Analysis of Electricity Grid Patterns & Performance

Prepared by: Yahya Fawzey

Date: August 18, 2025

View on GitHub: [Click here]

Executive Summary:

🎯 Project Goal:

The purpose of this project was to assess hourly electricity data from the three largest U.S. grid operators (i.e., CISO, ERCO and PJM) to extract basic consumption habits, track electricity flows between regions and evaluate forecast performance for demand.

💾 Data Source:

The data was sourced from the U.S. Energy Information Administration (EIA) API, covering the period from January 1, 2025, to August 17, 2025.

🔢 Key Steps

1. Data Ingestion – Fetched over 65,000 records from the EIA API using a pagination loop to handle the API's 5,000-record limit.
2. Data Cleaning & Preparation – Handled complex time zone issues by converting all UTC timestamps to the correct local time for each region.
3. Exploratory Data Analysis (EDA) – Analyzed and visualized daily, monthly, and seasonal demand patterns.
4. Comparative Analysis – Compared the demand profiles, seasonal peaks, and inter-regional power trading of the three grid operators.
5. Performance Evaluation – Assessed the accuracy of the day-ahead demand forecasts by calculating the Mean Absolute Percentage Error (MAPE) for each region.

🛠 Project Challenges & Solutions

Managing Unexpectedly Large Data:

- **Challenge:** The initial estimate for the dataset size was around 400,000 records, but we discovered the true size was over 1.5 million records. Downloading this blindly was time-consuming and inefficient.

  Solution: First, a diagnostic script was created to query the API for total record count, allowing for an informed decision before the full download of the data. Based on the large resulting count, it's strategic decision was to consider three relevant regions (CISO, PJM, ERCO), which allowed better focus on the project and made the data size more manageable.

- **Challenge:** Time Zone Inaccuracy – Initial plots showed a misleading late-night peak due to data being in UTC.

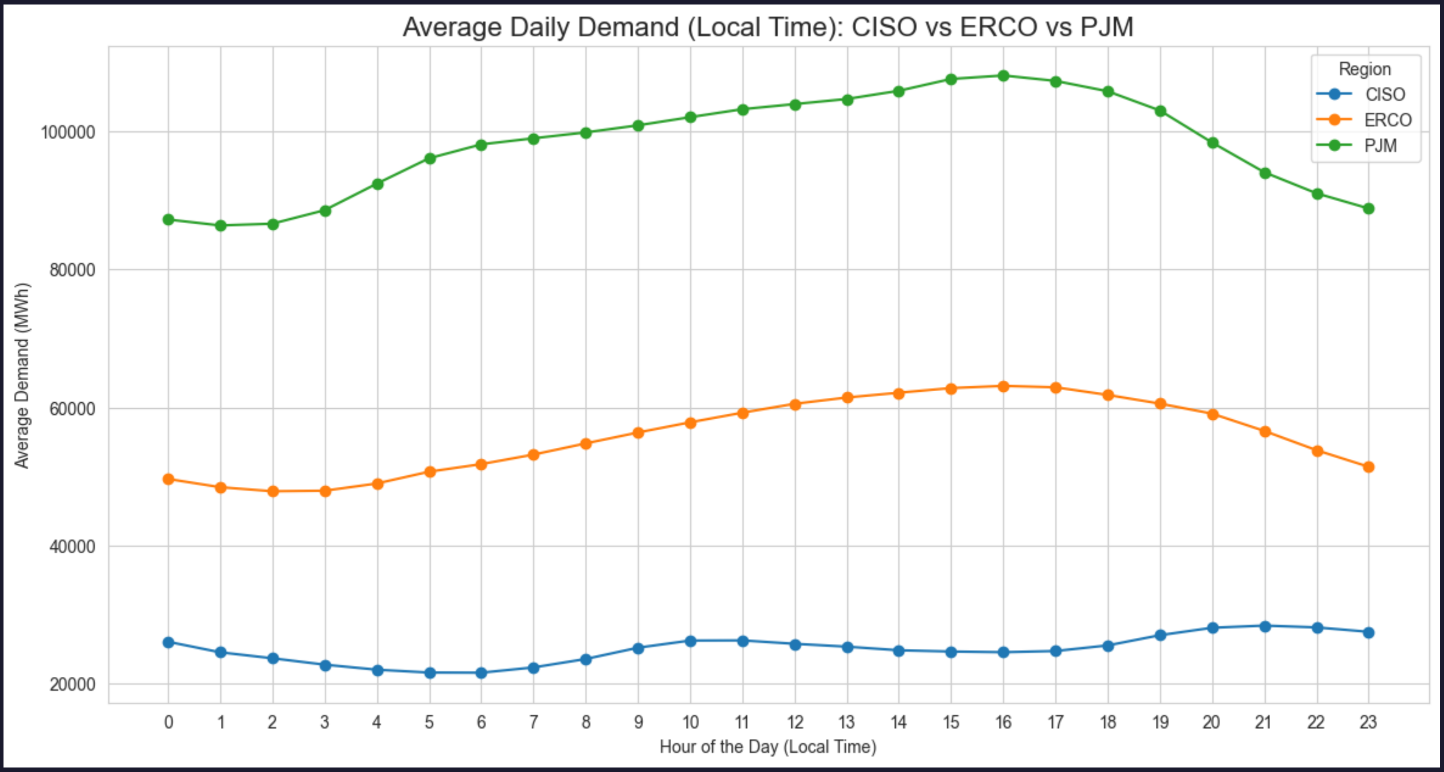
  Solution: Converted timestamps to the correct local time zone for each region, which fixed the demand curves.

