Summary Report

Assignment 1

This Python function, get_hours, calculates the number of hours that fall into specific peak type blocks for different ISOs (Independent System Operators) within a given period. The function takes three parameters: iso (one of PJM, MISO, ERCOT, SPP, NYISO, WECC, or CAISO), peak_type (one of onpeak, offpeak, flat, 2x16H, or 7x8), and period (which can be daily, monthly, quarterly, or annually).

Detailed Breakdown of Functionality:

- **Peak Hours Definition:** Clearly defines which hours are considered peak, off-peak, etc., for each ISO.

- **Period Parsing:** Accurately parses and handles different period formats, converting them into start and end dates.

- **NERC Holidays:** Incorporates major holidays into the calculation to ensure accuracy for onpeak and offpeak hours.

- **Date Range Handling:** Generates an hourly date range to evaluate each hour within the specified period.

- **Daylight-Saving Time:** Adjusts for daylight-saving time where applicable, ensuring accurate hour counts for each ISO.

- **Peak Hours Calculation:** Iterates through each hour, applying specific rules for each peak type and counting the total hours that meet the criteria.

- **Return Structure:** Provides a comprehensive summary of the results, including the total number of peak type hours within the specified period.

Based on the sample, I get the correct answer:

```
[ ] # Sample Run
    results = get_hours("ERCOT", "onpeak", "2019May")
    print(results)

{'iso': 'ERCOT', 'peak_type': 'ONPEAK', 'startdate': '2019-05-01', 'enddate': '2019-05-31', 'num_hours': 352}
```

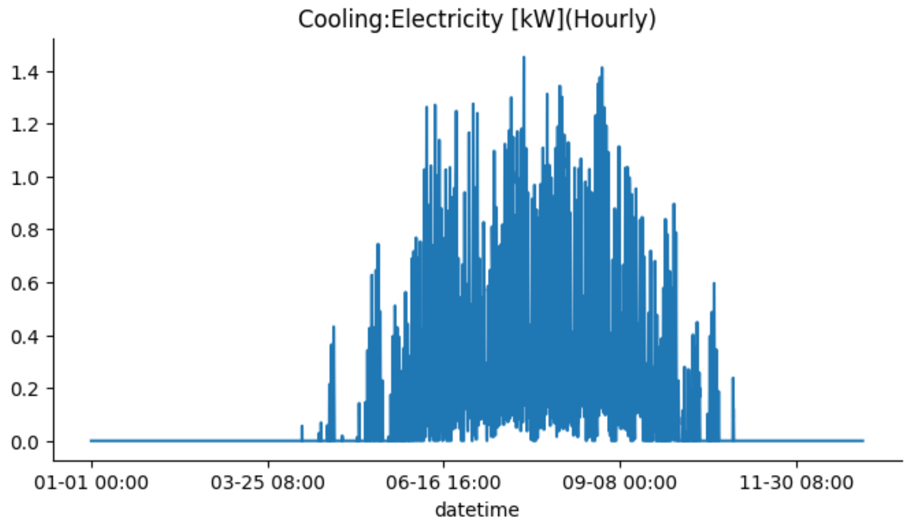
Assignment 2

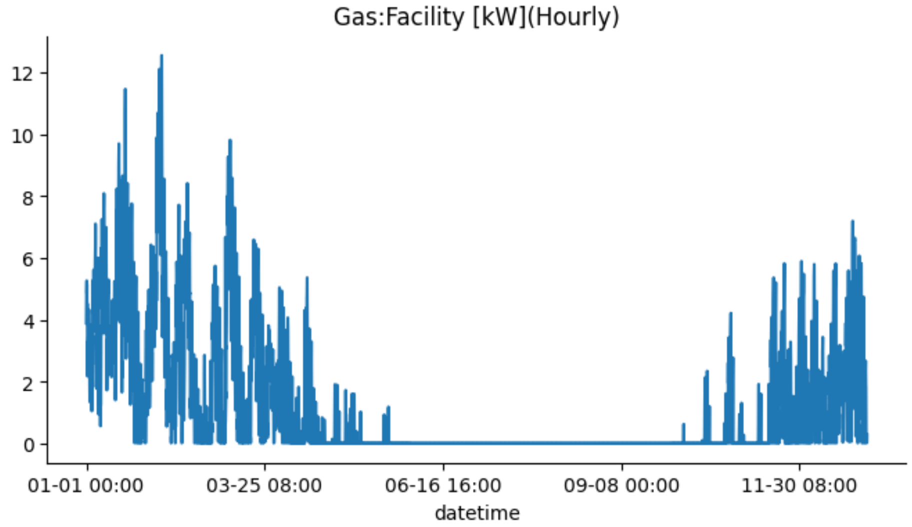
1. Merge different data sources – Check the code

