

Case study: Analysis of smart meter data

Sayani Gupta

25/08/2019

Smart meters provide large quantities of measurements on energy usage for households across Australia, and indeed many parts of the world. Households are distributed geographically and have different demographic properties such as the existence of solar panels, central heating or air conditioning. The behavioral patterns in households vary substantially, for example, some families use a dryer for their clothes while others hang them on a line, and some households might consist of night owls, while others are morning larks.

It is common to see aggregates of usage across households, total kwh used each half-hour by state, for example, because energy companies need to understand maximum loads that they will have to plan ahead to accommodate. But studying overall energy use hides the distributions of usage at finer scales, and making it more difficult to find solutions to improve energy efficiency.

One of the customer trial [smart-meter] conducted as part of the Smart Grid Smart City (SGSC) project (2010-2014) in Newcastle, New South Wales and some parts of Sydney provides customer wise data on half-hourly energy usage and detailed information on appliance use, climate, retail and distributor product offers, and other related factors. It would be interesting to explore the energy consumption distribution for these customers and gain more insights on their energy behavior which are otherwise lost either due to aggregation or looking only at coarser temporal units. The idea here is to show how looking at the time across different granularities together can help identify different behavioral patterns.

Let us compare the energy consumption distribution for two households from the customer trials. For this analysis, we only take the data for the year 2013. The first household has data from April and the second household has data from March. For the purpose of comparison, data for both households have been observed from April. Let us look at the distribution of energy across coarser temporal granularities and then deep dive into finer temporal granularities.

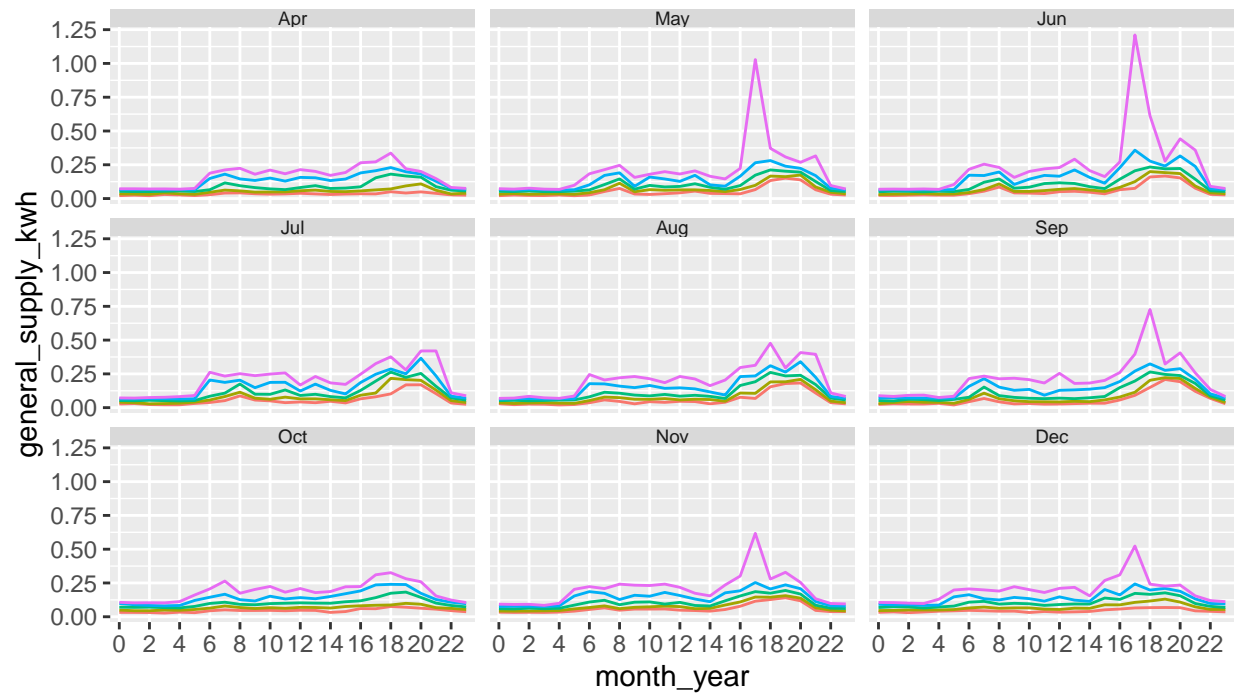
- Hourly usage across months of the year

