Key Factors Driving Load Growth in PJM and ERCOT

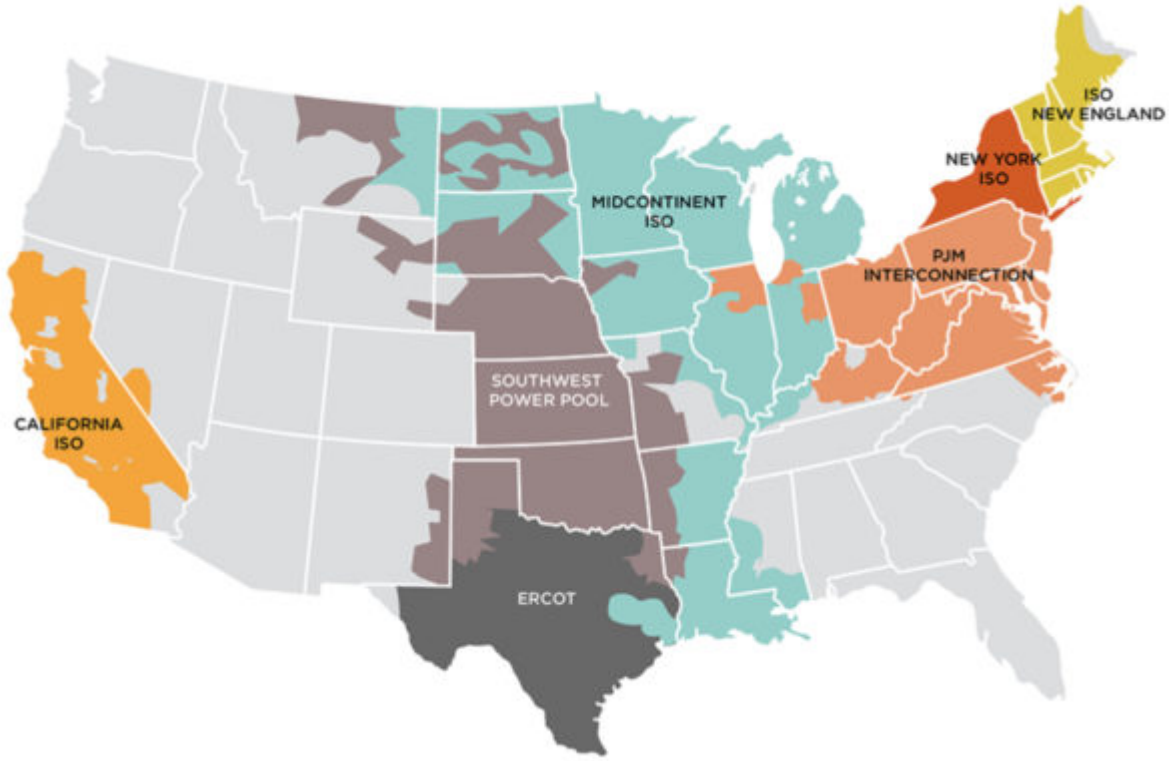
Hanna Karnei, Haru Koga, Jazmine Pritchett, Jaimie Wargo

Contents

Github Repository

https://github.com/jazpritch/KarneiKogaWargoPritchett_ENV872_EDA_FinalProject

1. Rationale and Research Questions



Written Section (Bold questions!)

2. Dataset Information and Wrangling

Written Section

2.1 PJM Load Data

Written Section

Table : Dataset Information for PJM

Detail	Description
Data Source:	JM Website - Metered Hourly Load
Retrieved From:	https://dataminer2.pjm.com/feed/hrl_load_metered/definition
Variables Used:	datetime_beginning_ept(Date), mw (Load_MW)
Date Range:	1/1/2010 - 12/31/2022

Table : Real-Time Dataset Information for PJM

Detail	Description
Data Source:	PJM website - Regulation Zone Preliminary Billing Data
Retrieved From:	https://dataminer2.pjm.com/feed/reg_zone_prelim_bill/definition
Variables Used:	datetime_beginning_ept(Date), mw (Load_MW), rmpcp (Price)
Date Range:	1/1/2013 - 12/31/2022

Table : Day-Ahead Dataset Information for PJM

Detail	Description
Data Source:	PJM website - Day-Ahead Scheduling Reserve
Retrieved From:	https://dataminer2.pjm.com/feed/dasr_results/definition
Variables Used:	datetime_beginning_ept (Date), mw (Load_MW), dasrmcp (Price)
Date Range:	1/1/2010 - 12/31/2022