- **Loading and Parsing:** Efficiently loads and parses time columns from two different data sources.
- **Resampling:** Converts minute-level data to hourly to match the granularity of the other dataset.
- **Merging:** Merges datasets based on datetime, ensuring no loss of information by filling NaNs with zeros.
- **Total Consumption:** Calculates total hourly consumption by summing all relevant columns.

2. Abnormal

a. Seasonal Variations in Cooling Electricity and Gas Usage





The two plots illustrate the hourly usage patterns for cooling electricity and gas over the course of a year.

Cooling Electricity Usage:

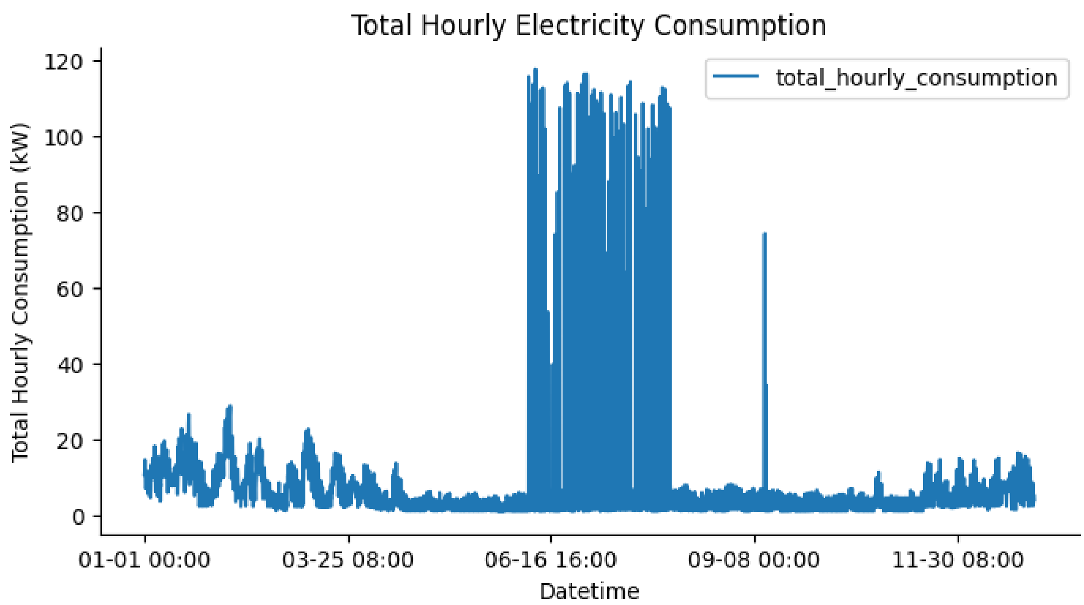
This plot shows that cooling electricity usage is concentrated between mid-June and mid-September. The usage peaks in this period, with the highest consumption reaching around 1.4 kW. Outside this period, the usage is negligible or zero, indicating minimal to no cooling needs during the rest of the year.

Gas Usage:

In contrast, gas usage is high during the colder months, from January to March and again from November to December. The usage peaks early in the year, with consumption levels reaching up to 12 kW. Between mid-June and mid-September, gas usage drops significantly to nearly zero, indicating little to no heating requirements during the warmer months.