Figure 1 shows the distribution of hourly consumption of both households across different months of the year. The two households are similar when it comes to their usage in the morning hours. The first household's extreme behavior is different from that of the second household in the evening hours of May and June (represented by higher 90th percentile energy usage). However, the 75th percentile of energy consumption for those two months are higher for the second household than for the first and their median consumption is more or less similar. This implies that consuming more energy in the evening is more typical for household 2. For the evenings of October, November, and December, it is clear that household 1 typically consumes much less energy than household 2. Moreover, the second household has a steep evening peak (75th percentile above 0.50 Kwh) across most months in contrast to the first household who has a steep evening peak only in the month of May and June with 75th percentile consumption falling below 0.25 Kwh. This might be an indication that of usage of a heater in the winter months for the first household, whereas the second household uses both air conditioner in hot months and heater in cold months or the number of members in the household vary widely. All of these can be explained with more demographics of the household and weather conditions in that area.

Let us dive deeper and see the hourly usage across different days of the week for these two households.

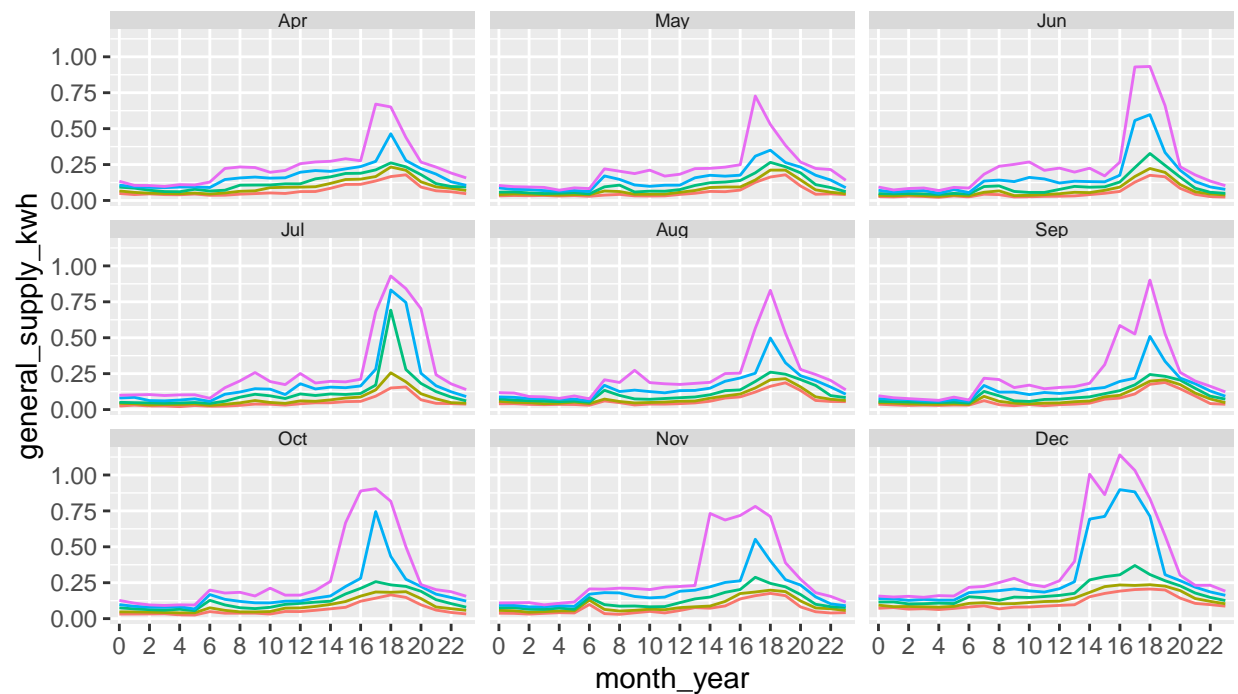
- Hourly usage across days of the week

Figure 2 shows the distribution of hourly energy usage across different days of the week through violin plots for both households. Energy consumption typically goes up between 16 to 18 hours. On weekends, the energy consumption for these hours is even higher. The straight lines in the violin plot for weekends at 17th and 18th hours suggest that the distribution of energy consumption is uniform demonstrating that any energy value has equal probability in falling in that hour. This suggests that these hours are extremely volatile based



quantile 10% 25% 50% 75% 90%

Quantile plot across hour-of-day by month-of-year for Household 1



quantile 10% 25% 50% 75% 90%

Figure 1: Quantile plots of hourly consumption across different months for the first household. 10th, 25th, 50th, 75th and 90th percentile are shown. For example, 90th percentile might denote behavior which are extreme and 50th percentile corresponds to a more typical behavior. Some extreme behavior is exhibited in the evening hours of May and June for household 1, however, overall energy consumption does not seem to vary much across months. Median and 75th percentile is less than 0.25 Kwh across most months. For household 2, distinct evening peaks for observed for few evening hours across all months. The median consumption also varies across months. They are likely to be using heaters in colder months and air conditioners in hot months like January.

