	438767.25–532294703.25	9,390,167
PA	336–1312	600.20	13.5–140.5	46.74	333846–28965544.5	1,970,894
TX	120–2361	1,108.82	6–723.75	76.26	167452–227923742	5,712,808

3 Exploratory Analysis

Insert exploratory visualizations of your dataset. This may include, but is not limited to, graphs illustrating the distributions of variables of interest and/or maps of the spatial context of your dataset. Format your R chunks so that graphs are displayed but code is not displayed. Accompany these graphs with text sections that describe the visualizations and provide context for further analyses.

Each figure should be accompanied by a caption, and each figure should be referenced within the text.

Scope: think about what information someone might want to know about the dataset before analyzing it statistically. How might you visualize this information?

The exploratory analysis of the data involved initial visualizations of the power outage frequency for each state. For each state, a chart was created of power outages from 2015 to 2023 and another chart of power outages by month group by year. This provided a preview of the overall trend and seasonal trend occurring in the power outage data, but further analysis is needed to quantify these trends.

3.0.1 California

Initial data exploration of California power outage data suggest a slight increasing trend (Figure 1). Furthermore, Figure 2 shows.

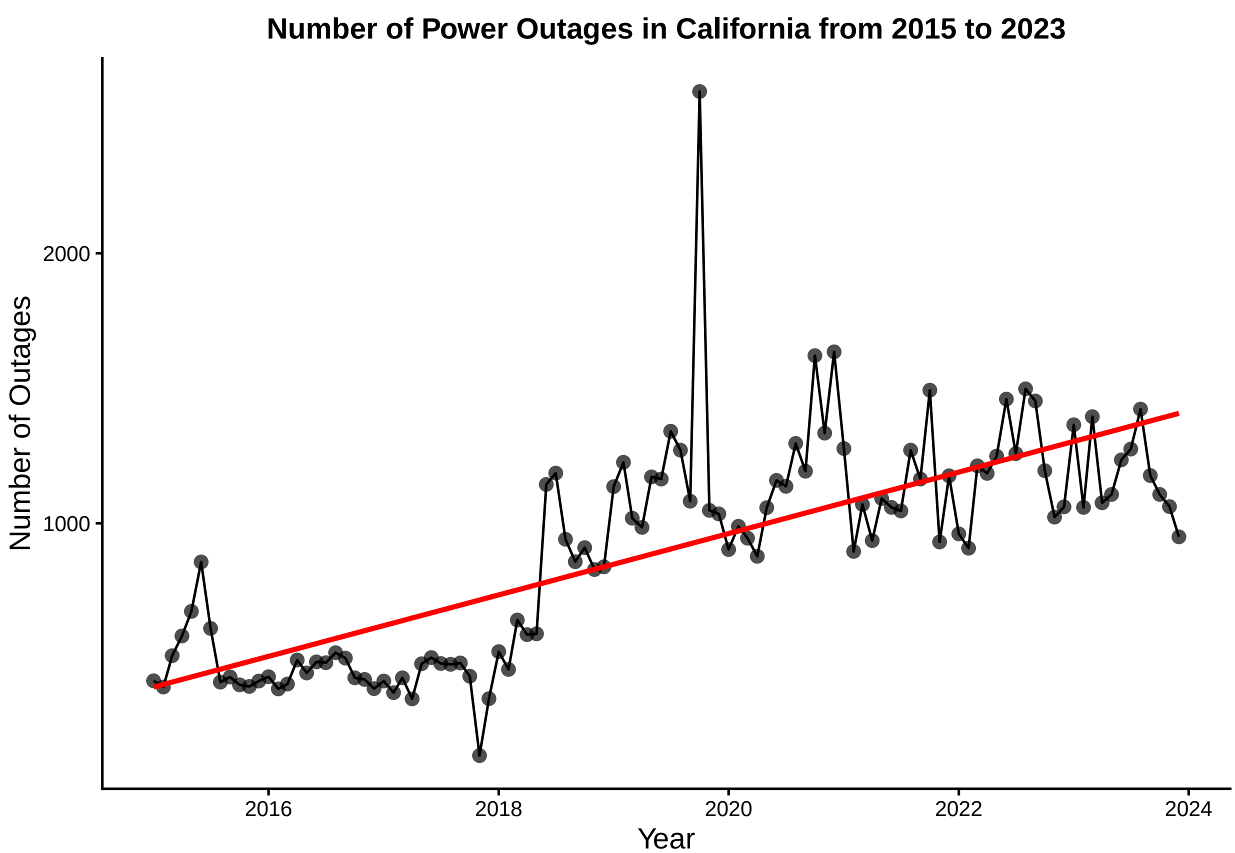

Figure 1: Yearly plot of power outages in California from 2015 to 2023.

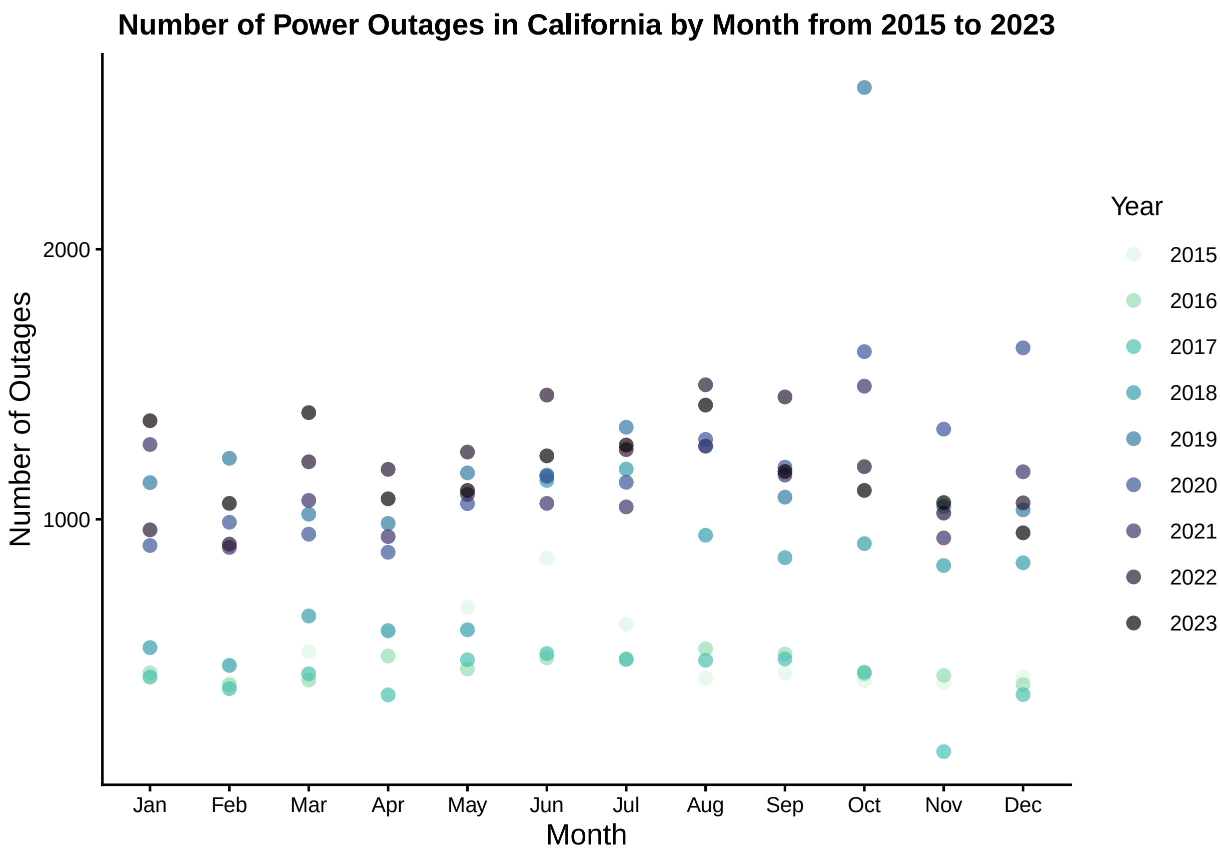

Figure 2: Monthly plot of power outages in California from 2015 to 2023.

3.0.2 Florida

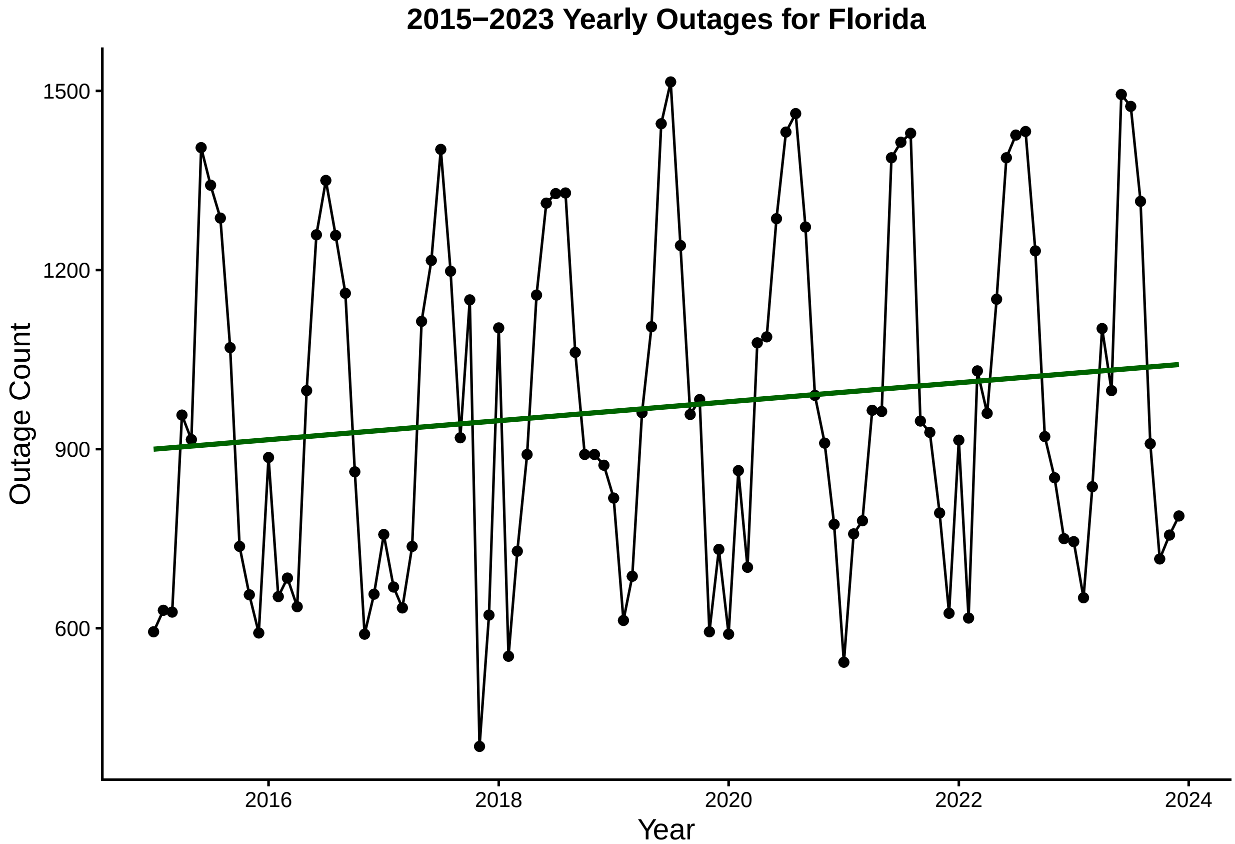

Figure 3: Yearly plot of power outages in Florida from 2015 to 2023.

