

# Case study: Analysis of smart meter data

*Sayani Gupta*

*25/08/2019*

## Case study: Analysis on smart meter data

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. 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 and 2 shows the distribution of hourly consumption of household 1 and 2 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 3 shows the distribution of hourly energy usage across different days of the week through violin plots for household 1. 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

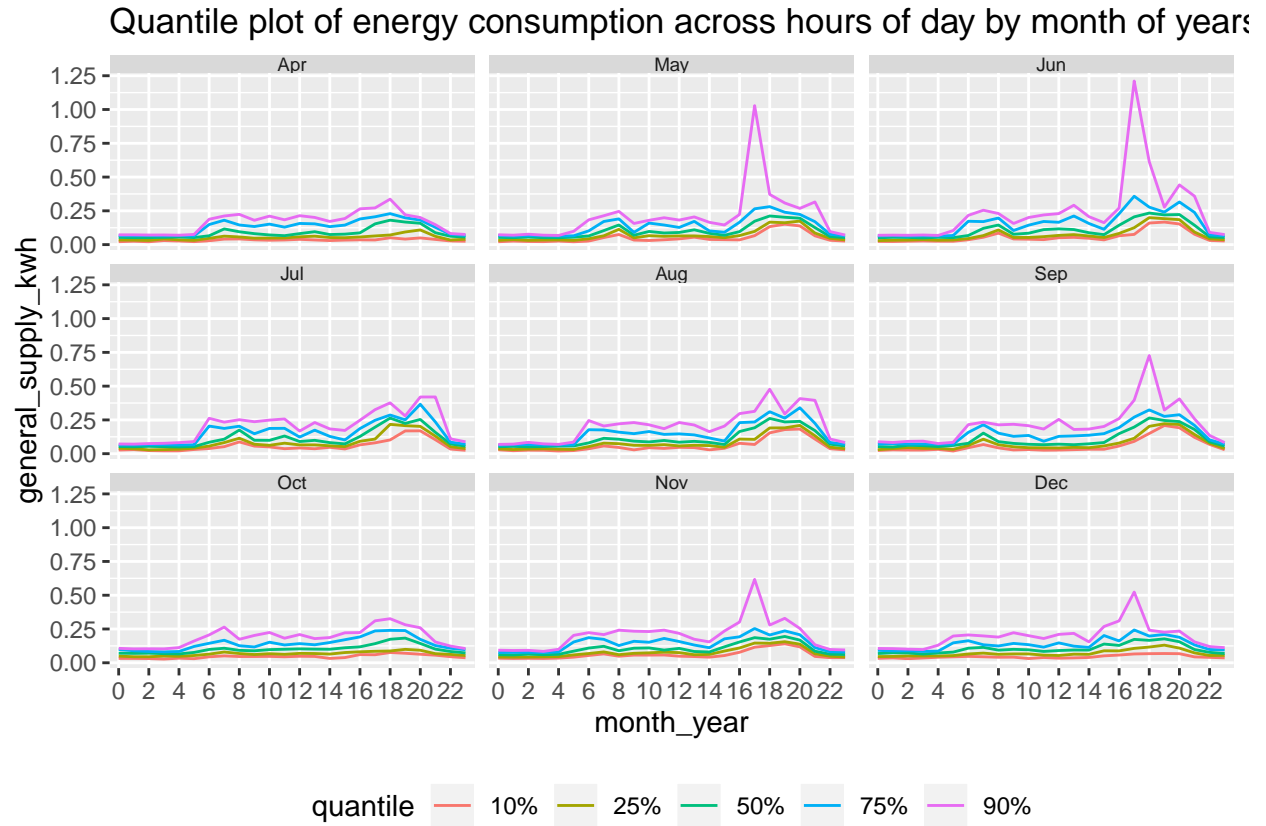


Figure 1: Quantile plots of hourly consumption across different months for the first household. 10th, 25th, 50th, 75th and 90th percentile are shown. For example, 90th percentile might denote behavior which are extreme and 50th percentile corresponds to a more typical behavior. Evening hours generally exhibits a peak. Some extreme behavior is exhibited in the evening hours of May and June, 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.

