

Visualisation and analysis of probability

distributions of variables defined on

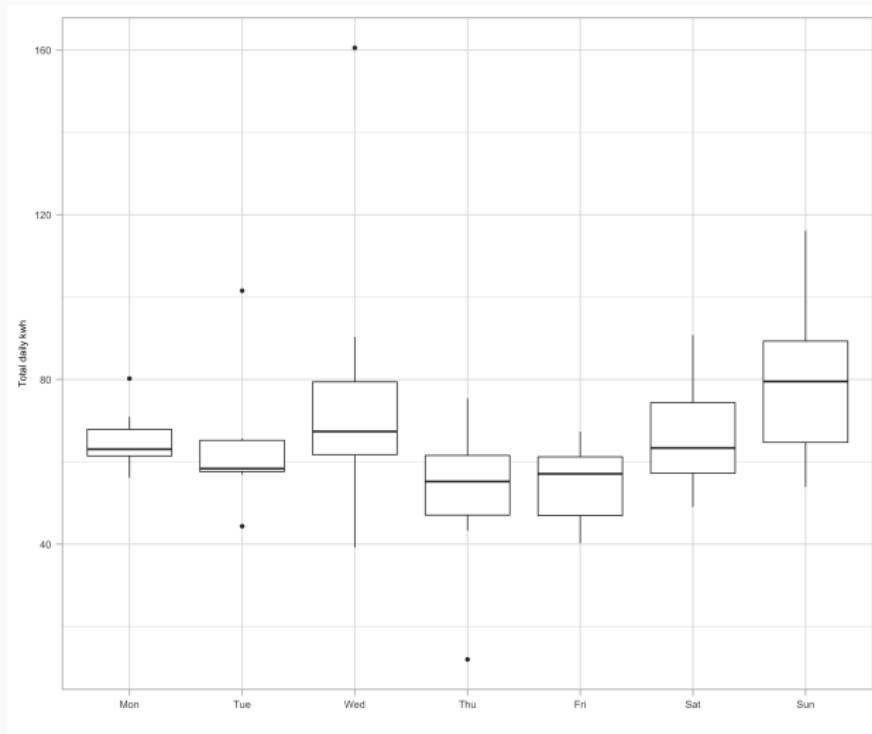
spatio-temporal granularities

Sayani Gupta

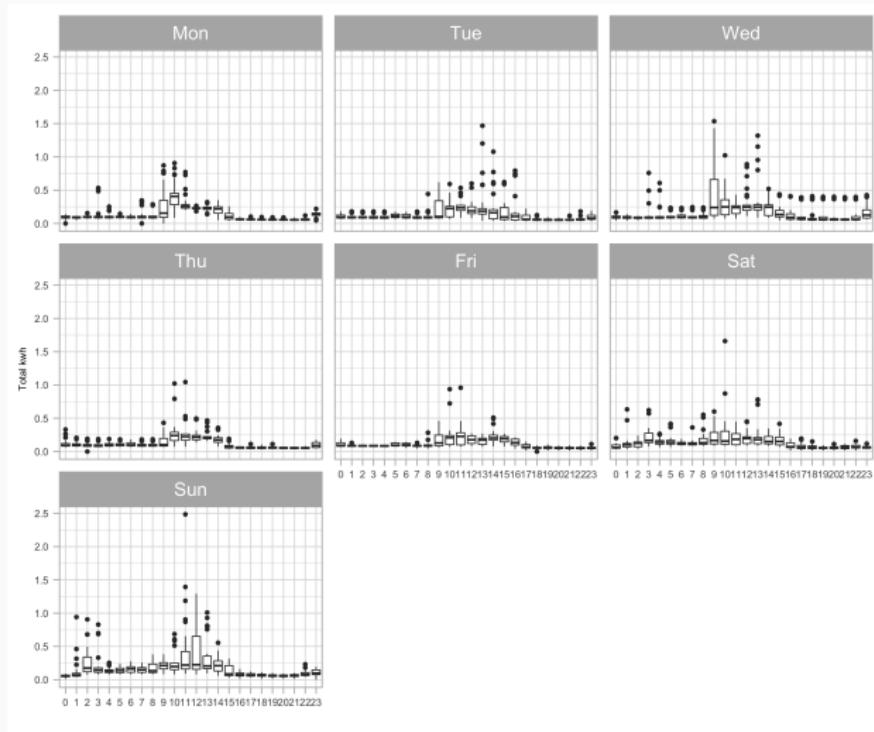
February 22, 2019