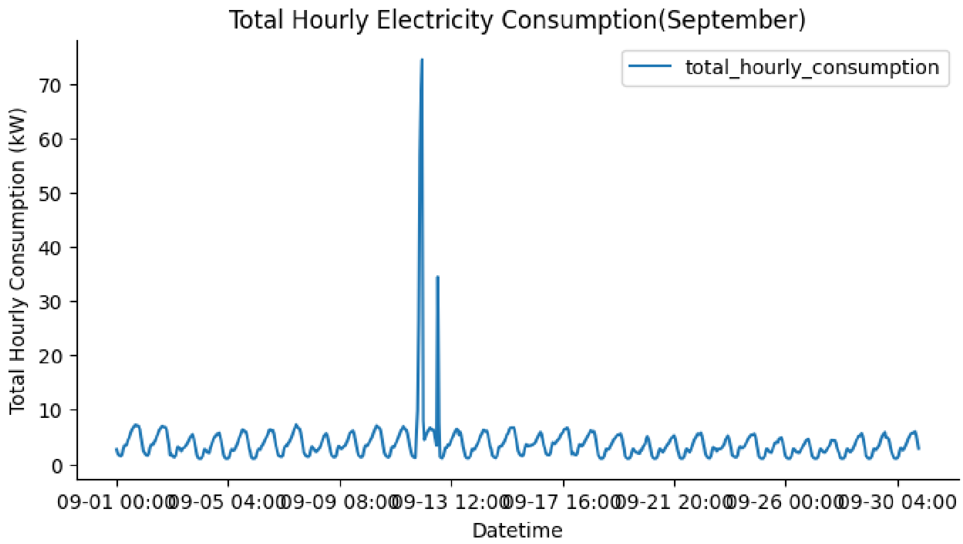
In summary, the cooling electricity usage and gas usage are inversely related: cooling demand is high during the summer months while gas usage is high during the winter months. This seasonal variation aligns with common sense, as cooling is needed in hot weather and heating is needed in cold weather.

b. Significant Mid-Year Spike in Electricity Consumption



From the plot, it's evident that the total hourly consumption sees a significant spike between June and September after merging these two datasets. This could be attributed to the usage of cooling equipment (appliance data from new.app4.csv) during the summer months. Another contributing factor might be the introduction and subsequent removal of a new instrument towards the end of the period.

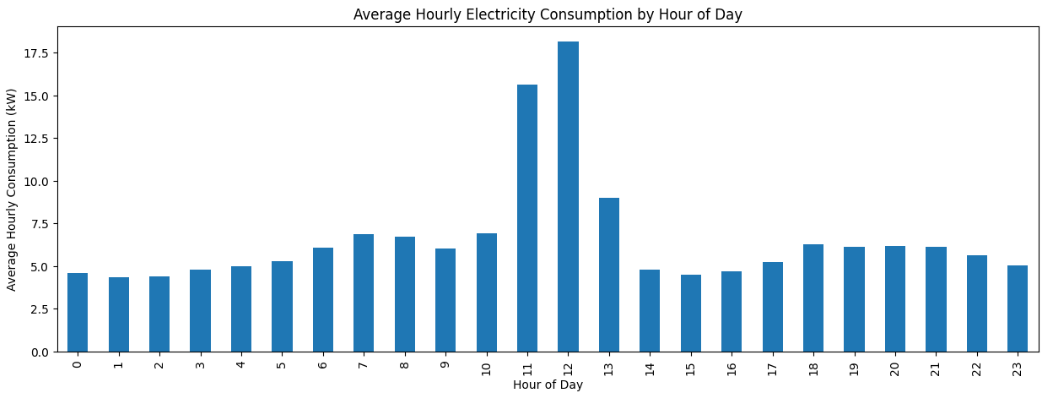
c. Single spike in September



The two spikes are temporary anomalies that significantly deviate from the regular consumption pattern.