2.2 ERCOT Load Data

Written Section

Table : Dataset Information for ERCOT

Detail	Description
Data Source:	ERCOT Hourly Load Data Archives
Retrieved From:	https://www.ercot.com/gridinfo/load/load_hist
Variables Used:	Hour_Ending, ERCOT
Date Range:	1/1/2010 - 12/31/2022

2.3 Demographic Data

Written Section

Table : ERCOT Linear Model Demographic Data

Detail	**Description**
Year:	2010 - 2022
Electricity Price:	Average prices of electricity to ultimate customers (https://www.eia.gov/totalenergy/data/monthly/pdf/sec9_11.pdf)
Median Income:	Income level at which half of the households in Texas earn more and half earn less (https://fred.stlouisfed.org/series/MEHOINUSTXA672N#)
Percent Renewable:	
Percent Minority:	Subtraction of white only percentages after dividing white popultion from totals (U.S. Census 2010-2022)
Population:	Total Population each year for Texas (https://www.census.gov/programs-surveys/popest/data/data-sets.html)
Mean Load:	Average load for ERCOT

3. Exploratory Analysis of Data

Written Section

3.1 Exploration of ERCOT

3.2 Exploration of PJM

4. Analysis

4.1 Time Series Analysis

HYPOTHESIS

4.2 Growth Analysis

4.3 Simple Linear Regression of PJM

HYPOTHESIS

Written Section

Table

Written Section

Table

Written Section

4.4 Multiple Linear Regression of ERCOT

HYPOTHESIS

To begin this analysis, we used Akaike's Information Criterion (AIC) to determine which explanatory variables should be used to prevent over fitting of the data. Table X outlines the results of this test. The lowest AIC value achieved was 173.58 with the average bill price and percent renewables as the chosen variables. This result can be confirmed by observing the correlation plot for the explanatory variables (Figure X).

Table : ERCOT Multiple Regression AIC Values

Variables	AIC Value
Mean_Load + population + percent_renew + med_income + percent_minority	178
average_price_tot + population + percent_renew + med_income	176.04
average_price_tot + percent_renew + med_income	174.26
average_price_tot + percent_renew	173.58

The multiple linear regression using average bill price and percent renewables as the explanatory variables for mean load in ERCOT has an adjusted R2 value of 0.9616, meaning that 96.16% of the variance in

average daily load can be explained by the explanatory variables. The p-value of this regression also indicates that the results of this model are statistically significant ($p < 0.001$, $df = 10$, $F = 151.1$). Table X shows the results of this regression. Assuming an alpha of 0.05, all variables are statistically significant, with percent renewables having the strongest significance with a p-value of 0.0001. For every 1% increase in renewable generation, there is a 358.75 MWh increase in average daily load. This is an interesting conclusion with no definitive explanation. It may have to do with the psychology around using more of a resource because it has a lower impact on the environment; for example, if someone has a solar heating system, they may take hot showers for longer because they feel that their energy consumption is cleaner. It could also be more of a technological advancement correlation since society is consuming more energy as we advance, and advancements in technology have resulted in more renewable energy to be captured (bigger wind turbines, more efficient solar panels, etc.) Overall, this finding warrants more study into this particular area.

We can also conclude that for every $1 increase in average bill price, there is a 1976.77 MWh increase in average daily load. Upon reflection, this last interpretation should likely have its causation reversed; instead, for every 1976.77 MWh added to the average daily load, there is a $1 increase in average bill price. This would make more sense because as load increases, more expensive generating plants need to be utilized, causing utility companies to have to charge higher rates to make a profit.

Table : ERCOT Multiple Regression P-Values

Variable	Coefficients	P-value
Intercept	15,132.37	0.0224
average_price	1,976.77	0.0083
percent_renew	358.75	0.0001

Figure X shows the residuals vs. fitted values, and Figure X shows the Q-Q residuals. These plots demonstrate the goodness of fit of the model. The residuals vs. fitted plot shows the distance of the residuals from the line of fit; it is important to ensure that these values are symmetric for a good fit, and this is the case with our model. The Q-Q plot points should stay close to the 1:1 line, and this is mostly true with our model as well, with some diversions towards the end with points 9 and 10.

5. Summary and Conclusions

6. References

Image Source: https://pv-magazine-usa.com/2022/06/29/faster-lower-cost-interconnection-by-combining-ercot-miso-pjm-approaches/