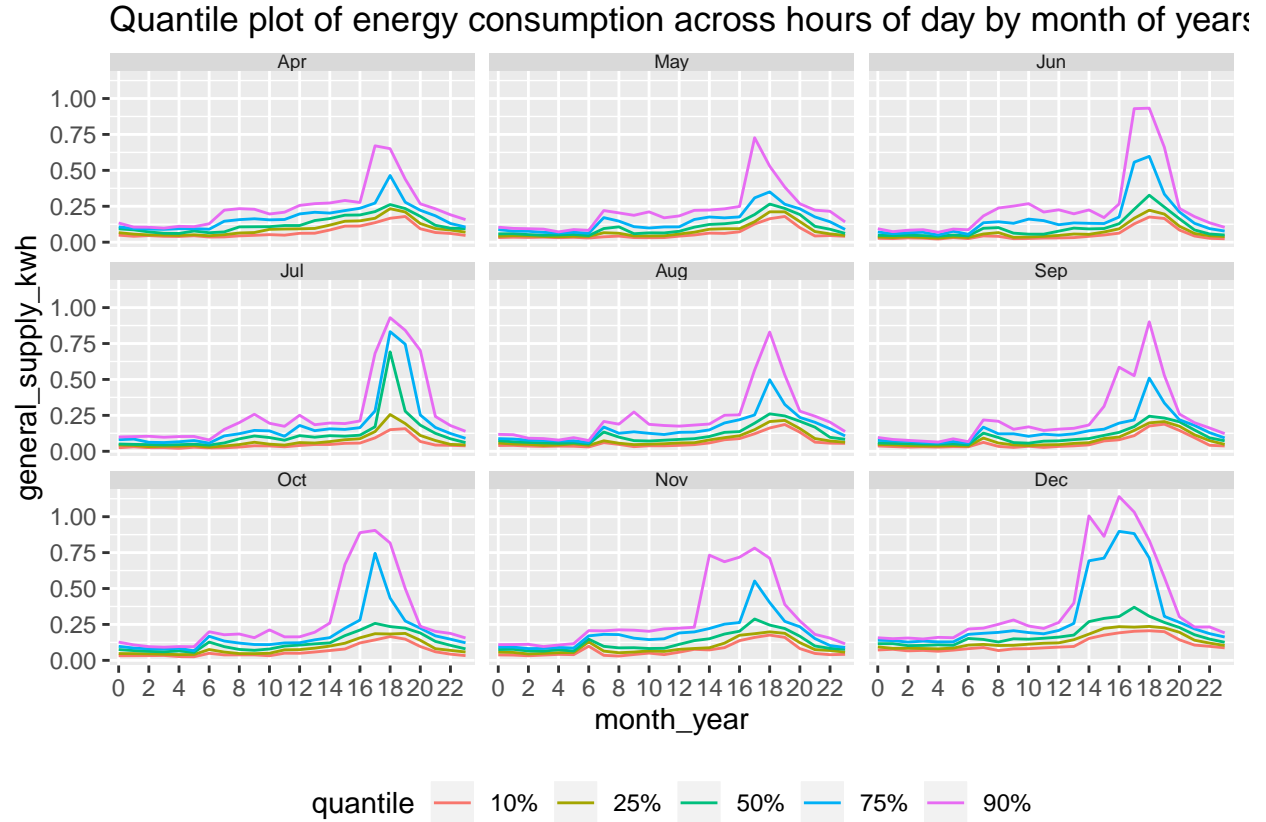


Figure 2: Quantile plots of hourly consumption across different months for the second household. 10th, 25th, 50th, 75th and 90th percentile are shown. Distinct evening peaks for few hours across all months. The median consumption also varies across months. This household varies their behavior with seasons and are likely to be using heaters in colder months and air conditioners in hot months like January. This is evident from the 75th and 90th percentile of consumption.