Potential Causes: The spikes could be due to the introduction of a high-energy appliance or equipment that was used temporarily. An unusual event, such as maintenance or testing of new equipment, could also cause a temporary increase in electricity consumption.

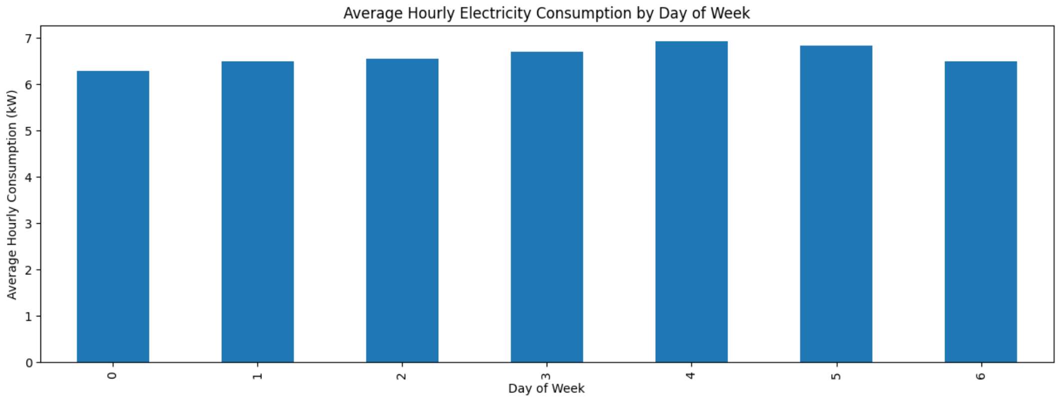
3. Summarize patterns by hour, weekday, and month



Hour of the Day:

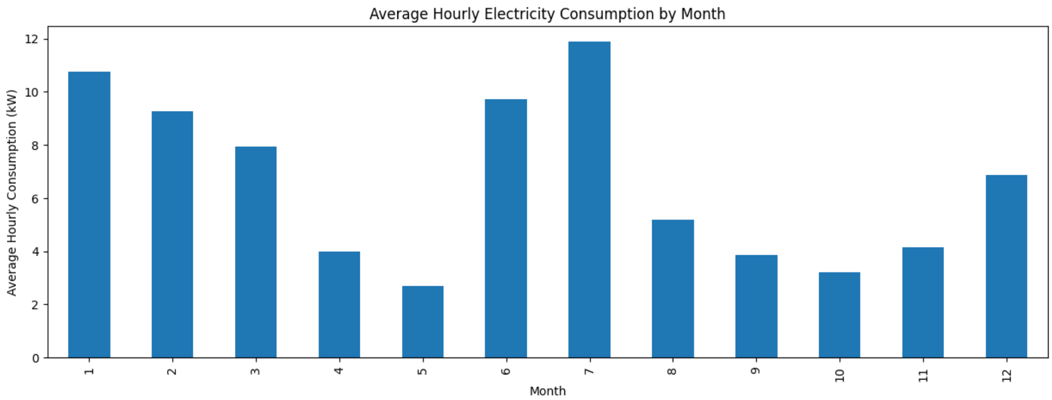
- ➢ Electricity consumption is lowest during the early hours of the morning (from midnight to 6 AM) and then gradually increases.
- ➢ A significant peak in consumption occurs between 12 PM and 1 PM, possibly due to increased activity during lunch hours.
- ➢ There is another smaller peak around 6 PM, potentially reflecting increased household usage as people return home in the evening.

➢ Consumption decreases gradually after 6 PM until it reaches its lowest levels late at night.



Day of the Week:

➢ There is a relatively consistent level of electricity consumption throughout the weekdays (Monday to Friday).



Month:

➢ Consumption decreases through February to April, likely reflecting milder weather reducing heating and cooling demands.