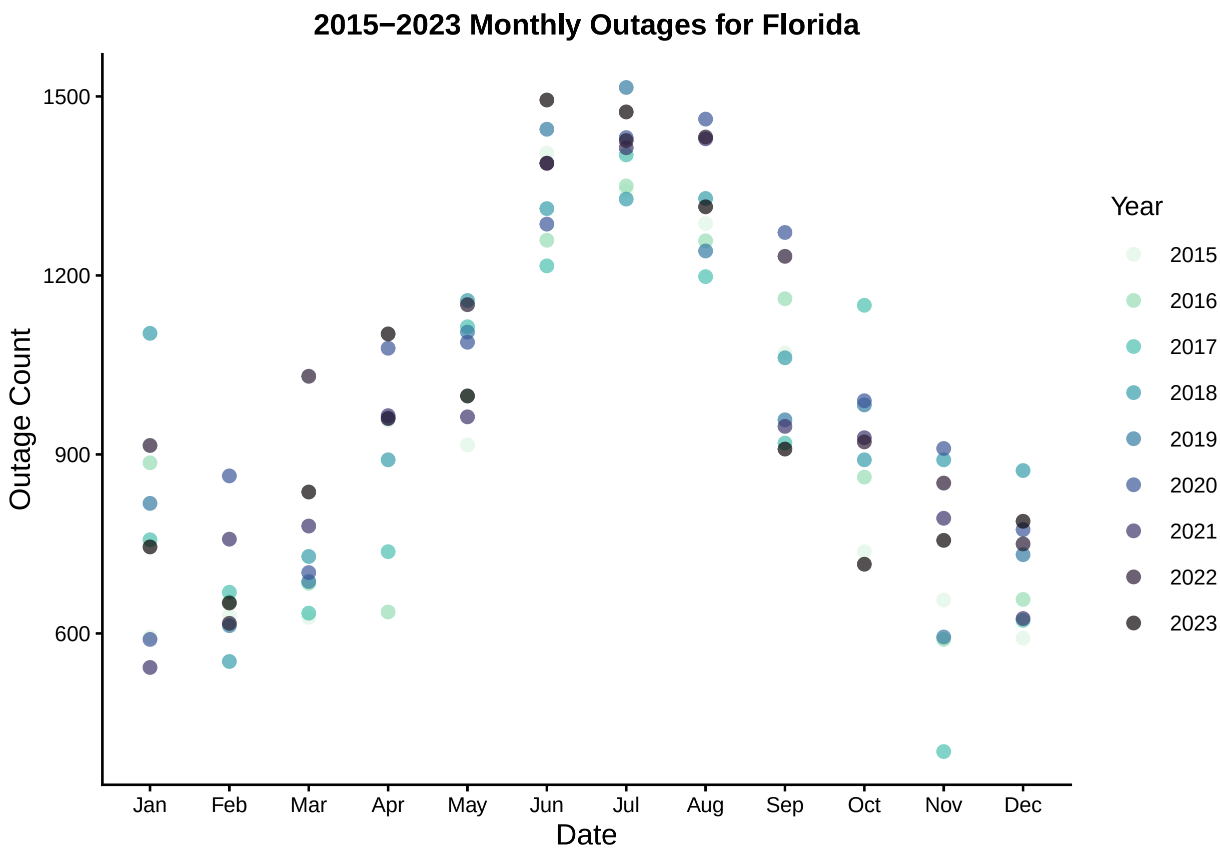

Figure 4: Monthly plot of power outages in Florida from 2015 to 2023.

3.0.3 Pennsylvania

From the initial data exploration of Pennsylvania, it is clear that power outages have gradually increased over the past decade. On a monthly basis, there is a noticeable rise in outages during the summer months, suggesting a seasonal pattern likely related to weather or energy demand. Overall, the data points to both long-term growth in outage frequency and predictable seasonal fluctuations.

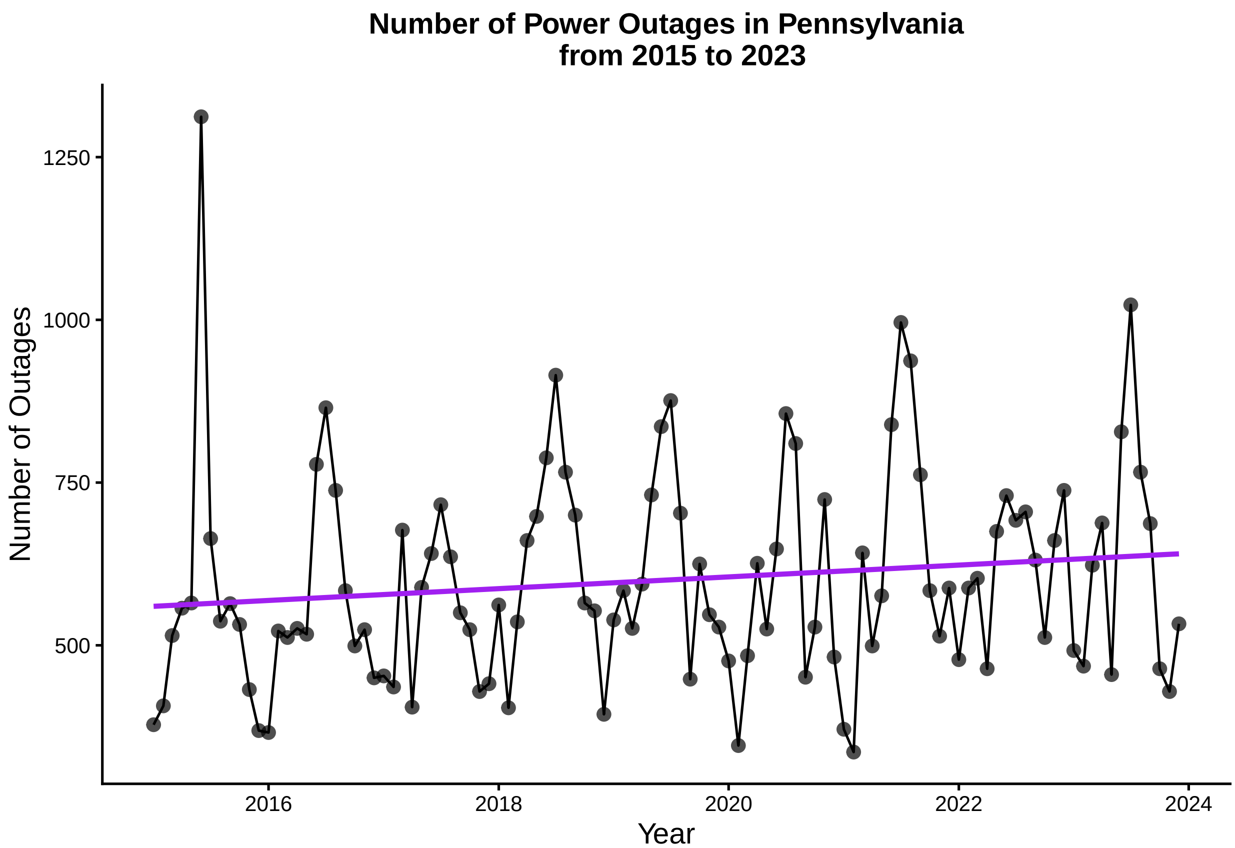

Figure 5: Yearly plot of power outages in Pennsylvania from 2015 to 2023.

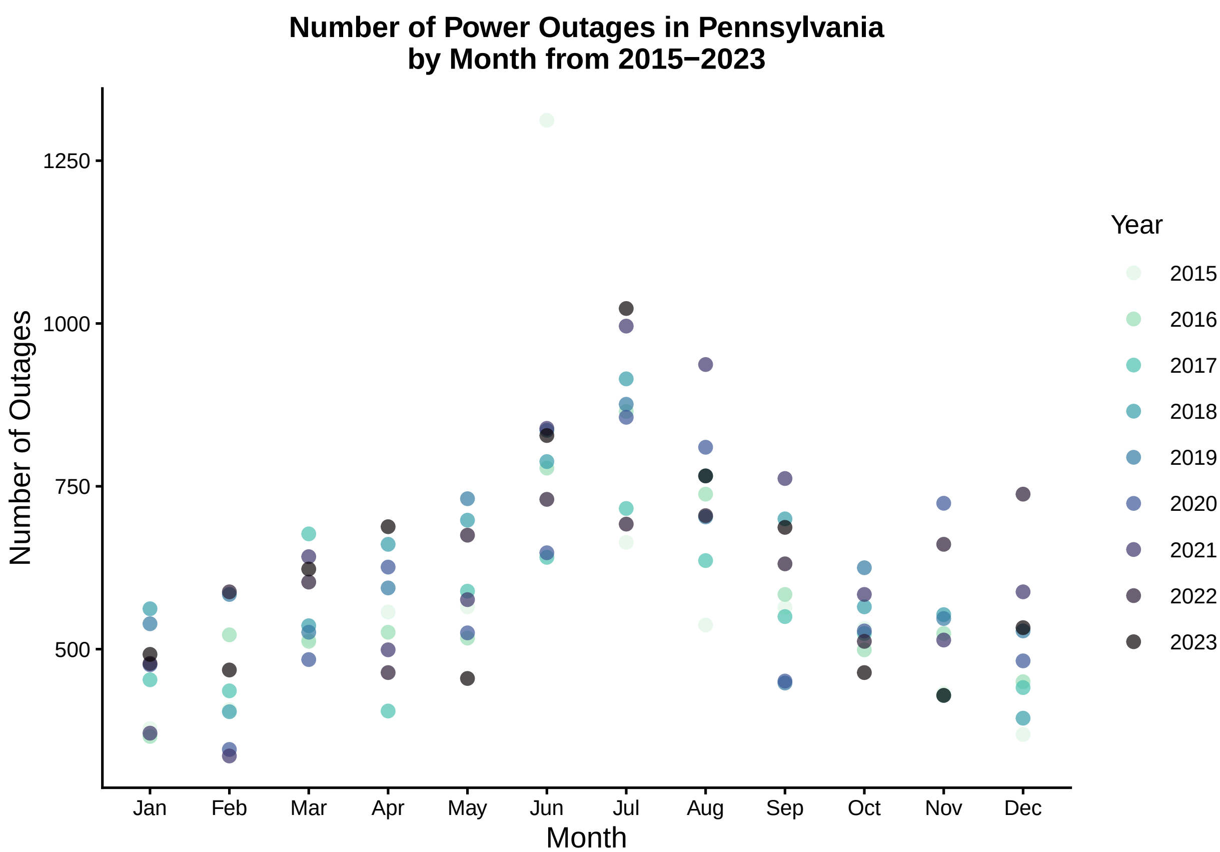

Figure 6: Monthly plot of power outages in Pennsylvania from 2015-2023.

3.0.4 Texas

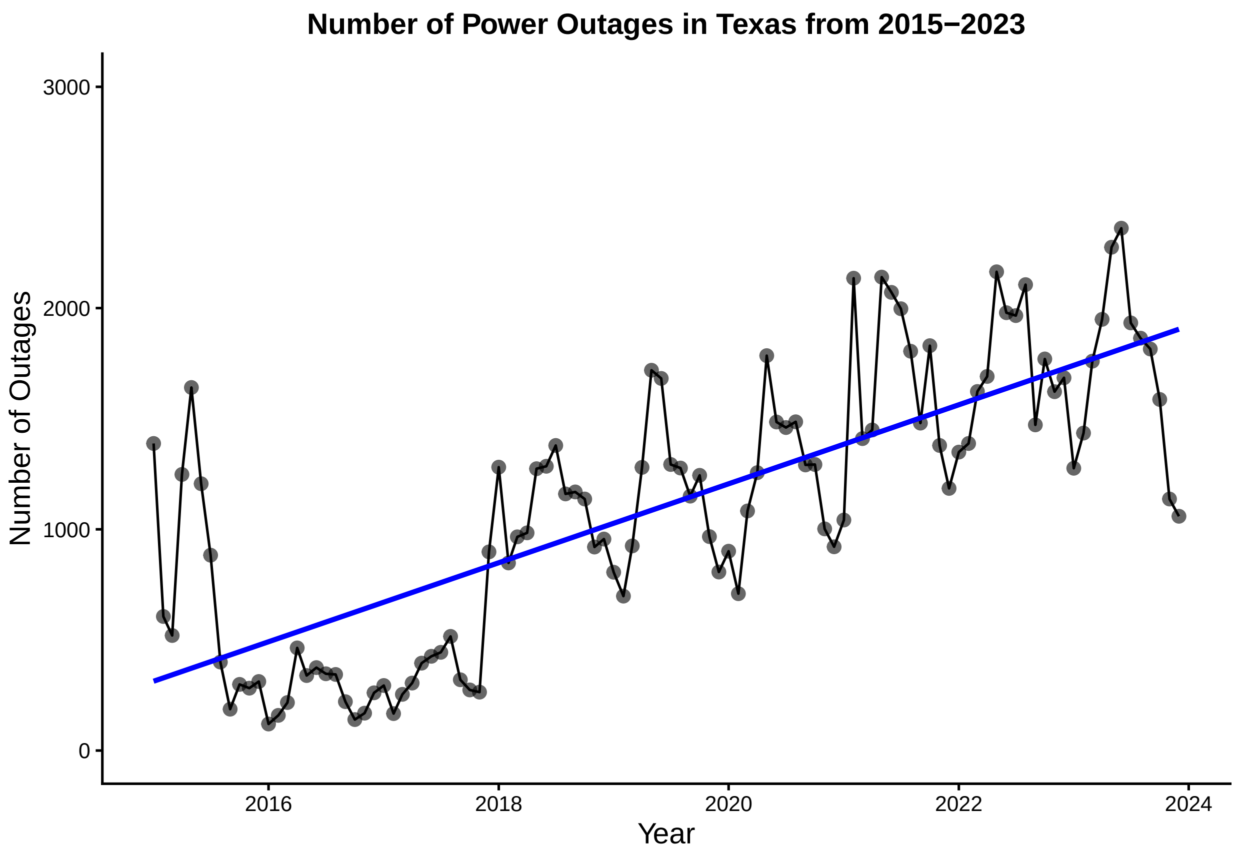

Figure 7: Yearly plot of power outages in Texas from 2015-2023.