on weather conditions/months. Energy consumption distribution in the morning hours of weekdays has a bulge between 0 to 0.25 KWh, indicating that if we pick up any typical weekday, we can say with reasonable certainty that the energy consumption would fall in that range. This is not true for the 17th or 18th hour of the day. The consistent behavior in the morning hours can be attributed to the fact that members in this household are not stay-at-home parents/ working from home professionals.

For the second household, the distribution of hourly energy usage across different days of the week through violin plots for Household 2. Contrary to household 1, household 2's energy consumption increases through noon to 7 pm. We can't say with certainty if the morning hours from 10 am would manifest a higher or lower level of energy consumption for this household. This indicates that this household might have kids or someone working from home.

- Across days of the week by months

A plot of the distribution across months of the year and days of the week might throw some light on the extreme behavior we saw in May and June. Was it only for the months May and June, that this household consumed more energy than they usually do ? Figure 3 shows that it is not only those months that saw high energy consumption on Saturday and Sundays. In fact, months like November and December saw higher energy usage than typical even on other weekdays. Still, for this household, it can be safely said that most of their activities which involve higher energy consumption are concentrated on weekends. The behavior of the weekends, as can be observed from the letter values of energy consumption, are different across months which can be attributed to weather conditions.

Figure ?? shows that the second household uniformly distributes their regular activity like using dishwasher or heater/air conditioner throughout the week. The distribution of energy consumption on Monday, Friday and Sunday look quite similar for December. Also, for other months, it is not very clear if weekend behavior varies distinctly from weekday behavior.

- Across week of the month and weekends

Another aspect which might of interest is how these households distribute their monthly chores between Saturday and Sunday across weeks of the month. "Monthly" chores might include washing mattress, comforters, bed and pillow covers, vacuum cloth furniture or cleaning dishwasher or oven. Also some households might visit friends and family or have them come over on weekends at least once a month. Figures 4 and ?? shows that the letter value plot of energy consumption for Saturday and Sunday across weeks of the months.

Figure 4 shows that median consumption of energy for both households. For household 1, it is higher on Sundays. Also Sundays have longer right-tails than Saturdays across all weeks of the month implying that Sunday is a little more active day for the household. And the forth weekend of the month, either they take it easy on themselves or stay out of their house for trips/visit friends. Household 2 has all-absorbing Sundays. They are either busy with weekly chores or have the entire family stay at home every Sunday. The first Saturday of the month looks pretty consistent compared to any other weeks of the month, implying either they are out for work/trips/visiting family. Also, 4th Sunday of the month has a long right tail implying they choose the last week of the month for doing their "monthly" chores.