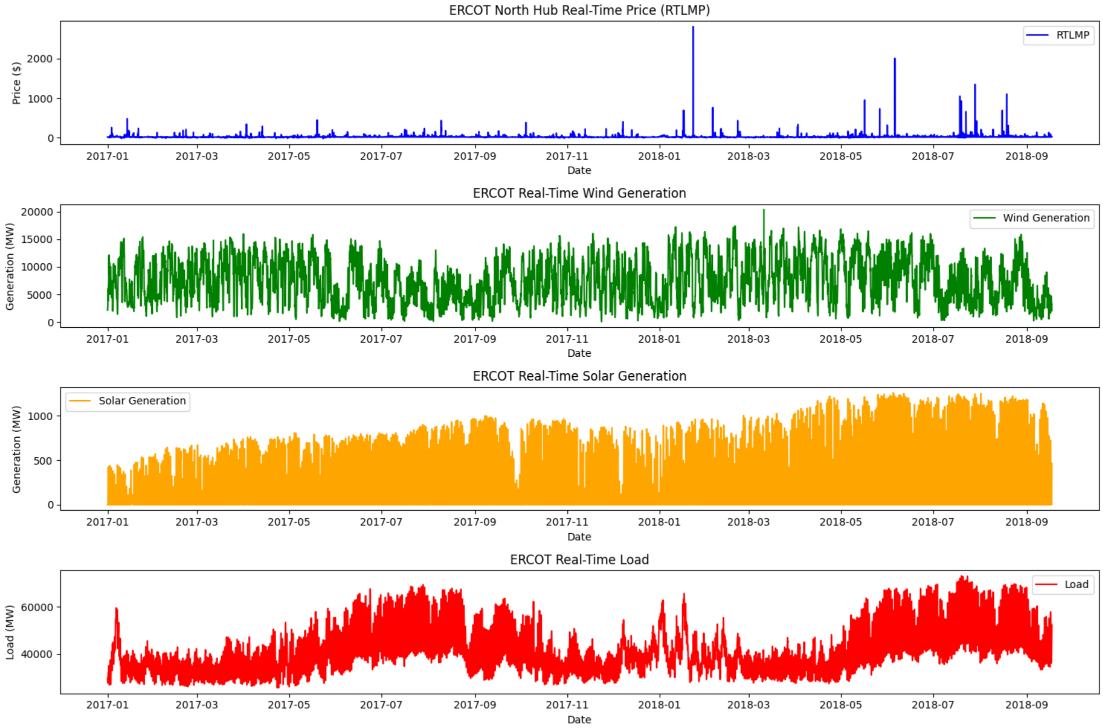
➢ A notable spike occurs in July, possibly due to the use of air conditioning in the hot summer months.

➢ Following July, there is a decrease through the cooler months of autumn (August to November).

➢ There is a slight increase again in December, likely due to heating requirements and possibly increased lighting and appliance use during the holiday season.

