Temperature vs. Electricity Use in U.S. States

University of New Brunswick, Saint John

CS 2704: Data Analytics using Python

Professor Jong-Kyou Kim

Group Members:

Aryan Bansal

Aniket Shah

Dhruv Thumar

Introduction and Background

Over recent decades, rising global temperatures have led to an increase in cooling demand, particularly through the use of air conditioning systems. This higher demand places additional stress on electric power grids, especially during the summer months when peak consumption occurs. Understanding the relationship between temperature and electricity use at the state level is important for utility providers, urban planners, and policymakers. Insights from this relationship can support more accurate demand forecasting, guide infrastructure investments, and help design effective energy policies to maintain grid stability and manage costs.

The Hypothesis

We hypothesized that an increase in average ambient temperature would lead to higher electricity consumption, primarily driven by increased use of cooling systems. To measure cooling demand more accurately than by using average temperature alone, we calculated Cooling Degree Days (CDD) for each state and year using the formula:

$CDD = \max(0, AvgTemp - 65°F)$.

CDD represents the cumulative number of degrees above 65°F, serving as an indicator of cooling energy requirements. By comparing models based on average temperature and CDD, we aimed to determine which metric better explains electricity consumption trends.

The Analysis and Implication

We merged two datasets, NOAA's annual average temperatures and the U.S. Energy Information Administration's yearly electricity consumption, by state and year. After addressing missing values and calculating CDD, we applied a logarithmic transformation to the electricity consumption data to address skewness.

Initial exploratory analysis showed a moderate positive correlation between average temperature and log-transformed electricity consumption, with a Pearson's correlation coefficient of approximately 0.55. This justified proceeding with regression modeling.

A simple linear regression model using only average temperature as the predictor was statistically significant ($p < 0.05$), but it explained only about 7% of the variance ($R^2 \approx 0.07$) in electricity consumption. When we introduced CDD as an additional predictor in a multiple linear regression model, the explanatory power improved slightly ($R^2 \approx 0.071$). Within this model, the coefficient for temperature was −9.64 and for CDD was 102.17, indicating that CDD had a stronger relationship with electricity use than raw temperature.

Regional patterns were consistent with expectations. Warmer states such as Texas and Arizona showed higher electricity consumption, while colder states like Maine and Vermont deviated from the trend, likely due to greater heating requirements. Although incorporating CDD improved the model slightly, the overall explanatory power remained limited, suggesting that other factors, such as building insulation quality, industrial consumption, and socioeconomic variables, also significantly influence electricity usage.

Overall, integrating CDD into forecasting models offers a meaningful improvement over using temperature alone, supporting better short-term demand planning and policy development for energy management.

Conclusion

This analysis confirms a positive association between temperature and electricity consumption across U.S. states, with Cooling Degree Days providing a stronger predictor of energy use than average temperature alone. Although the statistical models were significant, they explained only a small portion of the variance, highlighting the importance of considering additional variables when predicting electricity consumption.

Nonetheless, including CDD in demand forecasting can help utility companies and policymakers make more informed decisions to improve grid reliability and plan for future energy needs. Future research should expand on this work by incorporating factors such as demographics, building standards, and technological adoption to build more comprehensive and accurate predictive models.

References

Project GitHub Repository: https://github.com/dhruvthumar/2704project

NOAA Climate Data: https://www.ncdc.noaa.gov/cdo-web/

U.S. Energy Information Administration (EIA) Data: https://www.eia.gov/electricity/data.php