To summarise, we can posit the following for the two households:

- First household typically doesn't routinely use air conditioners in summer(December), or heaters in winter months (May and June). They are strictly 9 to 5 professionals, hence their energy consumption during day time is pretty consistent and only peaks up during a few hours in the evening. They do all energy-intensive chores (e.g. laundry, vacuum cleaning) mostly on the weekends. Also, part of the forth weekend of the month is spent out of the house visiting friends/family or going out for leisure trips.
- Second household typically use air conditioners in summer months and heaters in winter months. This might be a family with children with stay-at-home parents/ work-from-home professionals with extremely varying energy consumption peaking up from noon. Their weekday behavior is not distinctly different from weekends. Also, they are more likely to distribute their energy-intensive chores throughout the week and not leave all them for the weekends. Sundays are still pretty heavy on energy for this

Violin plot across hour-of-day and day-of-week for Household 1

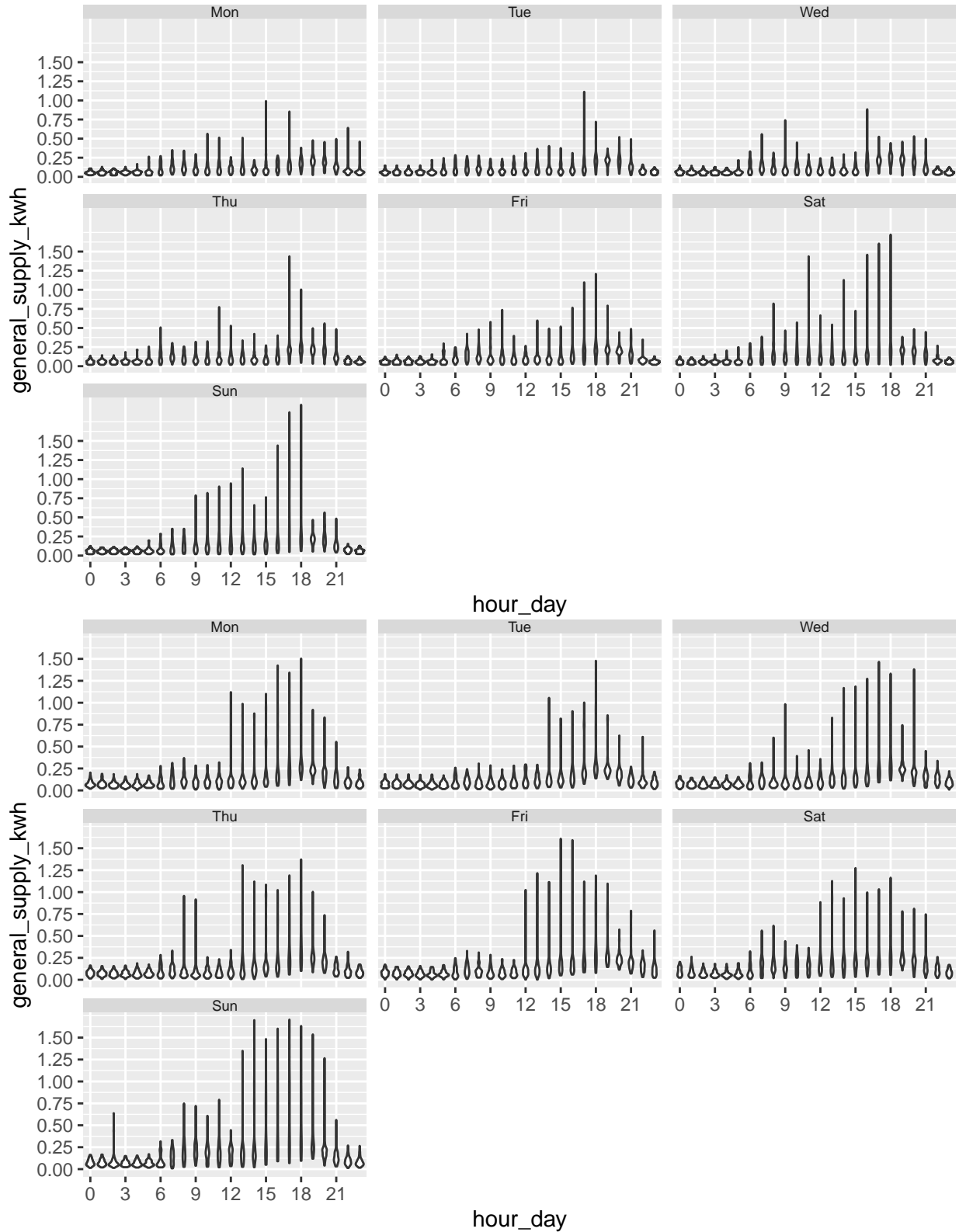


Figure 2: Violin plots of hourly consumption across different day of the week for the two households A long-tailed violin for an hour would imply that the energy distribution is skewed to either right or left, whereas a bulge shaped distribution implies that the energy behavior is more predictable and energy consumption in those hours can only vary in that range where the bulge appears. If an hour has a distribution that seems like a straight line, they are extremely volatile. For household 1, hours between 17 to 19 hours (weekdays) seem to be volatile compared to other hours. On weekends, energy consumption is volatile even in morning hours. For the second one, hours before 10am and beyond 10pm have a bulge implying those hours are pretty consistent. Hours in between these are variable.