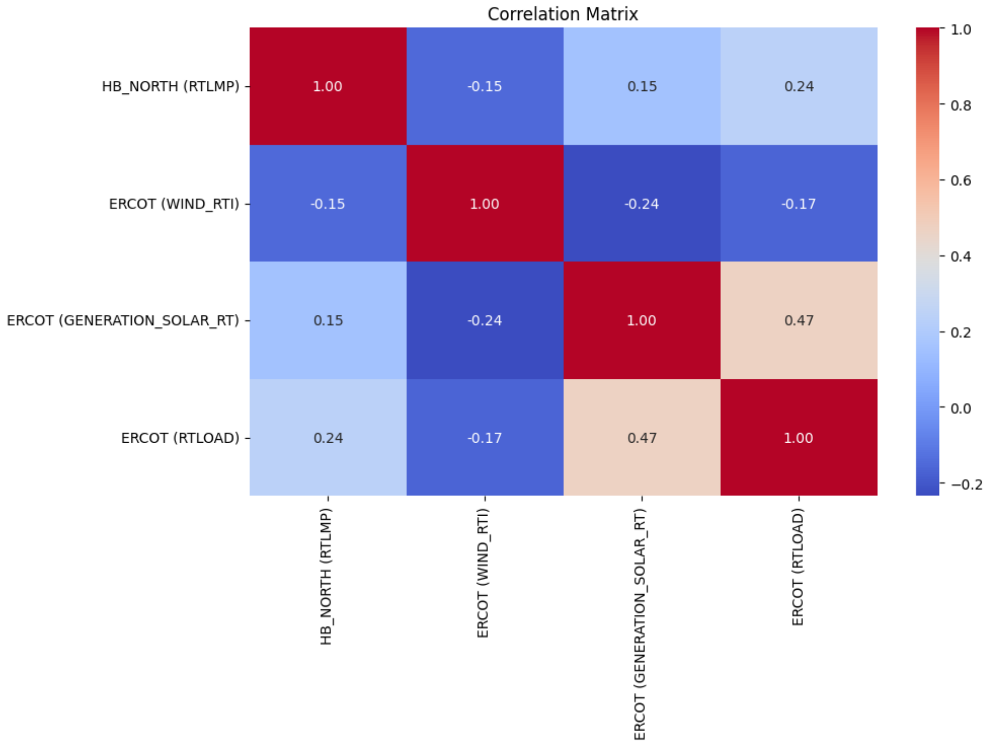
Assignment 3

1. Exploratory Data Analysis (EDA)

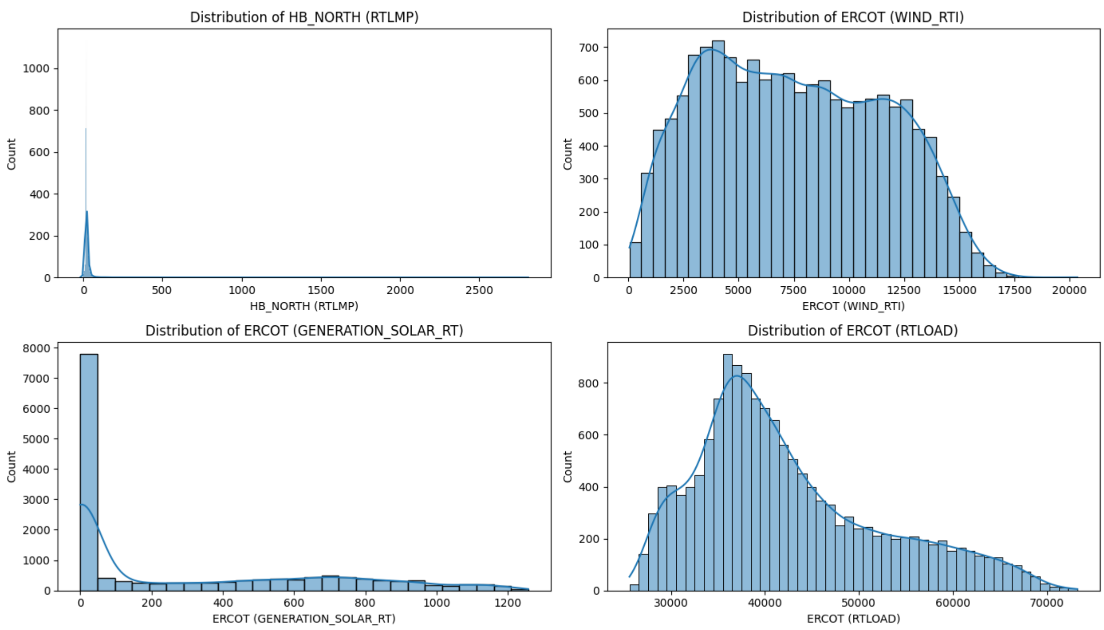


- RTLMP: Exhibits noticeable spikes and volatility.
- Wind Generation: Shows periodic fluctuations.
- Solar Generation: Characterized by many zero values reflecting nighttime hours.
- RTLoad: Appears to follow a more regular pattern, peaking at certain times.