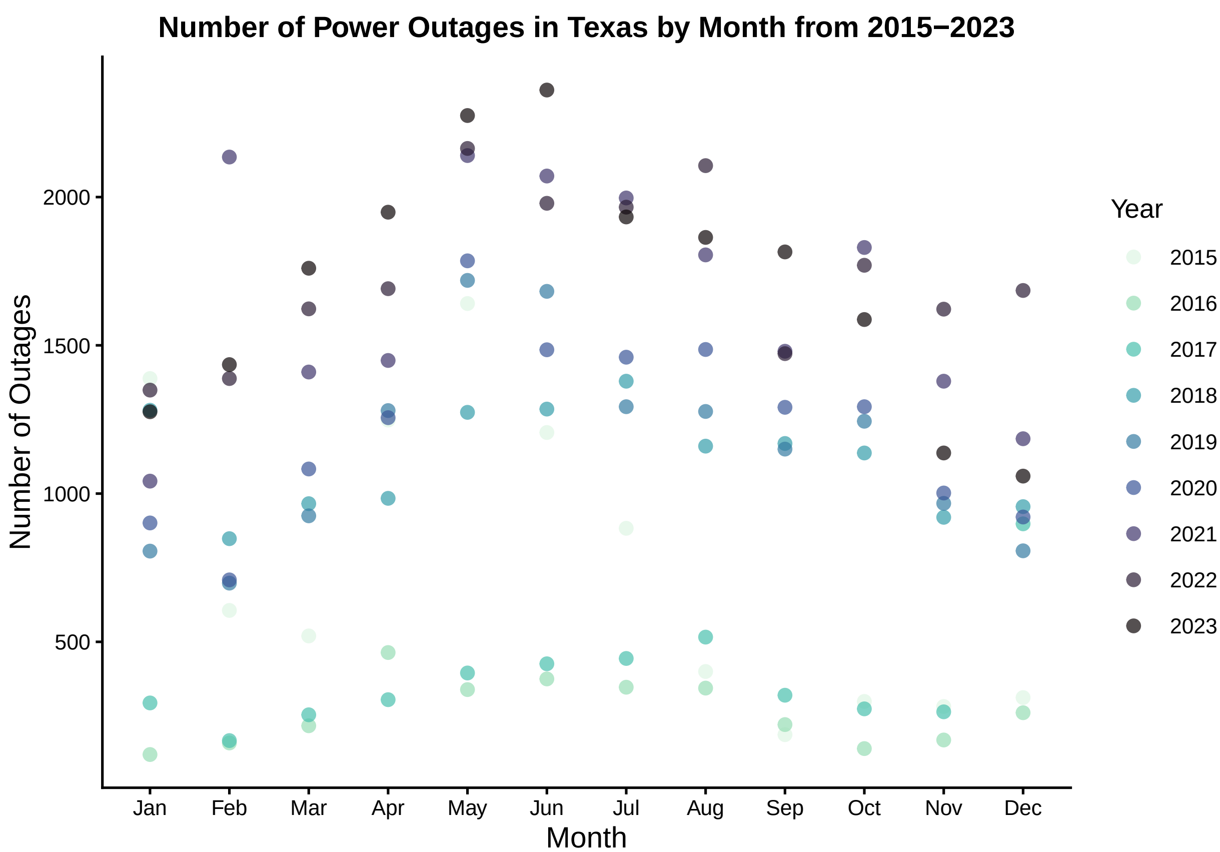

Figure 8: Monthly plot of power outages in Texas from 2015 to 2023.

4 Analysis

Insert visualizations and text describing your main analyses. Format your R chunks so that graphs are displayed but code and other output is not displayed. Instead, describe the results of any statistical tests in the main text (e.g., "Variable x was significantly different among y groups (ANOVA; df = 300, F = 5.55, p < 0.0001)"). Each paragraph, accompanied by one or more visualizations, should describe the major findings and how they relate to the question and hypotheses. Divide this section into subsections, one for each research question.

Each figure should be accompanied by a caption, and each figure should be referenced within the text

The power outage datasets for each state were analyzed by creating time series objects of the power outage data. The time series objects were decomposed to separate the trend, seasonality, and remainder of each dataset. New data frames were created with these three time series components. A non-seasonal power outage frequency component was created by subtracting the seasonal component from the trend and remainder components. This nonseasonal components was graphed along with the observed power outage frequency to display the visual fit of the trend. Mann-Kendall tests were performed on the nonseasonal component to statistically analyze the trends, providing the trend direction and statistical significance.

The seasonal component was isolated and graphed by grouping the yearly values (2015-2023) by month. This shows how the seasonal component changes over the year, allowing for the months with a high frequency of power outages to be identified. To statistically analyze the seasonality of power outages, an ANOVA test was performed to see if power outage counts varied significantly across months.

To investigate the severity of power outages, time series objects were created of the customer weighted hours data. The customer weighted hours' time series objects were also decomposed to separate the trend, seasonality, and remainder of each dataset. A non-seasonal power outage frequency component was created by subtracting the seasonal component from the trend and remainder components. This nonseasonal components was graphed along with the observed customer weighted hours to display the visual fit of the trend. Mann-Kenndall tests were performed on the nonseasonal component to statistically analyze the trends, providing the trend direction and statistical significance. [UPDATE BASED ON WHAT WE INCLUDE]

4.1 Question 1: How has the frequency of power outages changed over time?

4.1.1 California

Figure 9 displays the decomposed time series of the California power outage frequency from 2015 to 2023. There appears to be clear seasonal trend, which is analyzed further in part (5).

A MannKendall non seasonal trend analysis was applied to the California power outage frequency dataset with the seasonal component remove, a significant overall increase in power outages over time is observed (MannKendall; tau = 5.437e-01 p < 2.2e-16). Figure 10 visualizes the trend compared to the observed frequency of power outages in California. The red line represents the overall increasing trend.

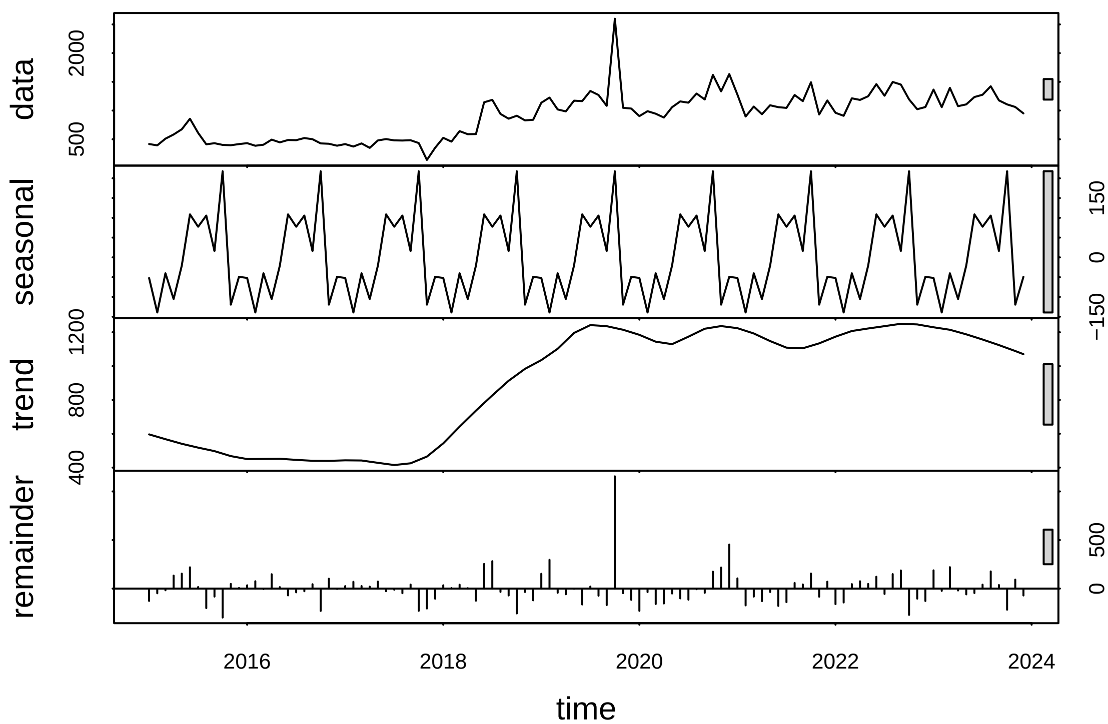

Figure 9: Decomposed components of the California power outage count time series.

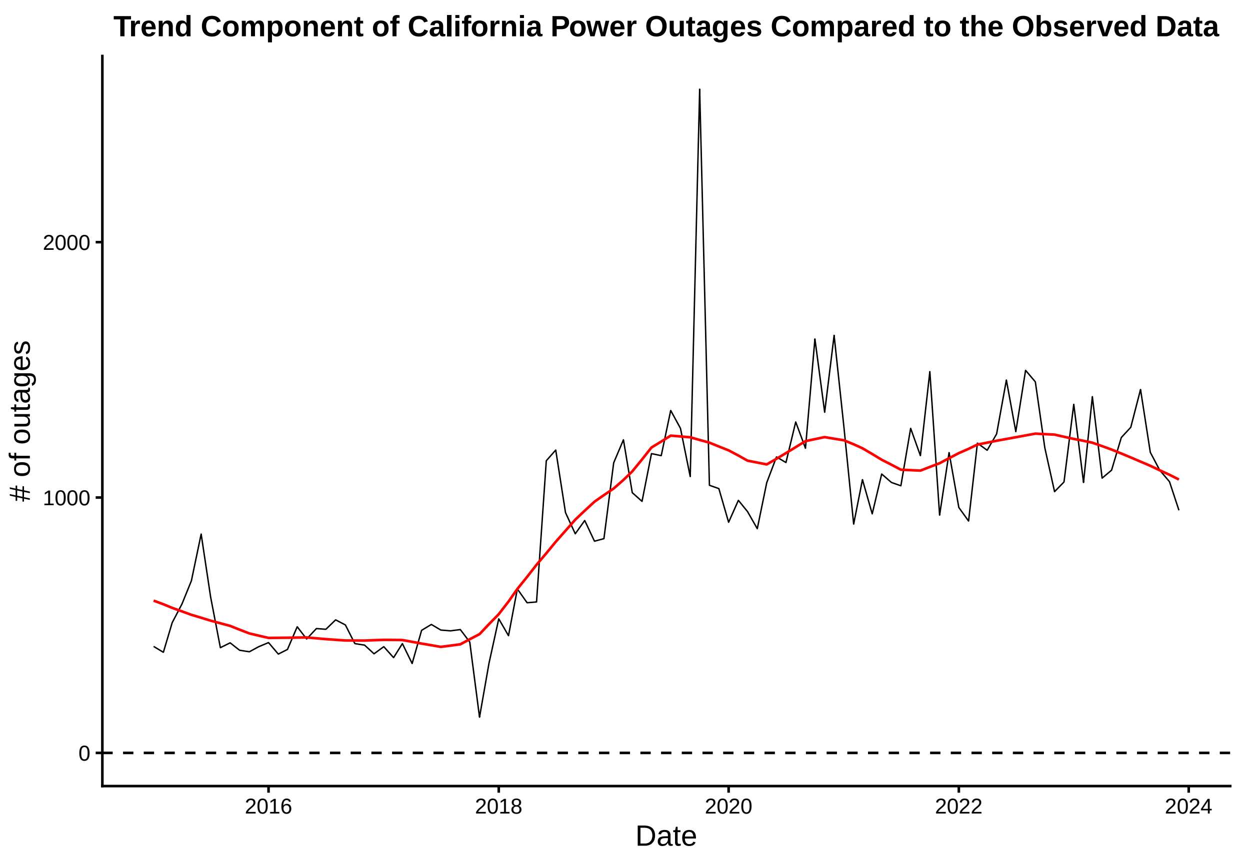

Figure 10: Trend versus observation of power outages in California (2015-2023).

4.1.2 Florida

Figure 11 shows the time series analysis for outage occurrences across Florida from 2015 to 2023. While a clear pattern of seasonality is evident within the time series, applying a non-seasonal Mann-Kendall test will determine if outage count across the years has a significant trend. The results of the test confirm that there is a statistically significant increase in outages in Florida (tau = 2.426865e-01, p = 0.0001997), independent from seasonality changes. Figure 12 shows the increasing trend.

