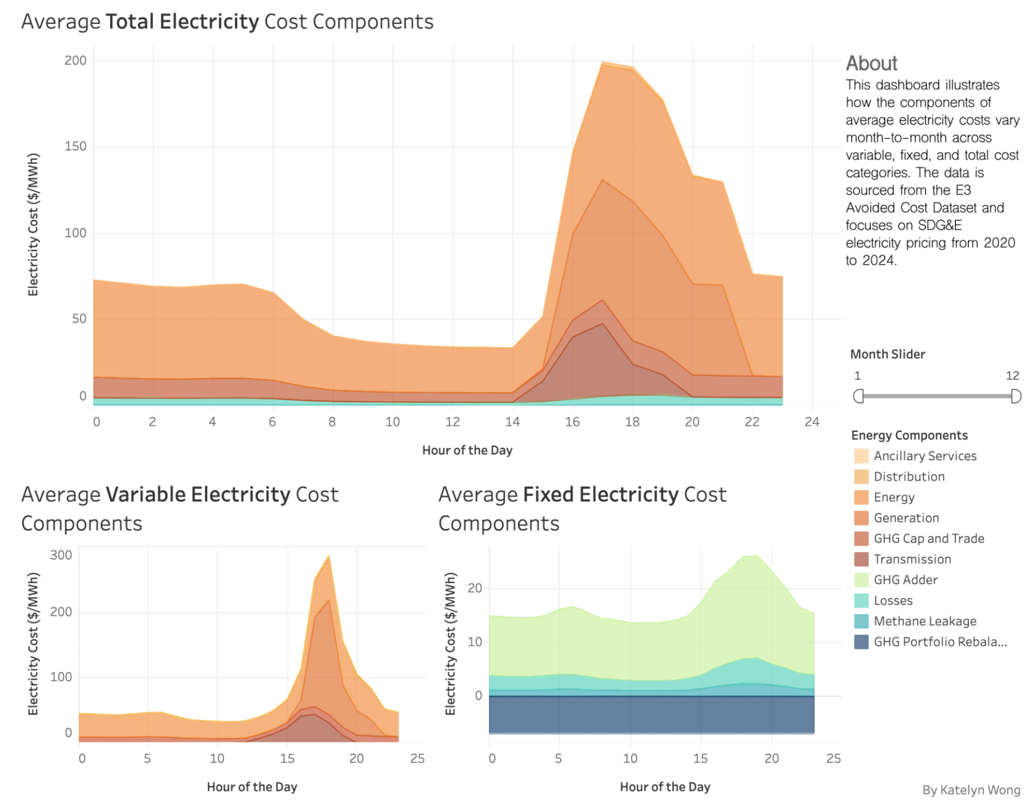
Optimizing Electricity Tariff Structures to Minimize Market Inefficiencies and Promote Equity and Sustainability

By Katelyn Wong



[Interactive Tableau Dashboard]

Consumers today often pay significantly more than the actual cost of producing, transmitting, and distributing electricity. When electricity prices fail to reflect real-time supply and demand, particularly in the fossil-fuel and renewable energy markets, consumers are not incentivized to shift their usage to times when clean energy is abundant. This mispricing leads to inefficiencies in the electricity market by discouraging the use of renewable energy sources and disproportionately affects lower-income households. The result is a system that rewards inefficient electricity uses and penalizes those least able to adapt.

My research focuses on optimizing electricity tariff rates using the Avoided Cost time series data from the Energy + Environmental Economics (E3) consulting firm. This dataset provides detailed records of electricity pricing and demand patterns across the climate zones in CAISO (California Independent System Operator), broken down into key electricity value components. The goal is to propose a dynamic pricing model that improves cost-efficiency, promotes renewable energy usage, and enhances transparency so that consumers can clearly see what they're paying for. Ultimately, this will encourage the transition to fairer pricing strategies and reduce instances where utilities overcharge customers.

Due to the complexity and scale of the CAISO dataset, careful data cleaning, modeling, and visualization were essential. Initially, I attempted to process the data using my personal laptop but quickly ran into limitations in computational power and efficiency. To work within these constraints, I was only able to analyze data from a single climate zone for 2024. However, my research required multi-year analysis across over 15 climate zones, along with the development of predictive models to optimize tariff structures and minimize excess consumer costs. This challenge led me to UCSD's Data & GIS Lab, where access to high-performance data analysis tools and software enabled me to scale up my analysis to include all SDG&E climate zones across five different years, significantly broadening the scope and impact of my analysis.

In the Data & GIS Lab, I used Anaconda Python as the main environment for data preprocessing and exploratory data analysis. The Pandas and NumPy libraries were important for managing and transforming large datasets, while Seaborn and Matplotlib enabled effective visualization of daily and seasonal price fluctuations. I used Scikit-learn for regression modeling and machine learning techniques such as K-Nearest Neighbors regression and Gaussian basis modeling. These helped identify peak pricing patterns for building hybrid regression models. For deeper multivariate analysis, I used the Statsmodels package to perform ordinary least squares regressions, which isolated important price determinants and informed predictions of fixed, marginal, and total electricity costs.

After establishing a strong exploratory data analysis foundation for the project and creating several predictive models in Python, I turned to Stata for running additional linear and logistic multivariable regressions to evaluate pricing responses to temporal and geographic factors. This allowed me to cross validate the statistical significance of my models generated in Python. After completing the modeling phase, I used the Lab's Tableau Public platform to create visualizations that presented my results in a clear, accessible way for a broader audience. I also reviewed academic literature on existing electricity pricing structures and market models to contextualize my findings, ensuring that my conclusions aligned with established research and the dynamics of the current electricity market.

So far, applying the resources from the Data & GIS lab to my research has allowed me to work beyond my personal computing limitations and taught me how valuable the UCSD Library's technical services are, especially for interdisciplinary research involving economics, policy, and data science. Through this project, I also gained experience in creating effective, interactive data visualizations using Tableau that can help consumers and policymakers understand complex pricing structures. The Geisel librarians also supported me in my research by helping me navigate new tools and maximize the use of the lab's software.

Looking ahead, I plan to expand my research by testing my proposed pricing models in other regional electricity markets outside of CAISO. This next phase will benefit significantly from the Lab's ArcGIS software, which will allow me to map utility territories and electricity infrastructure across different states, visualize socioeconomic disparities using income and energy usage data, and identify regions with high energy burden and grid congestion zones for further peak pricing model analysis.

By combining data science with spatial and economic analysis, I hope to contribute towards fairer, more transparent, and sustainable energy systems.