Violin plot of energy consumption across hour-of-day and day-of-week for

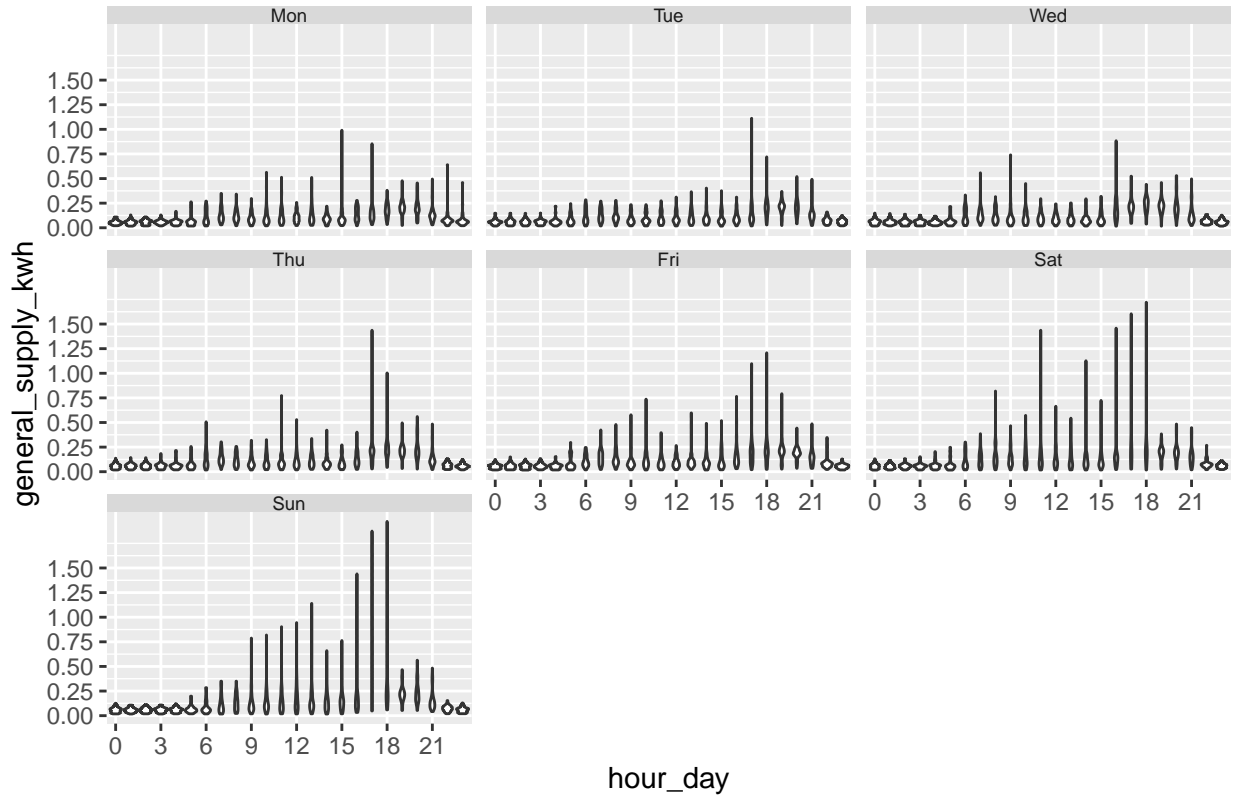


Figure 3: Violin plots of hourly consumption across different day of the week for the first household. A long-tailed violin for an hour would imply that the energy distribution is skewed to either right or left, whereas a bulge shaped distribution implies that the energy behavior is more predictable 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. On weekdays, hours between 17 to 19 hours seems to be extremely volatile compared to other hours. On weekends, energy consumption is volatile even in morning hours.

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.