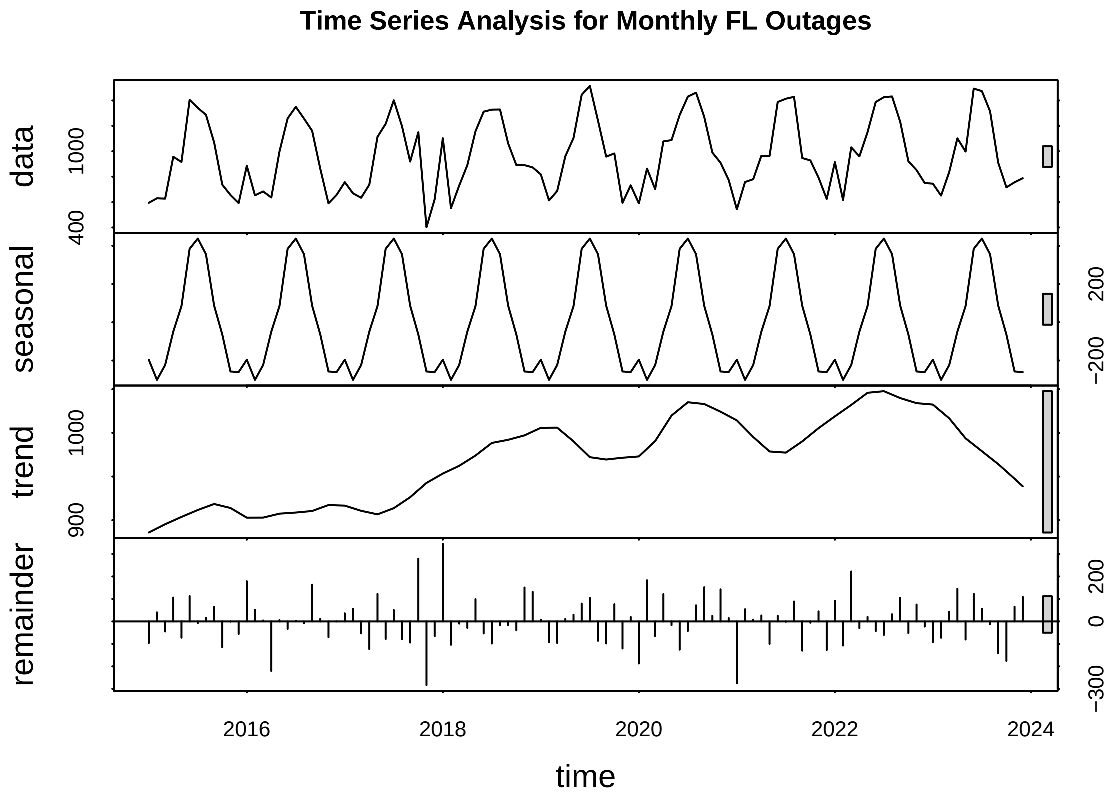

Figure 11: Time series for FL monthly power outages from 2015 to 2023.

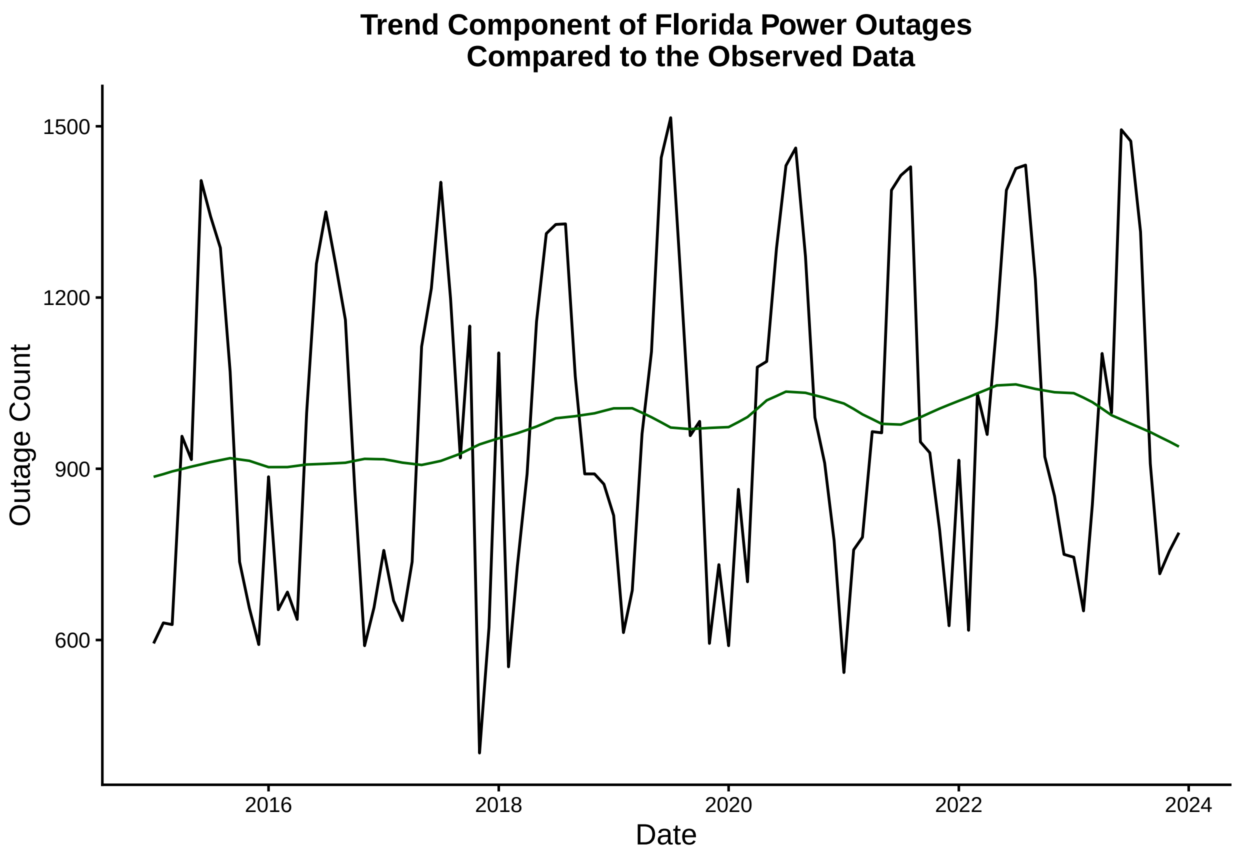

Figure 12: FL Trend Analysis for Monthly Outages.

4.1.3 Pennsylvania

After removing the seasonal component from the Pennsylvania outage time series, the Mann–Kendall trend test revealed a significant positive trend (z = 3.02, n = 108, p = 0.002), indicating that outage frequency has increased from 2015 to 2023. This result supports the hypothesis that outages have become more common during this time, independent of seasonal patterns. The decomposed time-series visualization shows this trend clearly, revealing that even after accounting for monthly fluctuations, the long-term component continues to rise.

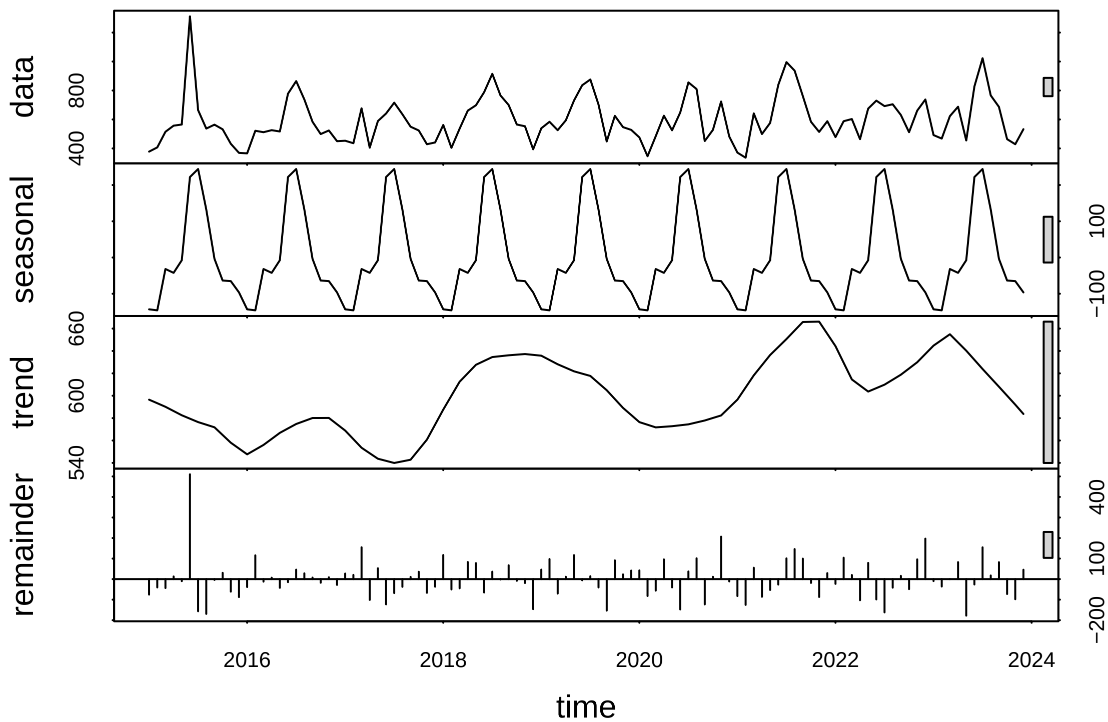

Figure 13: Decomposed components of the Pennsylvania power outage count time series.

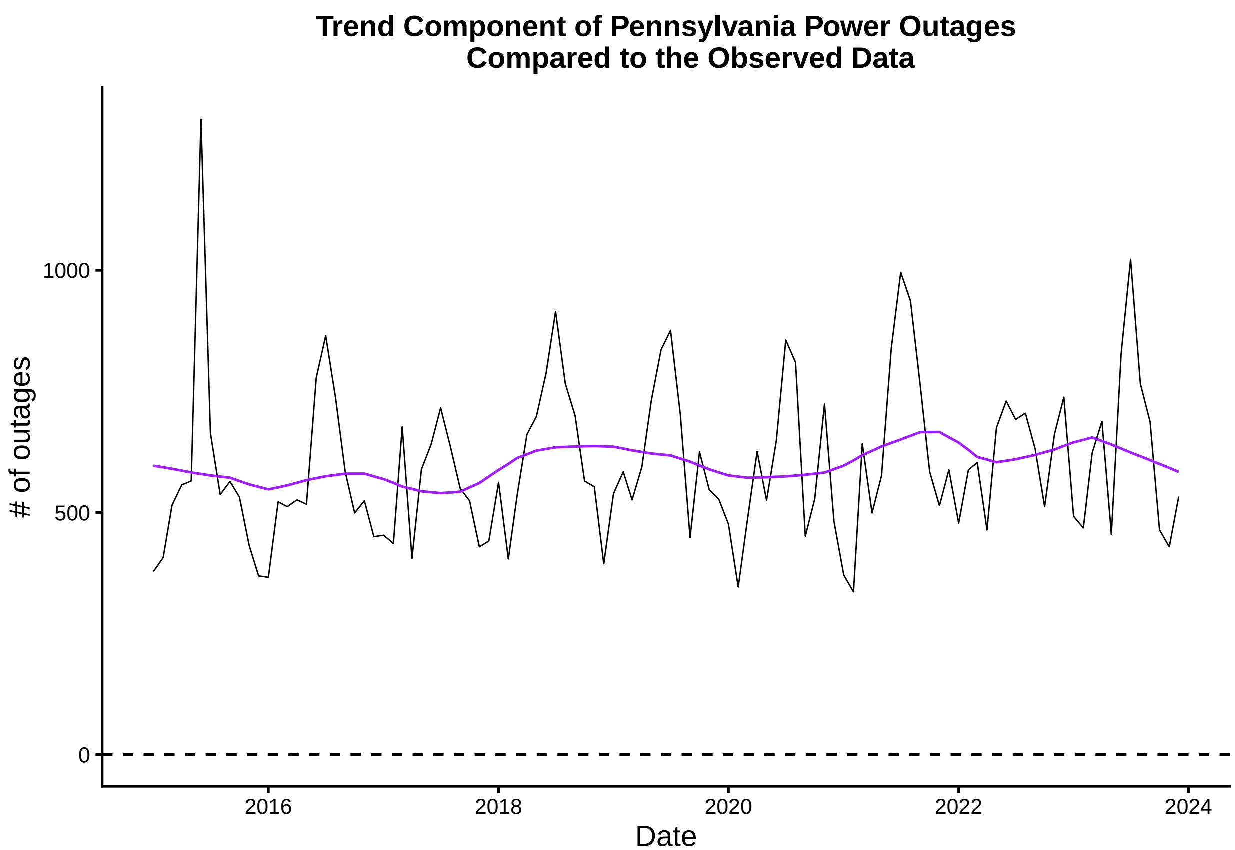

Figure 14: Trend versus observation of power outages in Pennsylvania (2015-2023).

4.1.4 Texas

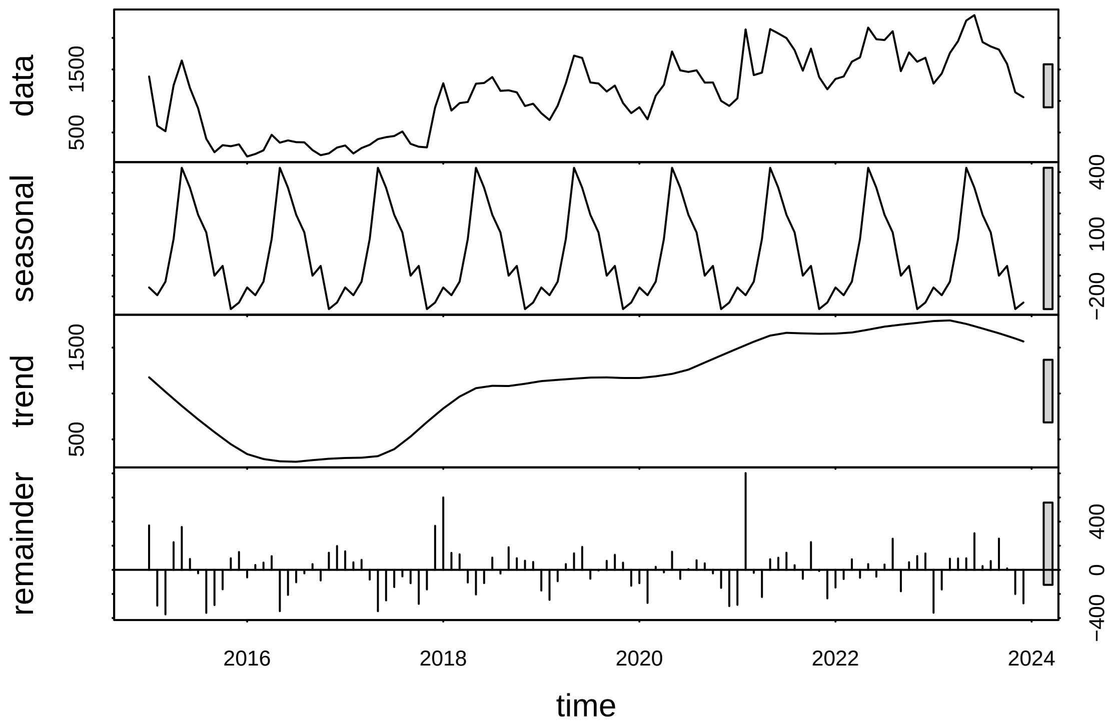

Figure 15: Decomposed components of the Texas power outage count time series.