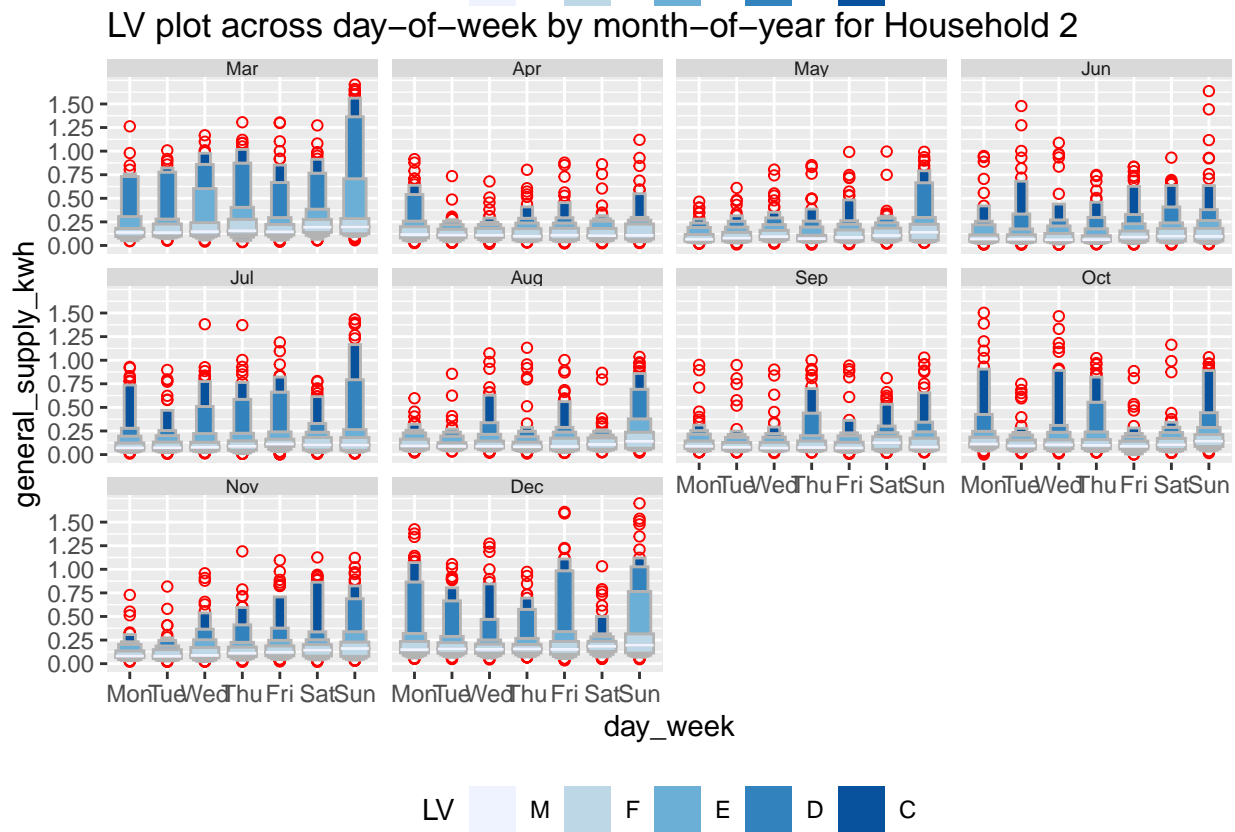
