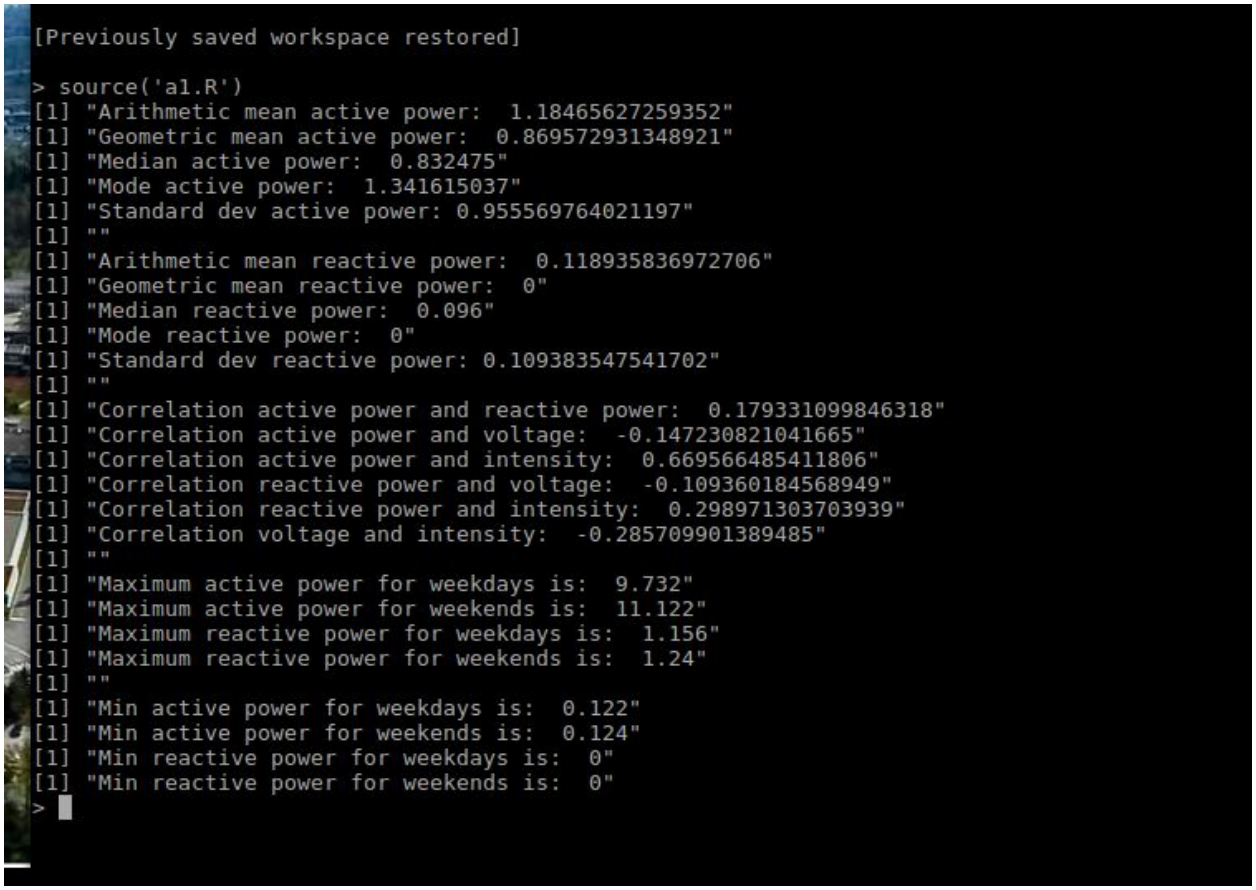
Observed Outputs



Assumptions & Explanations

We look at a power consumption dataset that determines a multivariate time series by describing various features over time. These features are Global Active Power (A), Global Reactive Power (B), Voltage (C), and Global Intensity (D).

We assume that our data and the packages used are accurate for the purpose of our analysis. Two of our team members achieved identical results running their code, while another team member achieved different results using the same function. It could be that the machines / OS were rounding numbers differently. We did not investigate these deviations further.

Since we could not find a straightforward Mode function in R, we used our own function that uses the data's unique property and returns the value with the highest frequency. This idea was used from stackoverflow. [1]

For the correlation function, if our input parameter for "use" was "everything" and the function found a null value ("NA"), then the correlation would be null. The "na.or.complete" or "complete.obs" settings did not have this issue, and we chose "na.or.complete" so we do not miss any details.

We found a positive correlation between active and reactive power, active power and intensity, and reactive power and intensity. This makes sense, because we expect more power "dissipation" (phantom power) as active power (real power) and intensity go up.

On the other hand, both active and reactive power had a negative correlation as voltage went up. We know that power over long distances is transferred in high voltage, and the data seems to reflect that low intensity and high voltage would reduce loss. [2]

We also found that more power consumption was used during the weekend, for both min and max results. Based on this, we think that the data includes business consumption as well as household consumption. Businesses are normally opened during the weekend as well, but people are more likely to stay at home and consume more power during the weekend. This would imply businesses consume consistent power throughout the whole week, while on average individuals consume more power when they can be home all day.

References

[1] https://stackoverflow.com/questions/2602583/geometric-mean-is-there-a-built-in
[2] https://electronics.stackexchange.com/a/152311