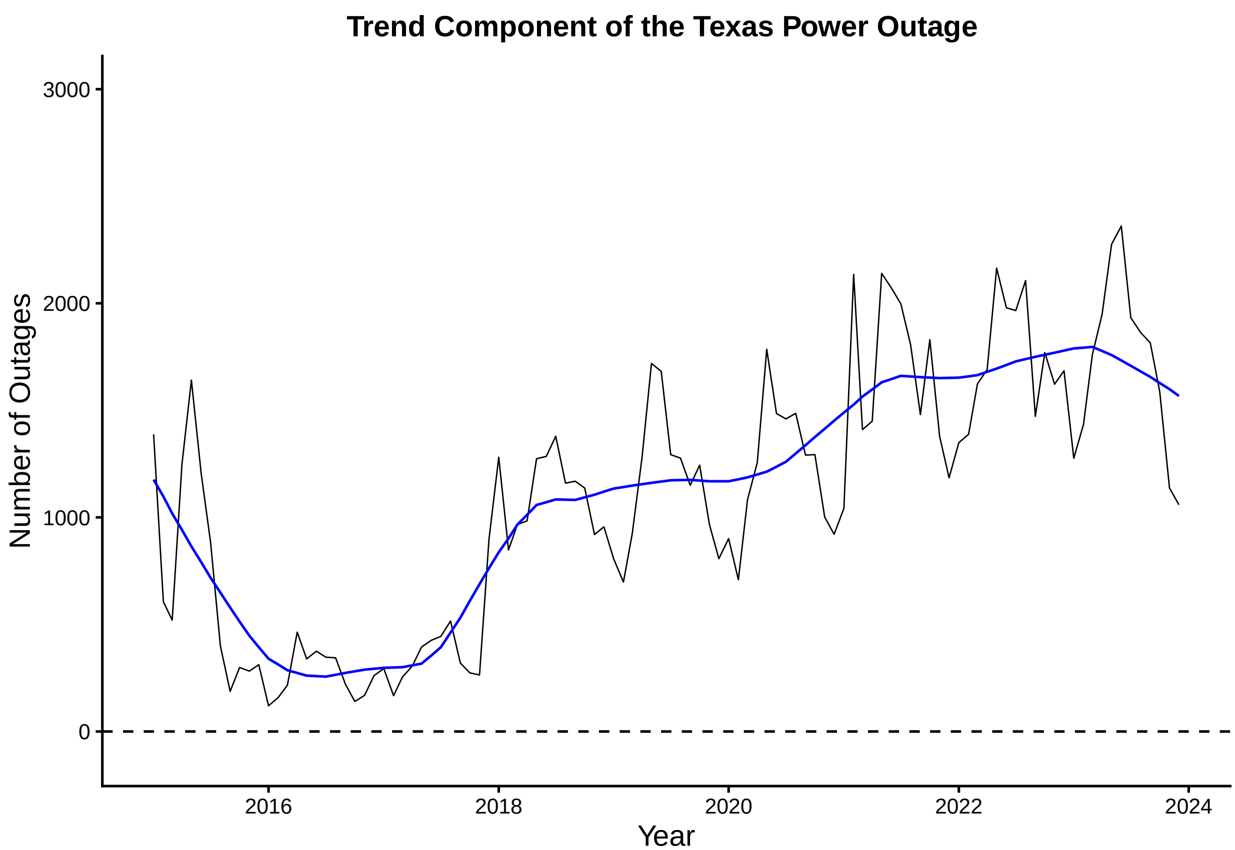

Figure 16: Trend versus observation of power outages in Texas (2015-2023).

4.1.5 Comparision between States

4.2 Question 2: Is there a seasonal trend? Are certain months more prone to outages?

4.2.1 California

Figure 9 displays a clear seasonal trend in the decomposed time series of the California power outage frequency. The seasonal trend was isolated and grouped by year to show the Seasonal Cycle of Power Outages in California in Figure 17. The seasonal cycle highlights a potential pattern of power outage frequencies peaking rising in the summer months, peaking in October. However, the frequency of outages in California was not significantly different among the months (ANOVA; df = 11, F = 0.704, p < 0.732).

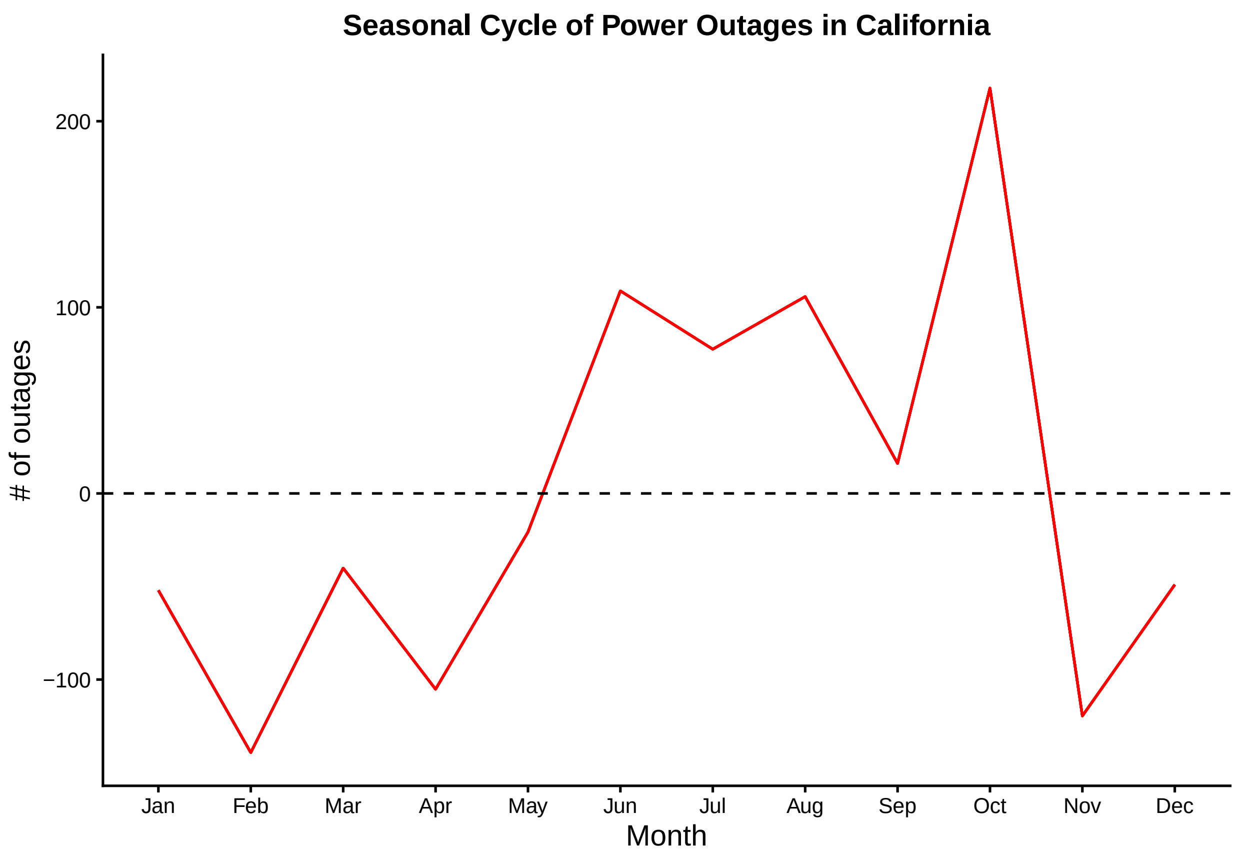

Figure 17: Seasonality component of the California power outage count time series analysis where yearly data is grouped by month.

4.2.2 Florida

Results from the time series for outages in FL displayed a pattern of seasonality. To further analysis the strength of that trend, months of every year were aggregated to show a cumulative trend across 2015 to 2023 in Figure 18. A clear normal distribution shows a concentration of outage occurrences in summer months, and a significant decline in winter months. This is statistically proven with an ANOVA test that confirms outage counts between months vary significantly, following a seasonal trend (Df = 11, F-value = 43.28, $p < 2e\text{-}16$).

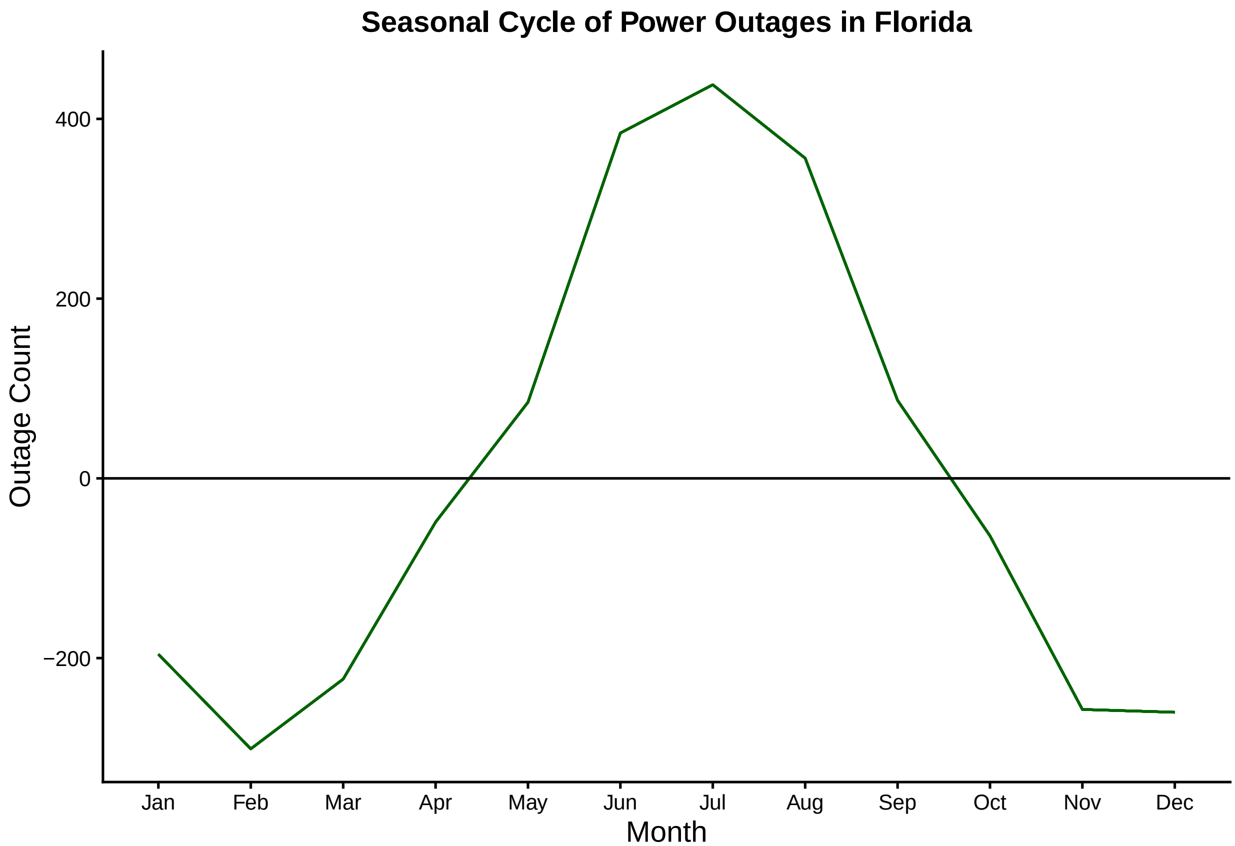

Figure 18: Seasonality component of FL outage TS analysis, years grouped by month

4.2.3 Pennsylvania

The Mann–Kendall seasonal trend test for Pennsylvania outage counts indicated a significant seasonal signal in the data ($\tau = 0.20$, p = 0.009), demonstrating that outage frequencies vary across months. This finding supports the hypothesis that outages are not evenly distributed throughout the year but instead follow a seasonal pattern. The visualizations further highlight this pattern, showing clear peaks in June, July, and August, when outage counts are highest. Together, these results suggest that summer months consistently experience elevated outage activity, potentially driven by seasonal weather conditions or increased energy demand.