💡 Key Insights

- **Peak Demand** – All regions show a clear evening peak between 4 PM and 7 PM local time.
- **Seasonal Patterns** – ERCO (Texas) and CISO (California) are *summer-peaking* grids, with ERCO's summer demand 20% higher than its winter demand. PJM is a *dual-peaking* grid with nearly identical demand in both seasons.
- **Grid Inter-dependency** – CISO is a major net importer of electricity, PJM is a major net exporter, and ERCO operates as a largely self-sufficient "energy island."
- **Forecast Accuracy** – ERCO had the most accurate forecast (MAPE ~2.5%), while CISO was the least accurate (MAPE >8%

Analysis & Findings:

Daily Demand Pattern:-

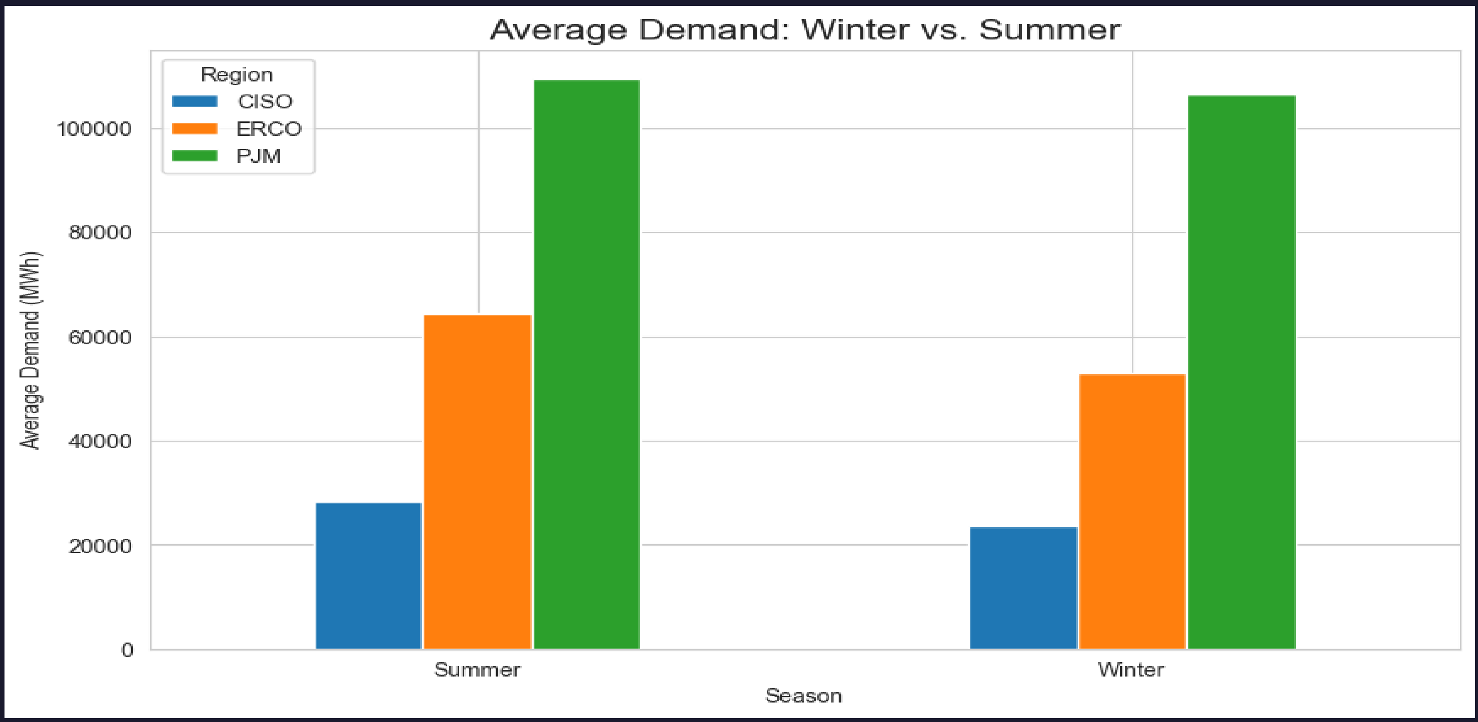


Key Insights:

- The plot clearly shows that for all three regions, the highest electricity demand consistently occurs in the late afternoon and early evening, between **4 PM and 7 PM local time**.
- **PJM** shows a particularly steep morning "ramp" in demand starting around 5 AM, reflecting a synchronized start to the business day across its dense metropolitan areas.
- **ERCO (Texas)** and **PJM** exhibit a very broad and sustained peak that remains high for several hours, a classic signature of regions with significant air conditioning load.

Actionable Decisions: This information gives grid operators lots of flexibility in optimized scheduling for power plants, having the ability to call on more expensive "peaker" plants during limited hours of peak demand, realizing fuel savings in the millions. If you are in a region like CISO (California), the data is also important for managing the "duck curve," in which solar power drops off just as this evening demand ramp begins.