	HB_NORTH (RTLMP)	ERCOT (WIND_RTI)	ERCOT (GENERATION_SOLAR_RT)	ERCOT (RTLOAD)
count	14987.000000	14982.000000	14983.000000	14987.000000
mean	25.766417	7532.436283	291.989714	42371.673703
std	46.361945	3992.884834	370.914596	9874.339631
min	-17.860000	54.440000	0.000000	25566.511248
25%	18.041250	4135.630000	0.000000	35431.636526
50%	20.057500	7281.445000	22.150000	39934.007113
75%	25.030000	10851.647500	608.635000	47873.100786
max	2809.357500	20350.400000	1257.540000	73264.662123



- RTLMP is negatively correlated with wind, suggesting that higher renewable generation tends to lower prices.
- RTLMP is positively correlated with real-time load and solar generation, indicating that higher demand increases prices.

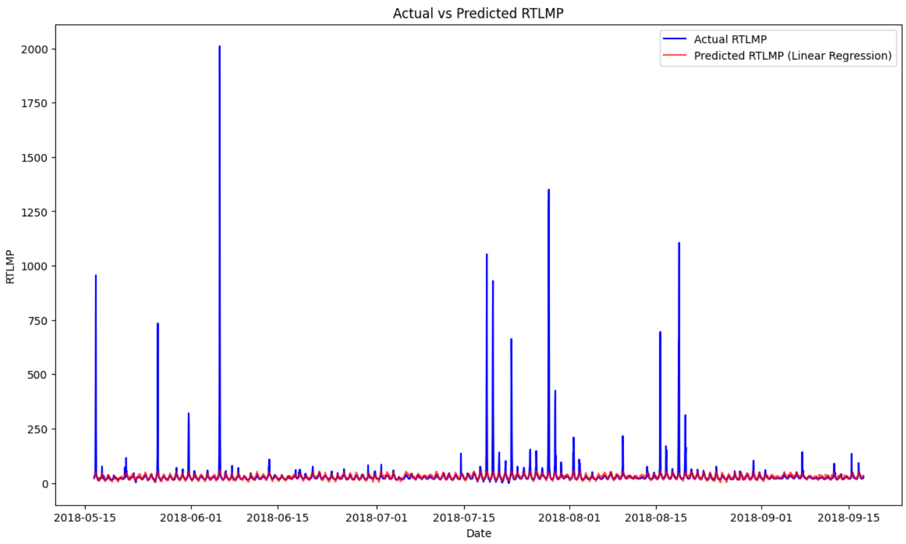


- RTLMP: Right-skewed with a few extreme values.
- Wind Generation: Right-skewed, with many hours having lower values.

- Solar Generation: Highly skewed with many zero values due to nighttime.
- RTLoad: More normally distributed but with significant variability.

2. Forecast Model

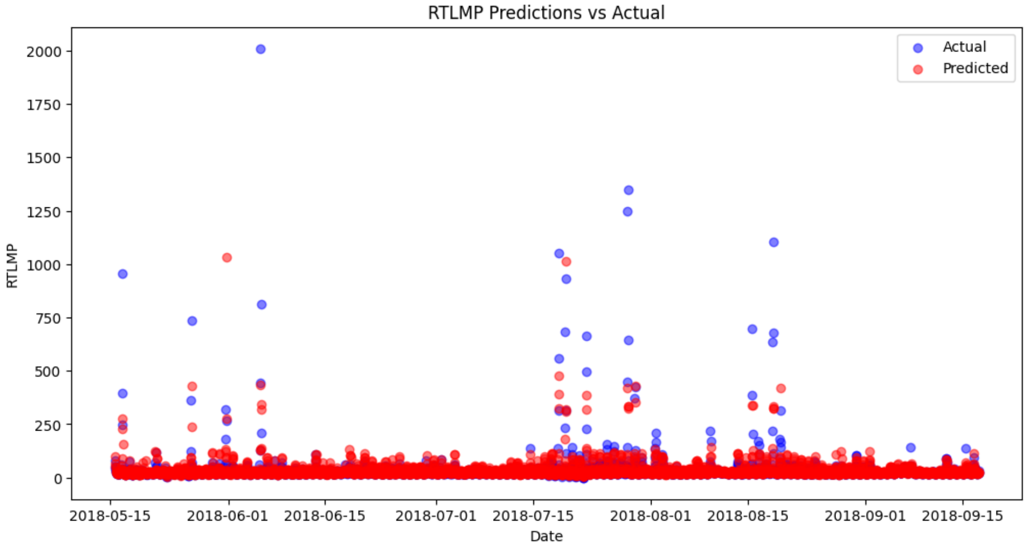
a. Linear Regression




```
Train RMSE: 33.20643176095252
Test RMSE: 74.0161874477612
Train MAE: 8.179845754673066
Test MAE: 12.668565251608067
```