Figure 4 shows 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 5 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

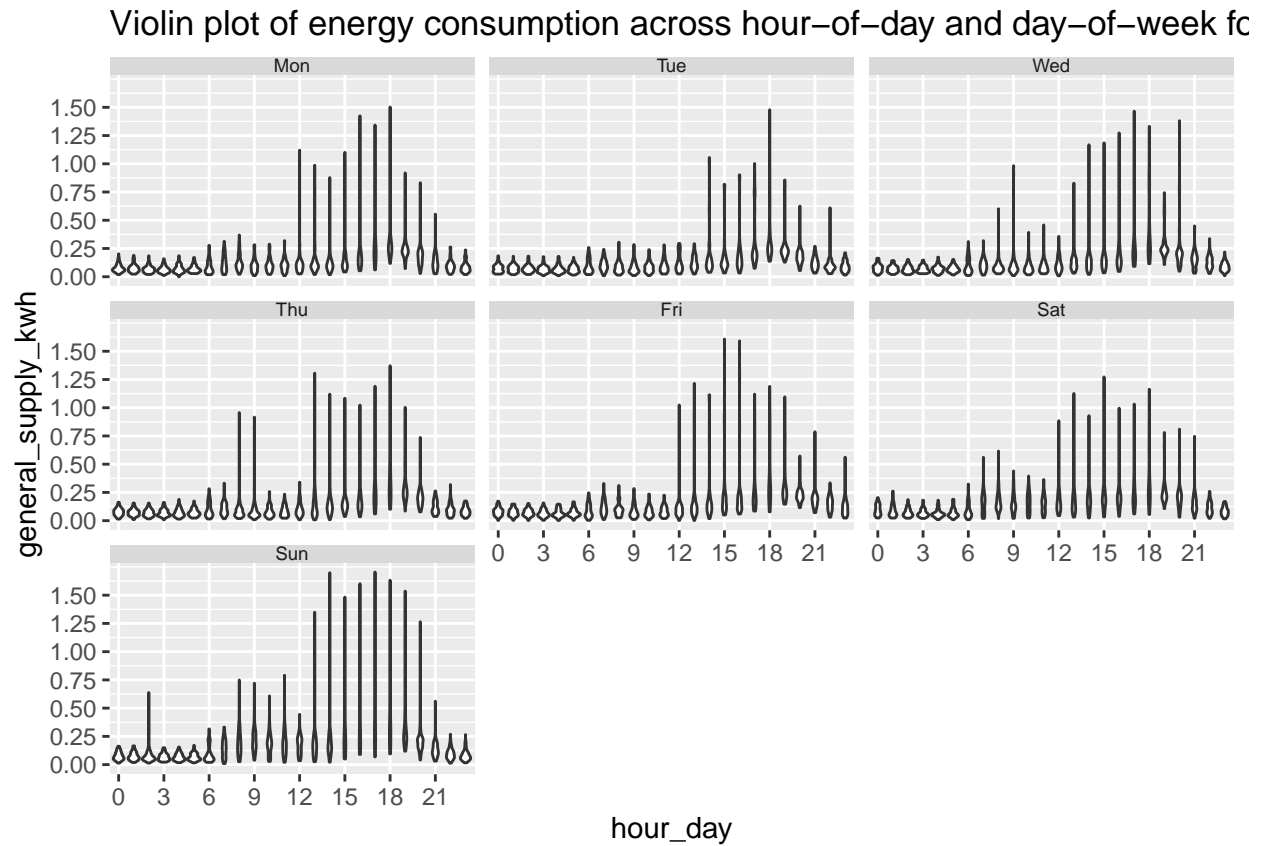


Figure 4: Violin plots of hourly consumption across different day of the week for the second household. Hours before 10am and after 10pm have a bulge implying those hours are pretty consistent. Hours in between these have a long right tail which are even higher for weekends.

weekends, as can be observed from the letter values of energy consumption, are different across months which can be attributed to weather conditions.

Figure 6 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.

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.
- 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.

We can keep on slicing and dicing the data similarly and derive more insights about the behavior of these households.