Seasonal Comparison:-

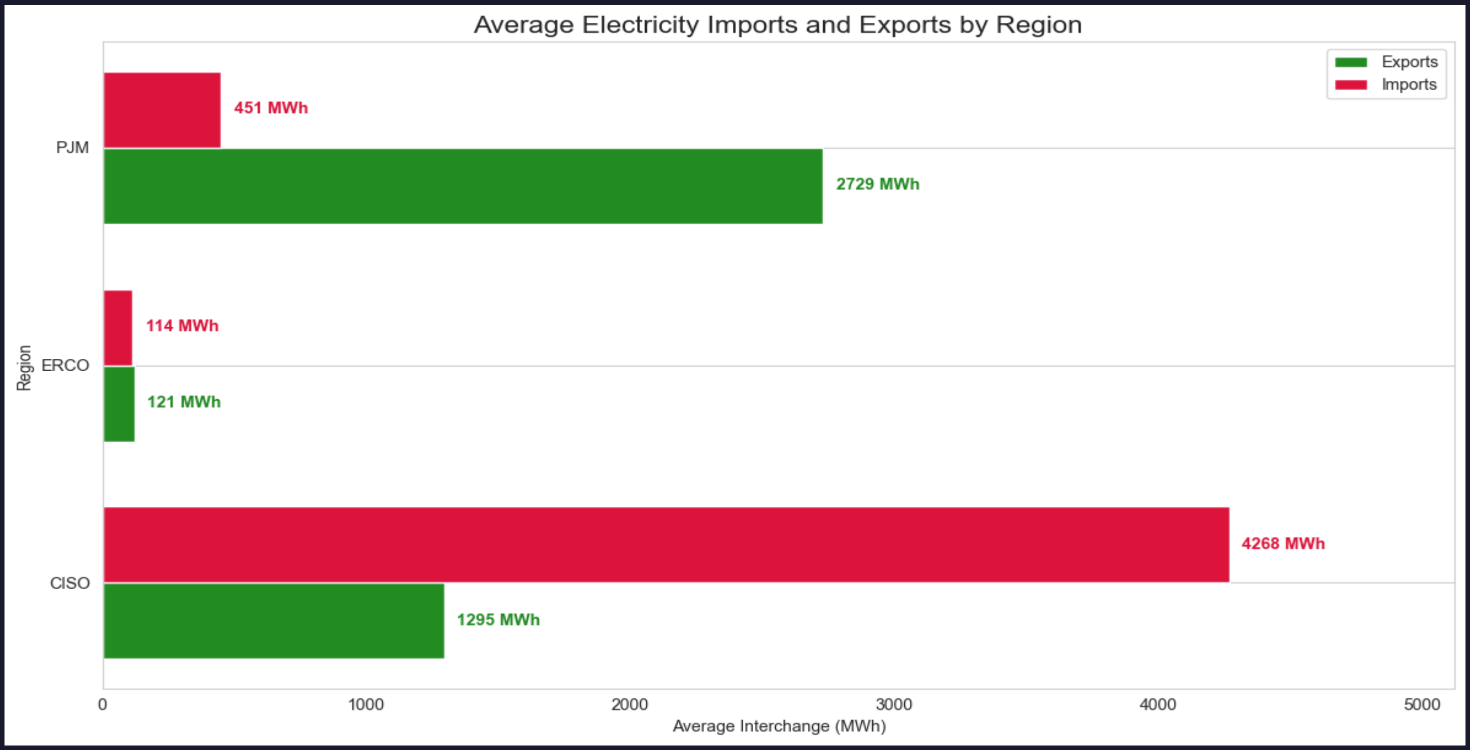


Key Insights:

- **ERCO (Texas)** and **CISO (California)** are "summer-peaking" grids, with significantly higher average demand in the summer due to air conditioning load.
- **PJM** is a "dual-peaking" grid, with nearly identical high demand in both winter (driven by electric heating) and summer, indicating a year-round strain on its resources.

Actionable Decisions: This analysis has implications for long-range strategic planning. Grids that peak in the summer, such as CISO and ERCO, have the spring as the best time period for maintenance. For dual-peaking grids like PJM, the maintenance windows become much tighter. This analysis also influences investment, as it indicates that CISO and ERCO need to invest in solar to address their summer peak energy requirements, while PJM needs to invest in reliable energy sources, like natural gas or nuclear energy, for year-round loads.

