We were tasked with analyzing electric usage data that spanned two years from eighty-eight buildings on campus equipped with smart meters. Long-term goals for the project included identifying highest users on campus and eventually recommending measures to reduce energy consumption and save money. Each building has one to six meters collecting data every thirty minutes, totalling over three hundred meters and seven million entries.

We began by exploring the data. We initially sorted the buildings by category and then plotted their usage against temperature and time to see if we could correlate usage with either of those variables. Most of the individual meter data showed a correlation with time, where usage began to increase early in the morning, peaked from midday to late afternoon, and then dipped at night. Aggregate data for the building categories also revealed a correlation between usage and temperature, although some categories like laboratories showed a stronger correlation than others. Another variable we looked at was building age; we graphed raw and normalized (by area) cumulative usage for the buildings against their construction date and found no correlation. We then looked at cumulative and normalized data and began to identify the top fifteen highest users while also looking at individual meter data to identify trends that might reveal what each meter was measuring. To properly gauge the university's energy efficiency, we compared its data with Cornell University's energy data and national energy data. We found that the university is much more efficient than Cornell University across all building categories, but it performs poorly when compared to national averages, indicating that there is plenty of room for improvement.

After our initial exploration, we investigated ways to aid the university -- its administration and those who impact its electric bill -- in understanding and reducing campus energy consumption. One approach constituted an anomaly detection program, which aggregates pertinent historical data to predict normal usage and compares that prediction with real-time meter readings. We tested various variables that we identified in our initial exploration and found temperature, school term, and whether a given day occurs during the workweek to be most significant. Following evaluation using recent data, the model appears useful for identifying unreasonably high usage quickly.

We also analyzed the characteristics of specific buildings' usage to develop recommendations for the university's facilities department to implement. Since laboratories use far more electricity per square foot on an annual basis than most other types of campus buildings, we identified numerous energy-saving measures such as a sash-closing campaign for fume hoods and the installation of motion-sensitive lighting in laboratories, changes that would not impede research. Dormitories, while lower-profile users than laboratories, maintain EUIs (energy use intensities) well above national medians provided by the Department of Energy, so we recommended reviving and strengthening a long-abandoned competition between dormitory buildings while integrating the anomaly detection program to identify differences between dormitory's usages and historical norms. Additionally, we recommended campus-wide initiatives. For instance, we found that installing energy-efficient lighting -- a relatively simple solution -- would pay for itself through energy savings within two to five years.

In the end, having discovered the utility as well as the shortcomings of the data at our disposal, we recommended the university rename its meters and aggregate information about the area each meter covers. Furthermore, we recommended the university consider making real-time and historical energy data available for researchers and interested members of the university community to promote collaboration, transparency, and, ultimately, efficiency.