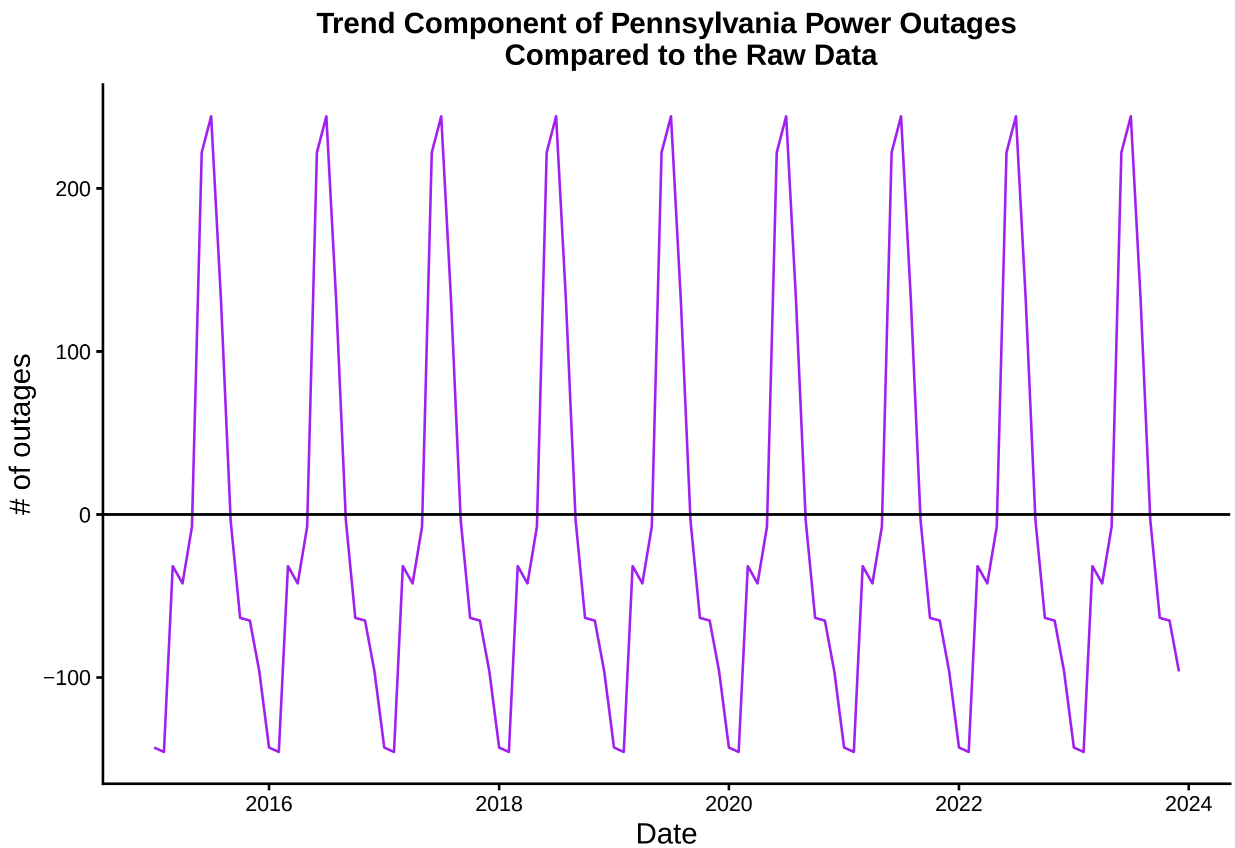

Figure 19: Isolated seasonal component of PA power outages.

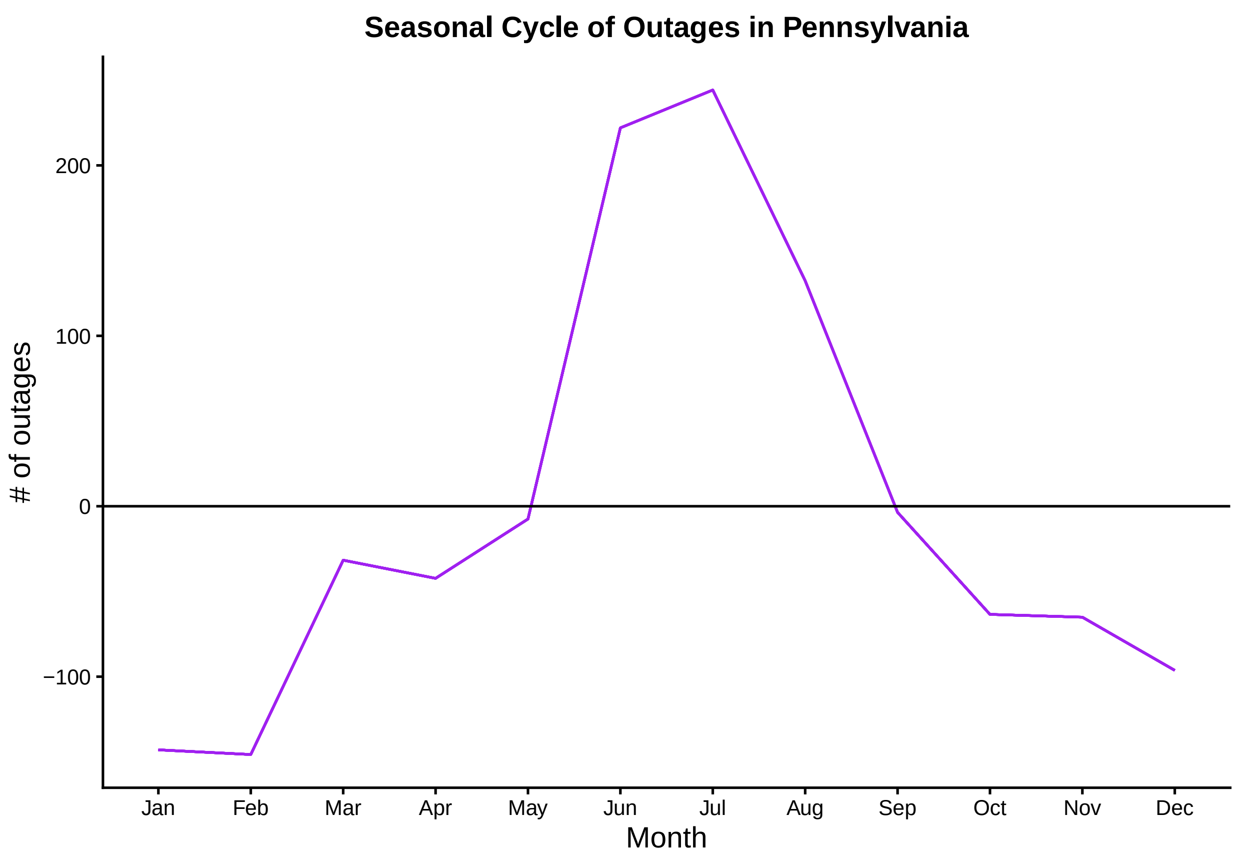

Figure 20: Single year breakdown of seasonal component of outages in PA.

4.2.4 Texas

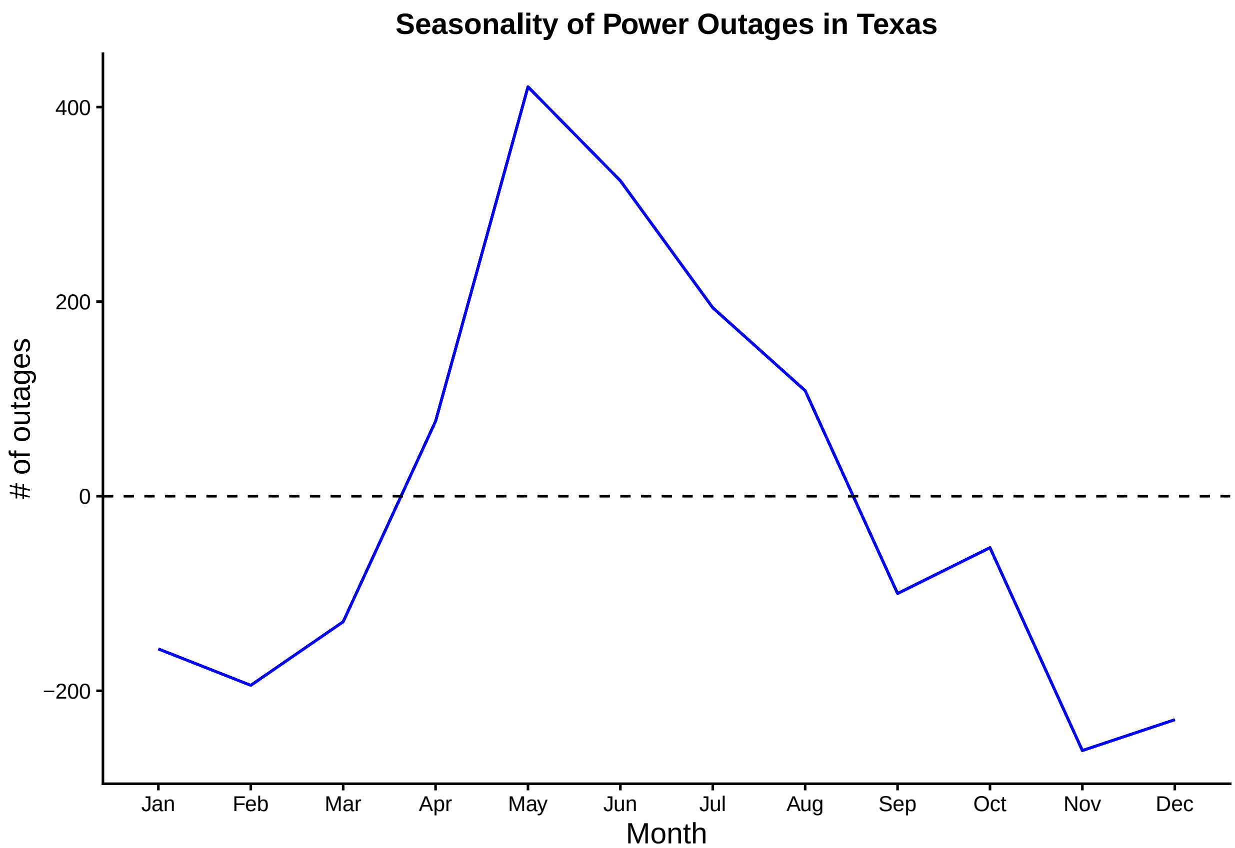

Figure 21: Seasonality component of the Texas power outage count time series analysis where yearly data is grouped by month.

4.2.5 Comparision between States

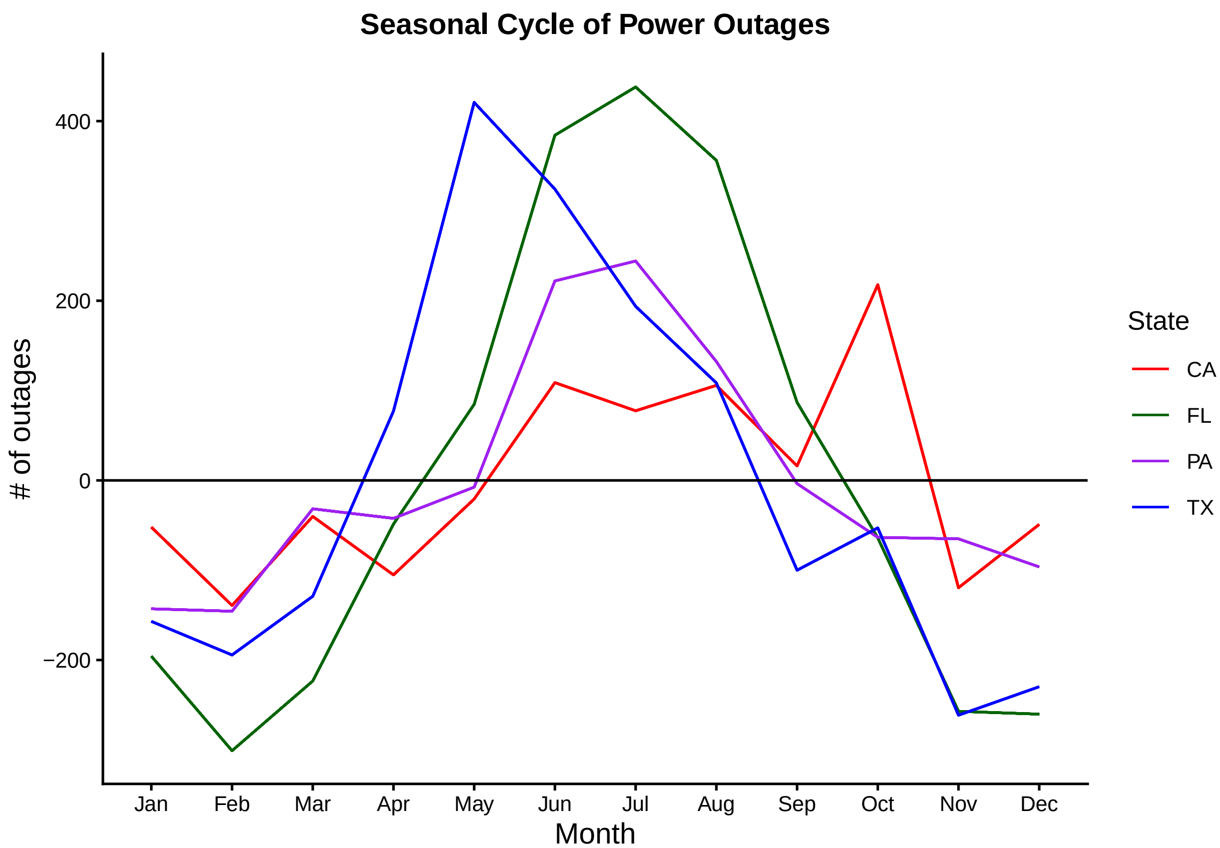

Figure 22: Seasonal cycle of power outages (2015-2023 combined).