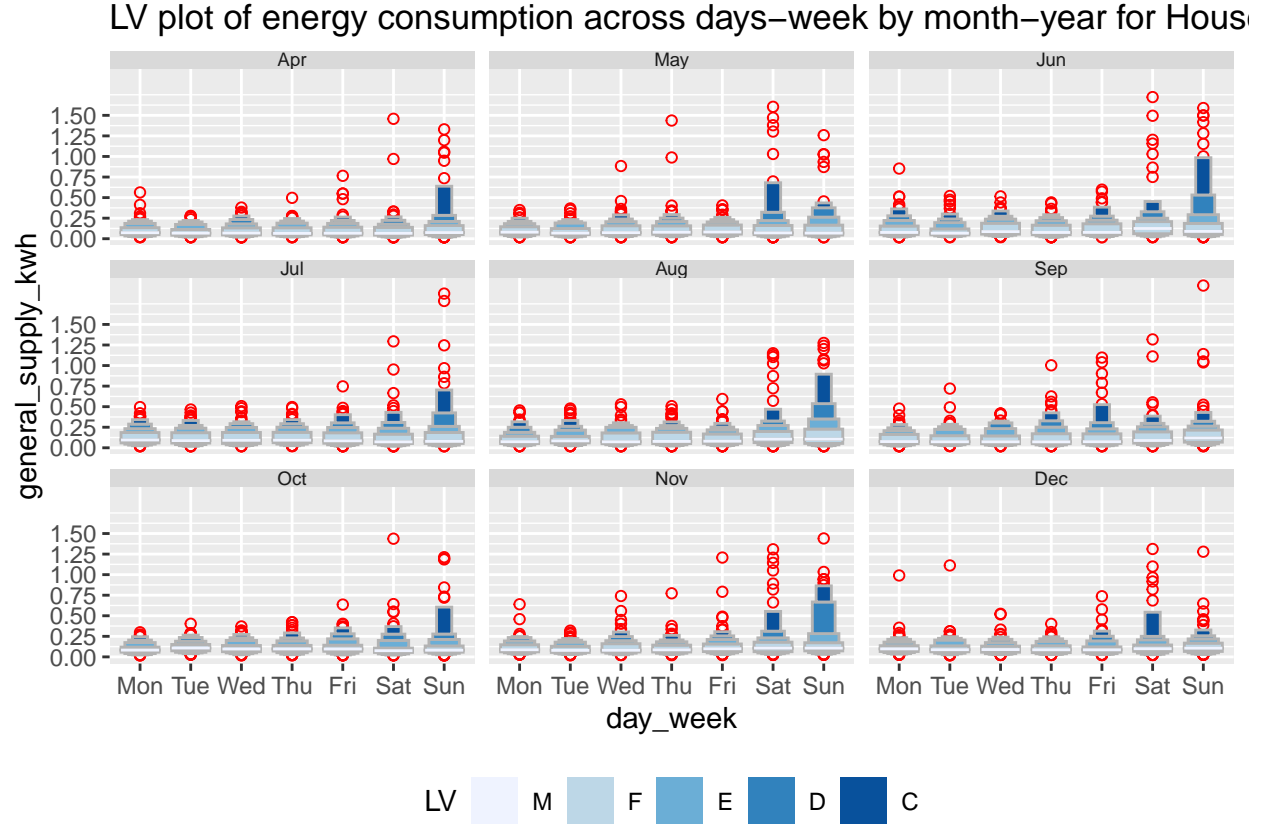


Figure 5: Letter value plots of energy consumption across days of the week and months for the first household. Letter value plots convey detailed information about tails of the distribution and outliers are unexpected observations rather than extreme observations. M, F, E, D and C represents 50, 25, 12.5, 6.25 and 3.13 percent of the tail area respectively. Outliers are colored in red. It can be noted that they do not typically consume more than 0.25 KWh in the weekdays. Their tail behaviors are not so distinctly visible because of the consistency in their behavior.