Interchange Analysis:-

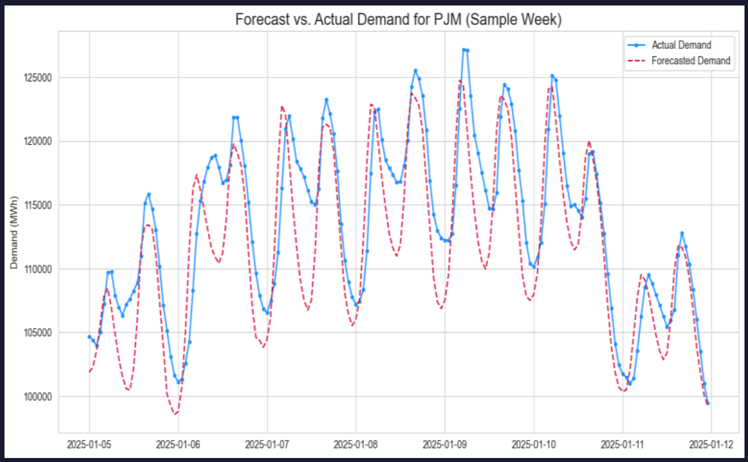


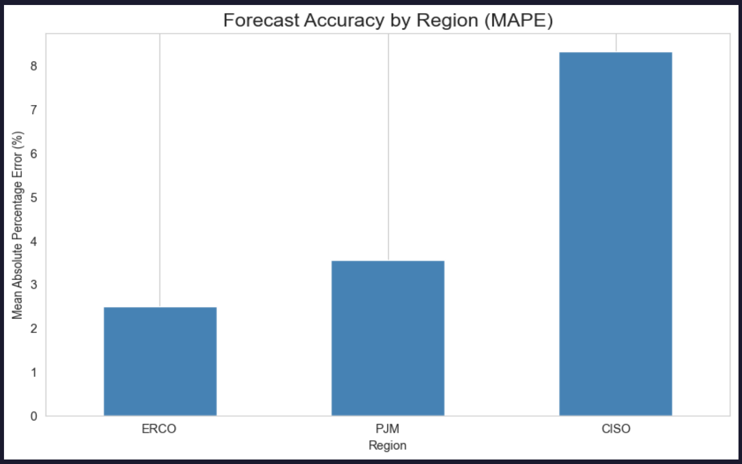
Key Insights:

- **CISO (California)** is a major **net importer**, buying significantly more power on average than it sells. This is largely driven by the need to import power in the evenings when solar generation ceases.
- **PJM (Mid-Atlantic)** is a major **net exporter**, consistently producing a surplus of electricity that it sells to neighboring grids.
- **ERCO (Texas)** is remarkably **self-sufficient**, with average imports and exports that are nearly identical and a tiny fraction of the other regions.

Actionable Decisions: This visualization demonstrates separate strategic vulnerabilities. CISO's import dependency presents operational cost and reliability risk and continues to drive investment in battery storage. PJM's export status is a consistent source of revenue and has optimization of generation for monetization, as the electric grid is not weighed down by reliability risk. ERCO's independence leaves it vulnerable during extreme weather, so speed-to-market policy decisions are driven by in-state grid resilience.

Forecast Accuracy Analysis:-





Key Insights:

- ERCO (Texas) has the most accurate forecast, with an average error of only ~2.5%, indicating high reliability in its predictions.
- CISO (California) has the least accurate forecast, with an error rate of over 8%. This highlights the difficulty of predicting demand in a grid with high penetration of variable renewables like solar.
- PJM's forecast is moderately accurate but was observed to systematically under-predict demand, especially at the peaks.

Actionable Decisions: This analysis helps drive operational enhancements. The systematic bias in PJM's forecast would prompt a recalibration of the model. CISO's high error rates would justify raising the "operating reserve" of standby power plants, serving as a “safety cushion” against uncertainties. ERCO's high accuracy allows for a more efficient schedule of plants, resulting in lower operational costs.

Conclusion

This assessment was able to unravel the intricate operating patterns of three important U.S. electricity grids. By adjusting for time zone issues and dealing with a large dataset from the API, it was able to expose the distinct daily and seasonal patterns for each region. The information on inter-regional dependency and forecast accuracy provides a compelling and data-endorsed basis for strategic planning associated with long-run grid reliability, monitoring costs, and investing in the energy transition as it constantly evolves.