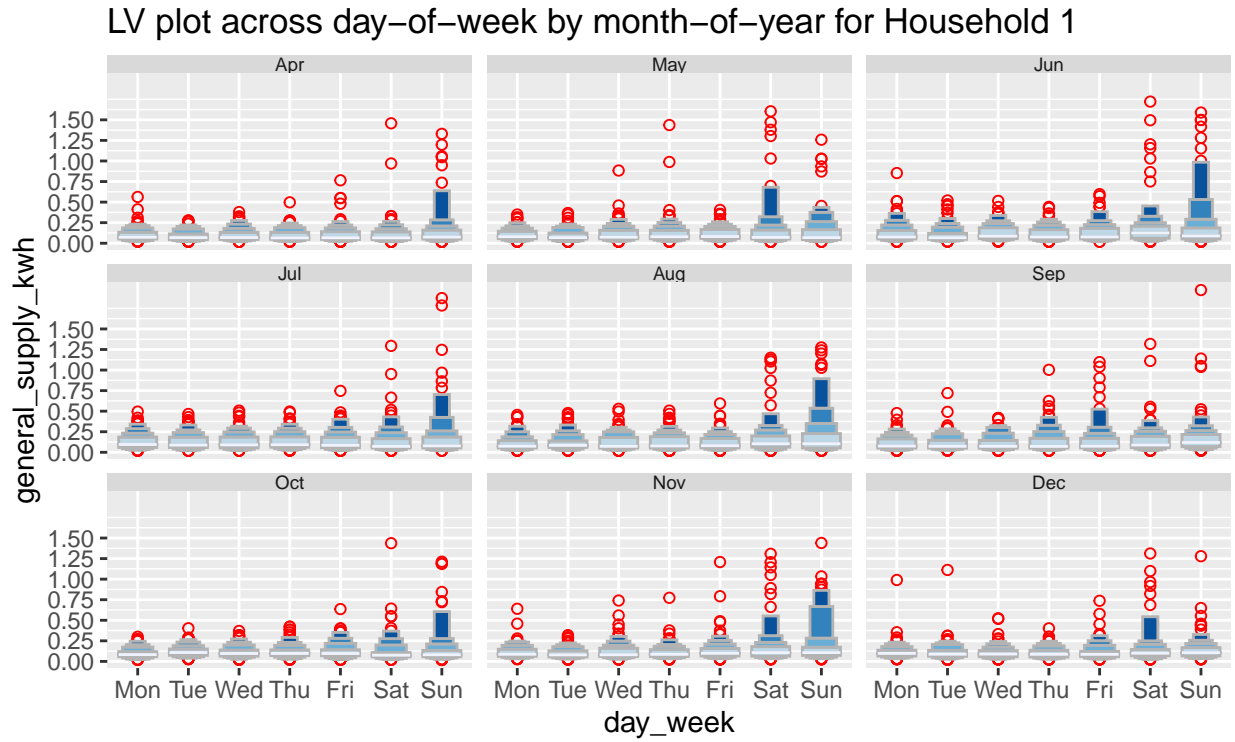