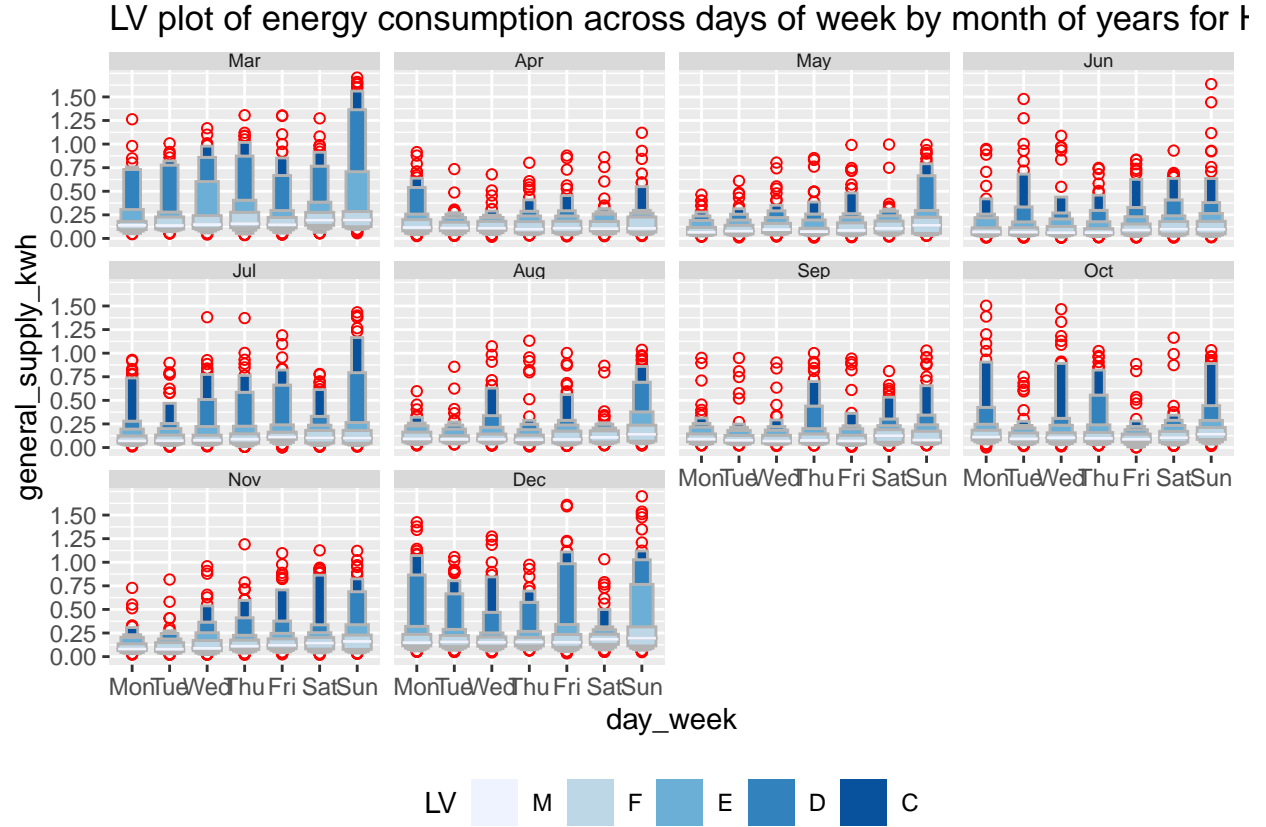


Figure 6: Letter value plots of energy consumption across days of the week and months for the first household. Contrary to the first household, outliers appear at an energy usage beyond 0.75 KwH. The extreme values in the tail of the distribution are represented by different letter values. In March, one can observe the F letter value to range between 0.75 to 1.25 KwH on weekends and between 0.25 to 0.75 KwH on weekdays -represented by letter value D, which is far more higher than subsequent months and comparable to summer month December. For unusual days or more volatile days of the week, one can expect to see more distinct letter values. This reinforces the fact that their behavior changes across months but not so much between days of the week.