4.3 Question 3: How has the severity of power outages changed over time?

4.3.1 California

Figure 23 displays the decomposed time series of the California customer weighted hours from 2015 to 2023. However, after removing the seasonality component the Mann-Kendall trend test suggested that there is a significant increase in the customer weighted hours of power outages in California, thus the outage severity (MannKendall; tau 3.513e-01 = $p < 7.178\text{e-}08$).

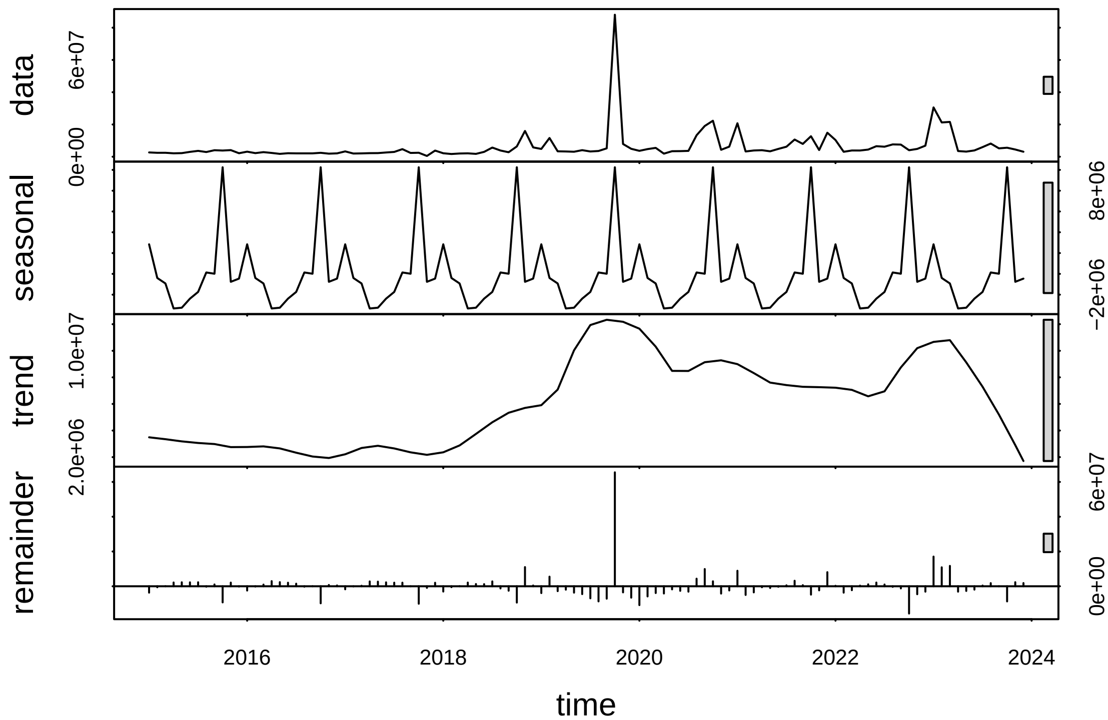

Figure 23: Decomposed analyis of customer weighted outage time in California

Figure 24 visualizes the time series analysis for customer weighted hours in Florida from 2015 to 2023. The results of the Mann-Kendall test on seasonality across the data returned inconsistent tau results suggesting seasonality was not a significant trend. Running a non-seasonal Mann-Kendall test showed far greater significance in an increasing customer weighted duration ($z = 1.2186$, $n = 108$, p-value $= 0.223$). Unlike the other states, the p-value for non-seasonal trend is greater than 0.05, suggesting that the customer weighted hours follow more of a seasonal trend than may be present in other states.

4.3.2 Florida

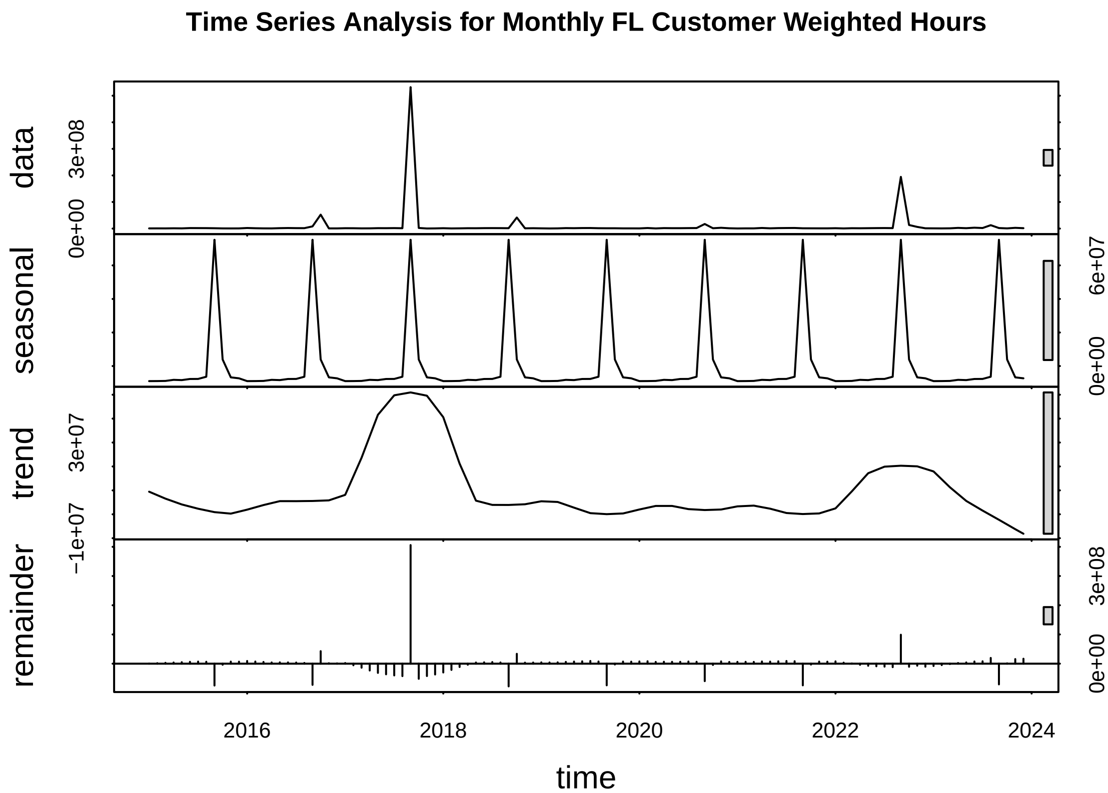

Figure 24: Time series analysis for FL customer weighted hours from 2015 to 2023.

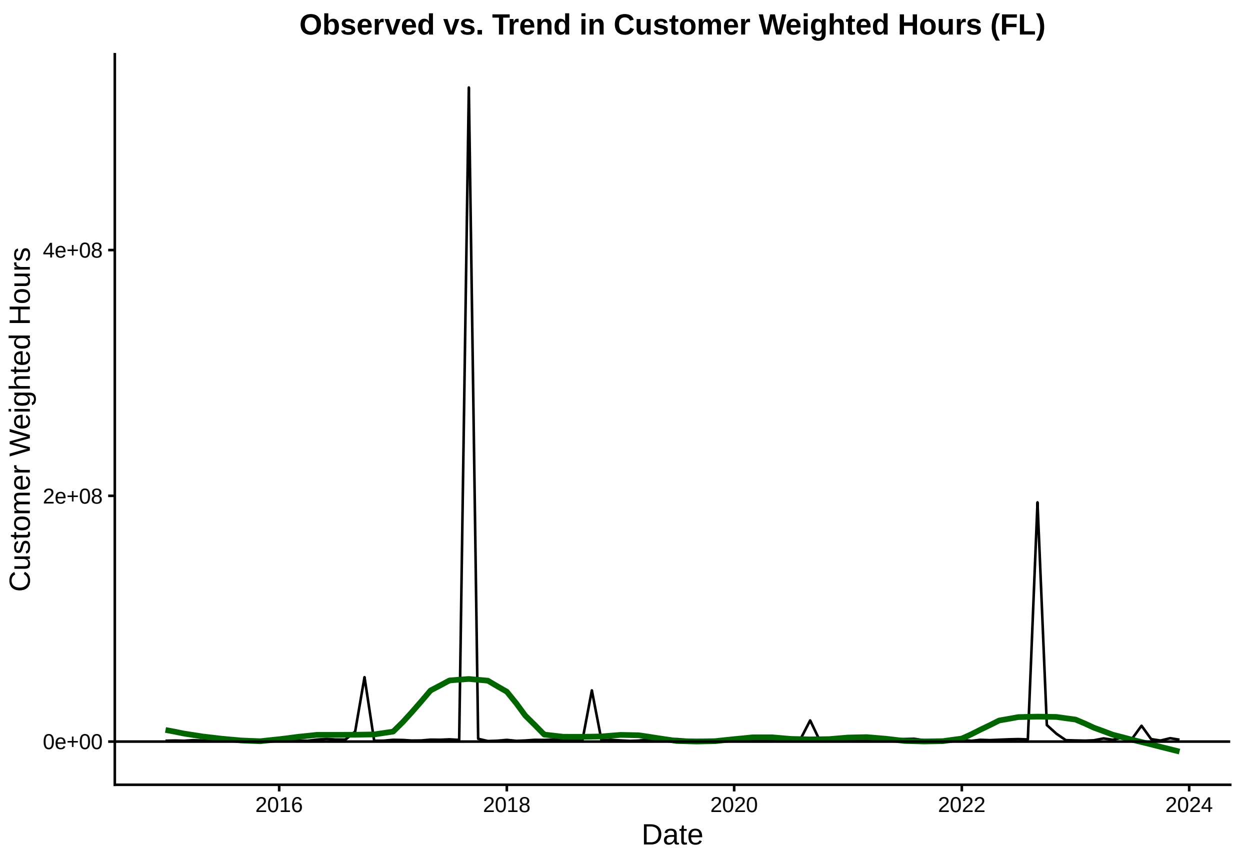

Figure 25: FL Trend Analysis for Customer Weighted Hours.

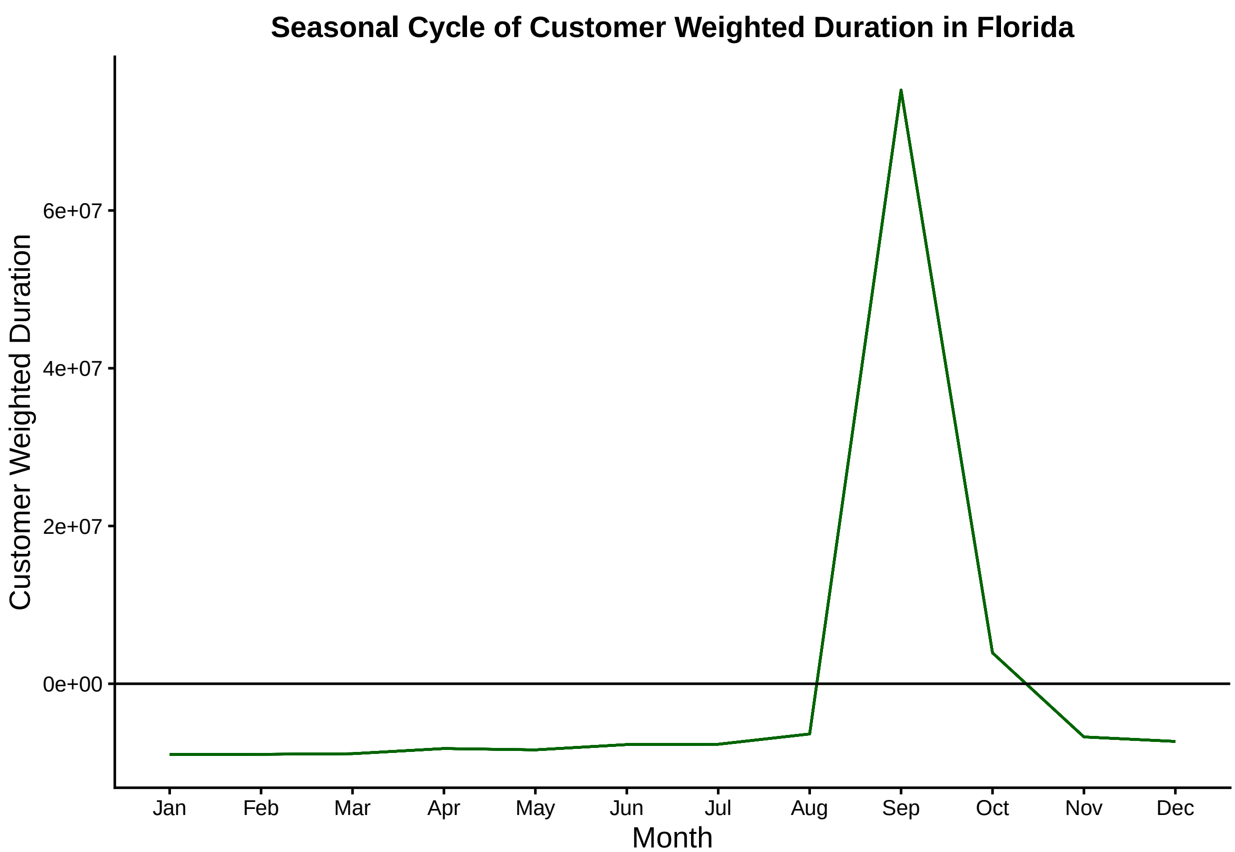

Figure 26: FL Seasonal Analysis for Customer Weighted Hours, years grouped by months.

