Effects of weather conditions on energy usage

Pereviazkin Egor, 2134292

Introduction

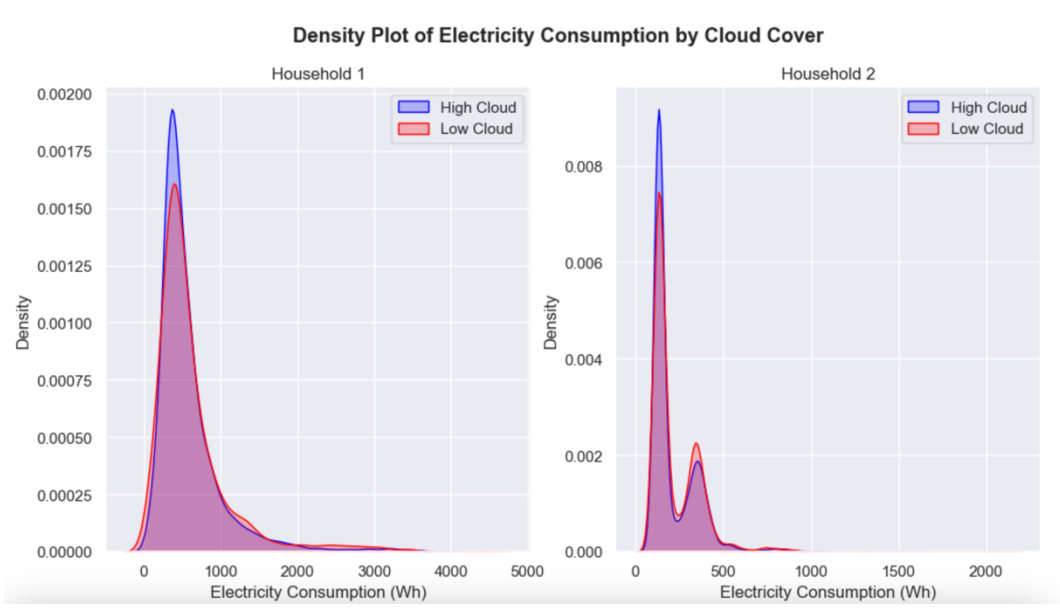
During the completion of GA1 I found out that electricity and gas consumption may depend on different weather conditions, and that produced my problem statement: ‘Do weather conditions, different from temperature and wind speed, affect energy consumption?’, and refined hypothesis: "The average electricity consumption in 2013 was higher during hours with a cloud cover of 5 octants or more compared to hours with less than 5 octants of cloud cover.” This refined hypothesis is based on the idea that during cloudy days (with cloud cover at 5 octants or higher), households would require more electricity due to the reduced natural sunlight.

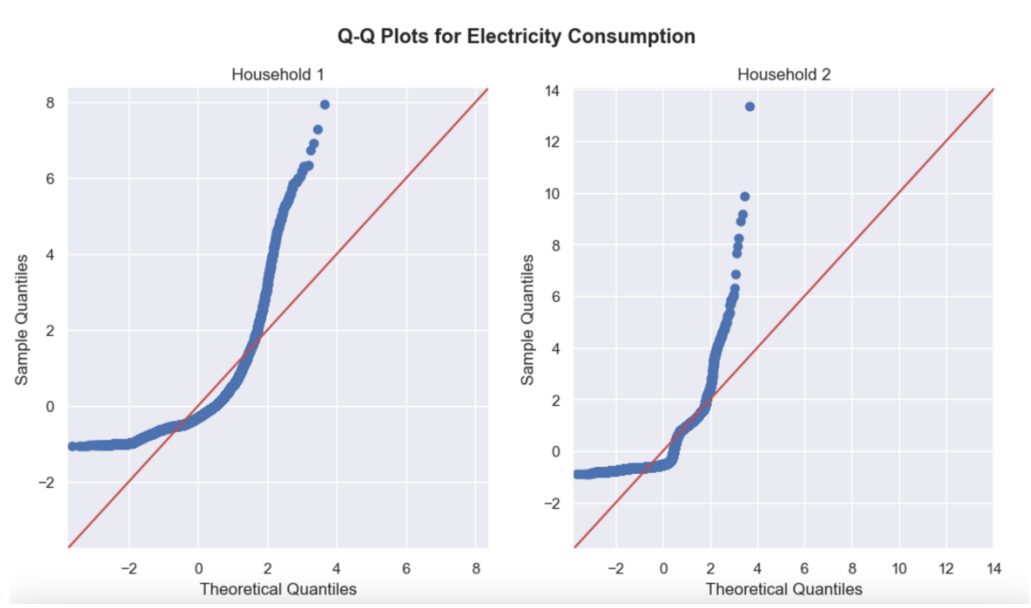
Methods

I removed all the readings with missing values of cloud cover and gas or electricity consumption to avoid unwanted effects on data analysis, incorrect conclusions and issues with carrying t-tests. Additionally, by checking periods with long-term extreme deviations of consumption readings from average electricity and gas usage, I found atypical data that could be caused by faulty equipment, extreme weather conditions and holiday activities. These records can significantly influence further data analysis and should be withdrawn.

Results

We aim to test if there is a significant difference in electricity consumption between periods with high cloud cover (≥ 5 octants) and low cloud cover (< 5 octants) in two households in 2013. We performed an independent two-sample t-test using the ttest_ind function because we have two independent samples and we want to compare their means. Since the p-values of household 1 (0.021) and household 2 (0.0036) are less than significance level 0.05, we reject the null hypothesis for both households. This means there is a statistically significant difference in electricity consumption between periods of high and low cloud cover observed in both households. KDE plots bellow show that household 1 and 2 support the use of the t-test because of their bell-shaped distribution, but second peak in the right plot suggests that data should be examined more closely. Q-Q plots suggest small to moderate deviations from normality, primarily caused by high-consumption outliers.





Discussion

The choice of using a t-test is appropriate for comparing the means of electricity consumption in high-cloud and low-cloud conditions, ensuring that any difference observed is statistically significant and not due to random chance. Furthermore, the process of data cleaning helps to ensure the robustness of the results. Electricity consumption is likely influenced by the time of day and other confounding variables such temperature. Removing atypical data also introduces the risk of missing unique behaviors (home renovation or parties). In order to improve reliability of the results I could take into account confounding variables, deeper analyze atypical data and causes of second peak in KDE plot of electricity consumption of household 2, and include more households into investigation.

Conclusion

Answering the problem statement I can conclude that weather conditions (in particular, cloud cover) significantly influence electricity consumption in households, with higher electricity usage observed during periods of heavy cloud cover. Understanding these relationships can inform strategies for energy efficiency, particularly in regions with frequent cloud cover. The findings from this analysis provide valuable insights for further exploring the impact of weather conditions on energy consumption.