The evaluation of the linear regression model shows that the Train RMSE and Train MAE are significantly lower than the Test RMSE and Test MAE, indicating that the model performs better on the training data than on the test data, suggesting potential overfitting. The relatively high Test RMSE and Test MAE indicate less accurate predictions on the test set, especially during spikes in the actual RTLMP values. Visual analysis reveals that while the predicted values (red line) generally follow the trend of the actual values (blue line), the model fails to capture the extreme spikes in RTLMP. The model performs reasonably well for lower and more consistent RTLMP values but struggles with high volatility and extreme values.

b. Randon Forest



Since the data is non-linear, I use Random Forest to make the forecast. To improve the prediction, I introduce one lag of RTLMP as a variable for the forecast. From the plot, we can see that the red points cover most of the normal blue points. Additionally, the red points also cover some of the blue points more effectively than the previous Linear Regression Model.

Assignment 4

Product 1: Power Futures - ERN

ERCOT North 345KV Real-Time Peak Fixed Price Future is a monthly cash-settled futures contract based on the average of daily peak hourly electricity prices at the ERCOT North 345KV Hub. The peak hours are defined as 7:00 AM to 10:00 PM Central Prevailing Time (CPT), Monday through Friday, excluding NERC holidays. The contract is traded in USD with a minimum price fluctuation of $0.01 per MWh. The contract symbol is ERN, and it can be traded up to 50 consecutive monthly periods. The final settlement price is based on the average of the specified prices published by ERCOT for the delivery period (Ice) (Ice).

Product 2: Natural Gas Futures - H

Henry Hub Natural Gas Futures (H) is a futures contract for natural gas delivery at the Henry Hub in Louisiana. These contracts are a benchmark for natural gas prices in North America and are widely used for hedging and speculative purposes. The settlement method is physical delivery of natural gas, and the contracts are traded in USD per million British thermal units (MMBtu). The contract size is 10,000 MMBtu, with a minimum price fluctuation of $0.001 per MMBtu (Ice).

Product 3: Heat Rate Futures - XPR

ERCOT North 345KV Physical Heat Rate Peak Futures is a futures contract that measures the efficiency of converting natural gas into electricity. The contract is traded based on the peak hours of 7:00 AM to 10:00 PM CPT at the ERCOT North 345KV Hub. The settlement is done through physical delivery, and the contract size is typically specified in MWh. The heat rate is calculated as the ratio of the electricity price to the natural gas price, providing a hedge against changes in both fuel prices and electricity prices (Ice).

Conclusion of Hedging

- Hedging Effectiveness: The hedging strategy effectively reduced the volatility of the portfolio compared to the unhedged physical power position. This was evident from the comparison of the hedged vs. unhedged positions over time.

- Weekly Rebalance: Regular adjustments to the natural gas futures position ensured the hedge remained responsive to market changes.

Overall, the hedging strategy demonstrated that using a combination of natural gas futures and heat rate futures can effectively manage the risks associated with physical power holdings. The approach of weekly rebalancing and the use of relevant hedge ratios were key factors in maintaining the hedge's performance over time. This exercise highlights the importance of dynamic hedging strategies and the ability to adapt to market conditions to achieve desired risk management outcomes.