**30 minutes interval smart meter
data from Monash Residential
Services from April 4, 2018 to
May 31, 2018**

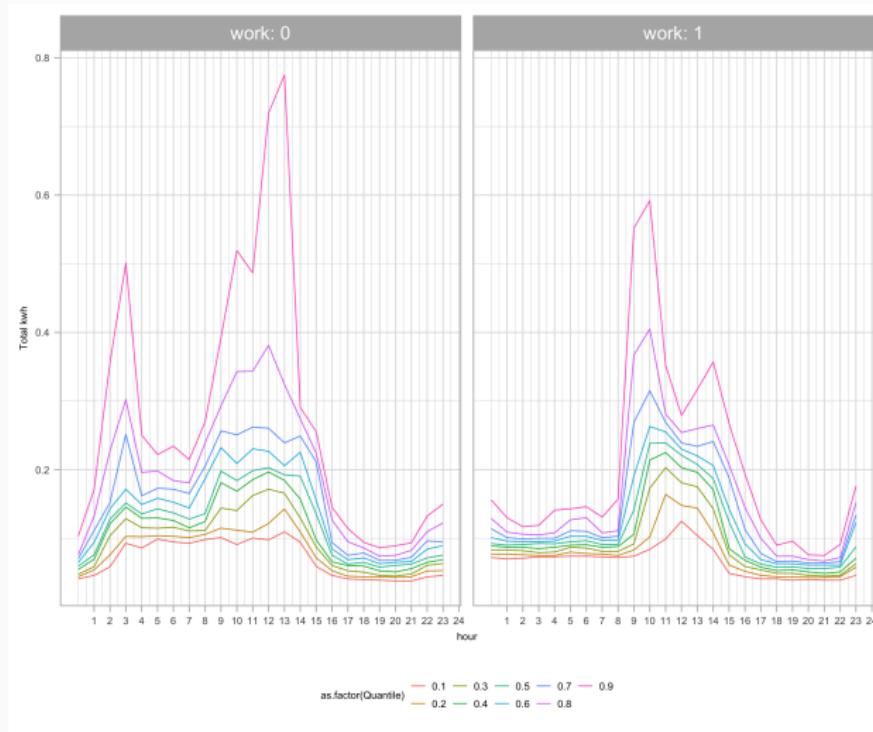
Across days



Across hours, by day of the week

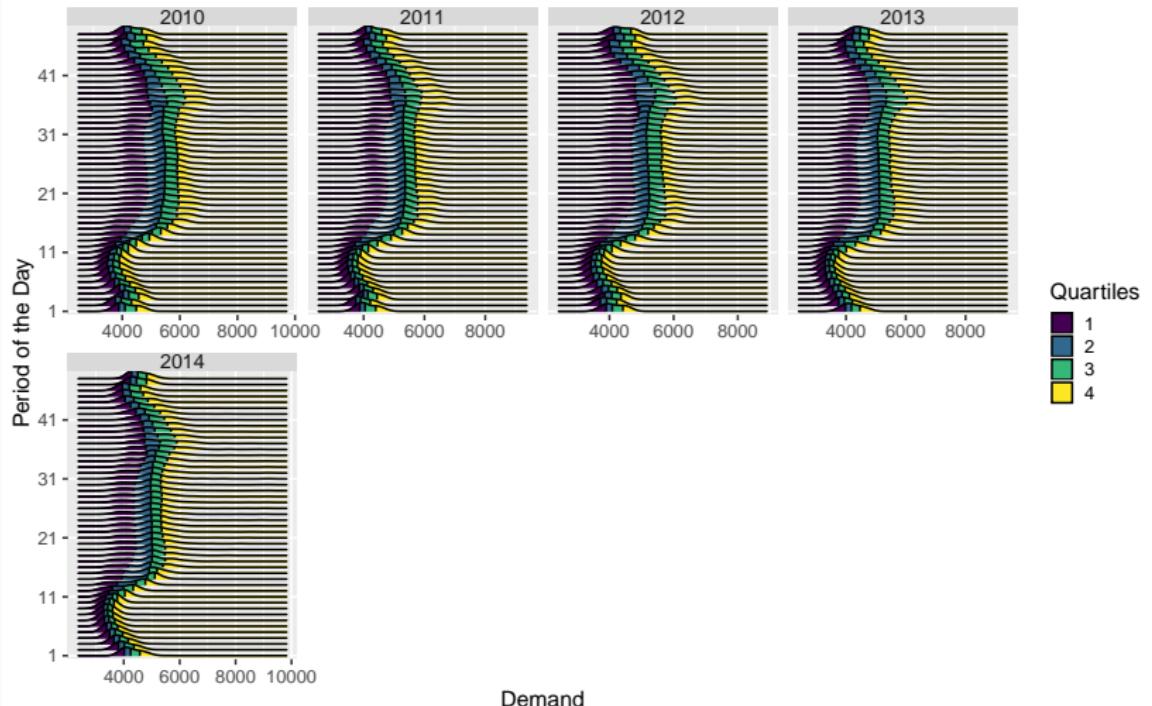