Figure 3: LV plots across days of the week and months. They convey detailed information about tails where outliers (colored in red) are unexpected observations rather than extreme ones. M, F, E, D and C represents 50, 25, 12.5, 6.25 and 3.13 percent of the tail area respectively. Household 1's tail behaviors are not distinctly visible because of the consistency in behavior on weekdays. For the second household, the F letter value seem to range between 0.75 to 1.25 Kwh on weekends and 0.25 to 0.75 Kwh on weekdays in Mar, which is very different from Dec. This reinforces the fact that their behavior changes across months but not so much between days-of-week.

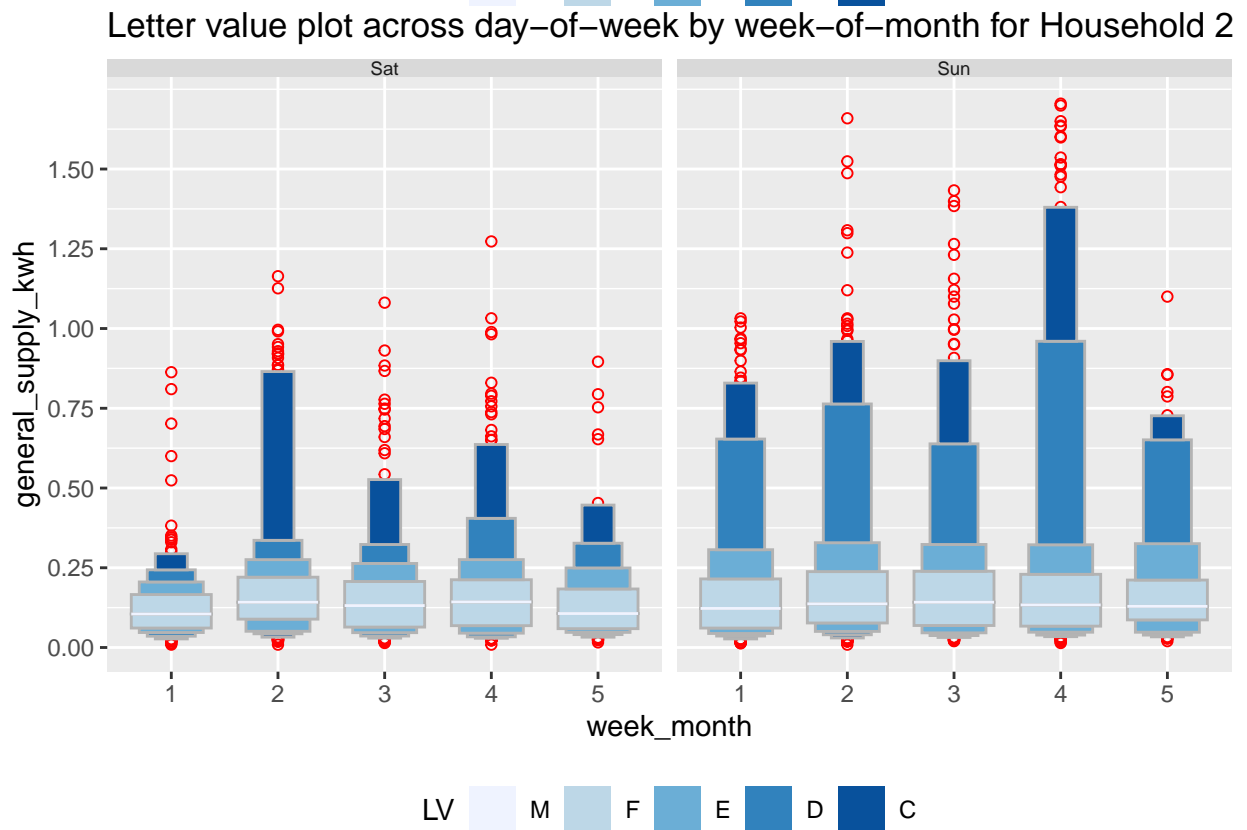
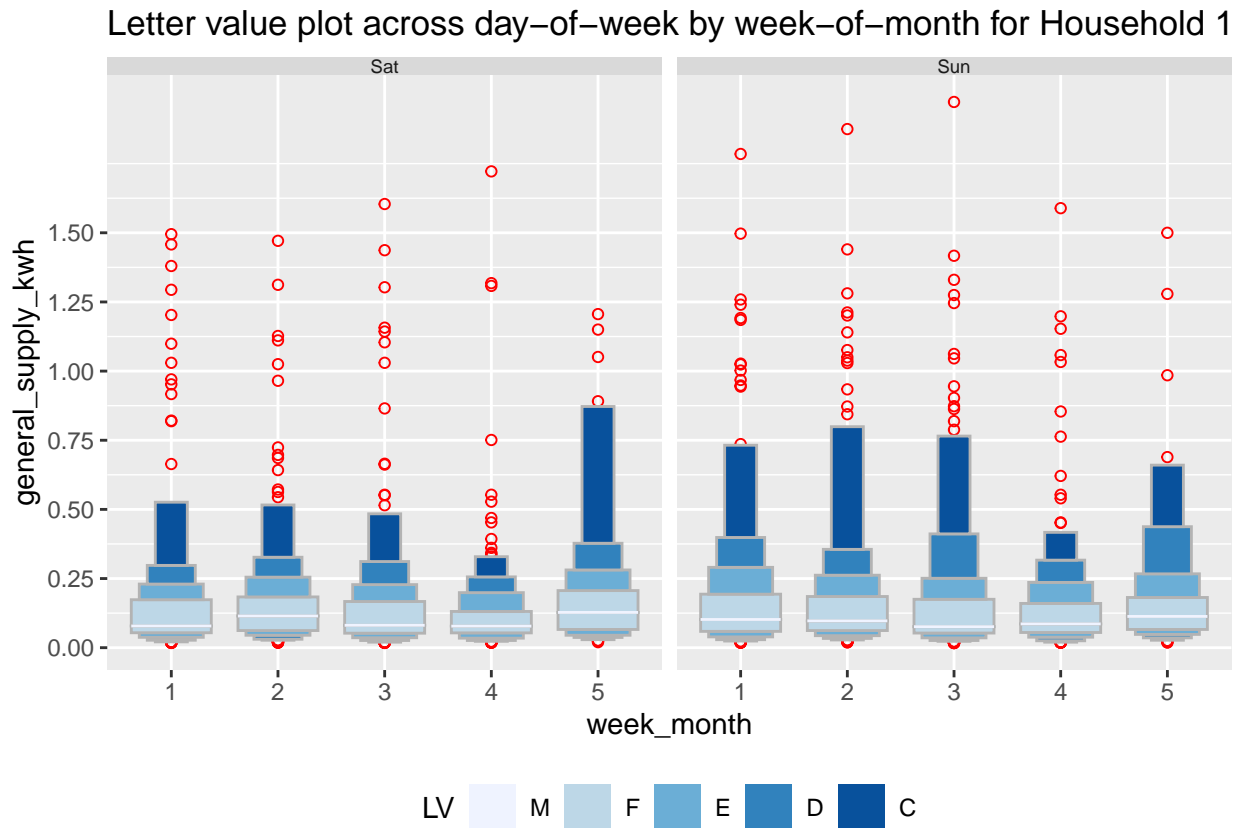


Figure 4: LV plot for Saturday and Sunday across weeks-of-month. For household 1, the 4th week of the month sees low energy consumption for both days. Typically Sundays are heavier on energy than Saturdays. For household 2, the fourth Sunday has high energy consumption and long right tails implying more variable behavior. The first Saturday of the week seems quiet and consistent for this household.

household when all members stay at home. Also, they sometimes spend the last Sunday of the month doing “monthly” chores or have people come over.