4.3.3 Pennsylvania

For customer-weighted hours, the seasonal test was not statistically significant (tau = 0.13, p = 0.09), indicating no consistent monthly pattern in outage impacts on customers. However, after de-seasonalizing, the Mann-Kendall trend test suggested a marginally increasing trend in customer-weighted hours over time (z = 1.89, n = 108, p = 0.06). Overall, these results suggest that while outage counts are increasing and show clear seasonality, the impact on customers is also trending upward, though less strongly.

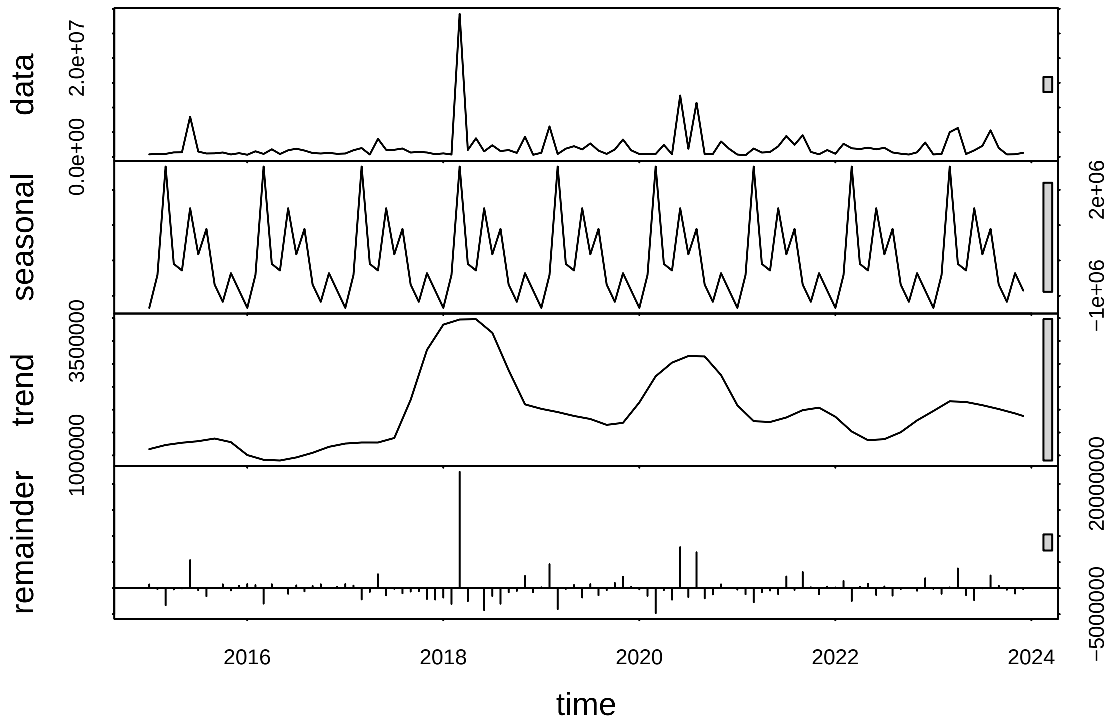

Figure 27: Decomposed analyis of customer weighted outage time in Pennsylvania.

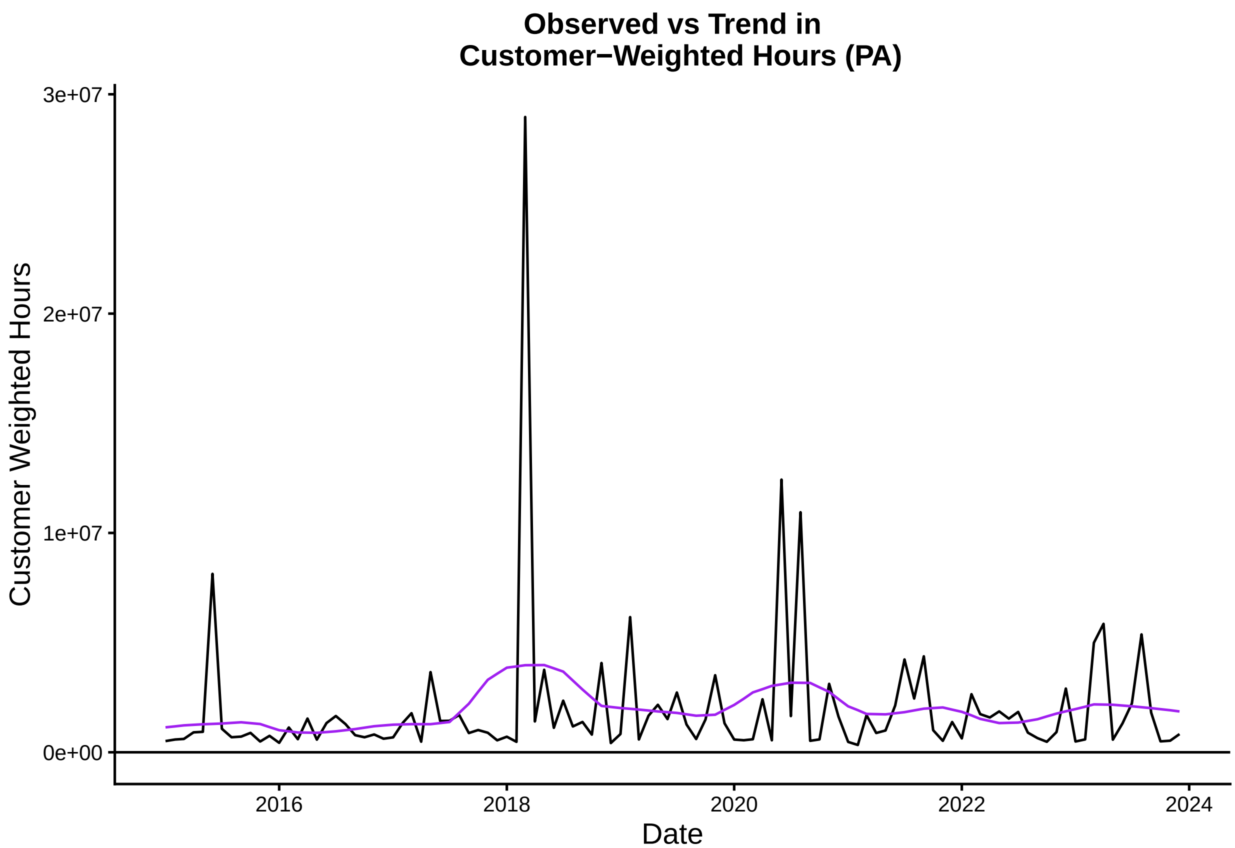

Figure 28: Trend versus observation of customer-weighted hours in PA (2015-2023).

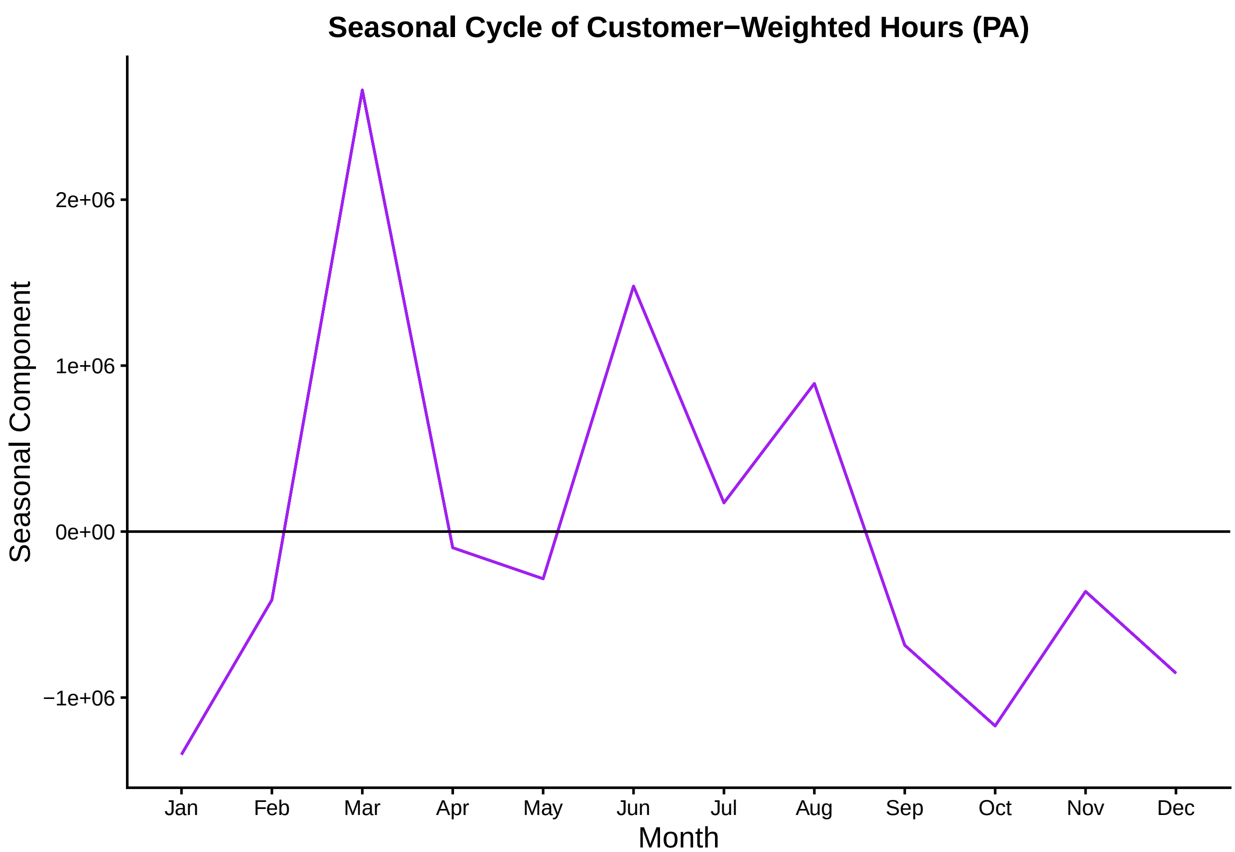

Figure 29: Single year breakdown of seasonal component of customer-weighted hours in PA.

4.3.4 Texas

4.3.5 Comparison between States

5 Summary and Conclusions

Summarize your major findings from your analyses in a few paragraphs. What conclusions do you draw from your findings? Relate your findings back to the original research questions and rationale.

After accounting for seasonal patterns, the analysis shows that the frequency of power outages in Pennsylvania has increased from 2015 to 2023. Outage counts display a clear seasonal trend, with the highest numbers occurring in the summer months of June, July, and August. In contrast, customer-weighted hours do not show a strong seasonal pattern, but there is a slight upward trend over time, indicating that the impact of outages on customers is also gradually increasing. Overall, while outages are becoming more frequent and seasonally concentrated, the effect on customers is rising more moderately.

6 References

- IBM. Power grid modernization—Strategies and tactics for resilience and energy transition. (2025, March 21). IBM. https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/power-grid-modernization
- Wendel, J. (2024, July 29). Extreme Heat, Hurricanes, Wildfires: How Summer's Extremes Disrupt the Power Grid. Pacific Northwest National Laboratory. https://www.pnnl.gov/news-media/extreme-heat-hurricanes-wildfires-how-summers-extremes-disrupt-power-grid
- Wing, I. S., Larsen, P. H., Carvallo, J. P., Sanstad, A., Wei, D., Rose, A., Baik, S., Smith, J., Ramee, C., & Peterson, R. (2025). A Method to estimate the economy-wide consequences of widespread, long duration electric power interruptions. Nature Communications, 16(1), 3335. https://doi.org/10.1038/s41467-025-58537-4