Across hours, by work days and non-work days

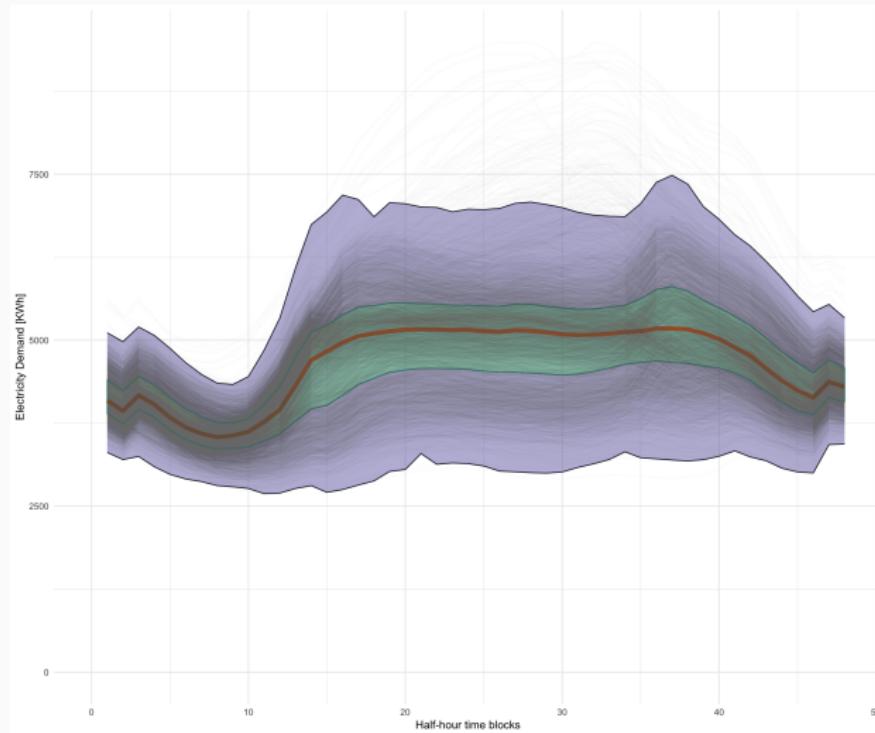


**15 minutes interval state level
smart meter data of Victoria
from January, 2002 to December,
2014**

Across half-hours of the day, by year

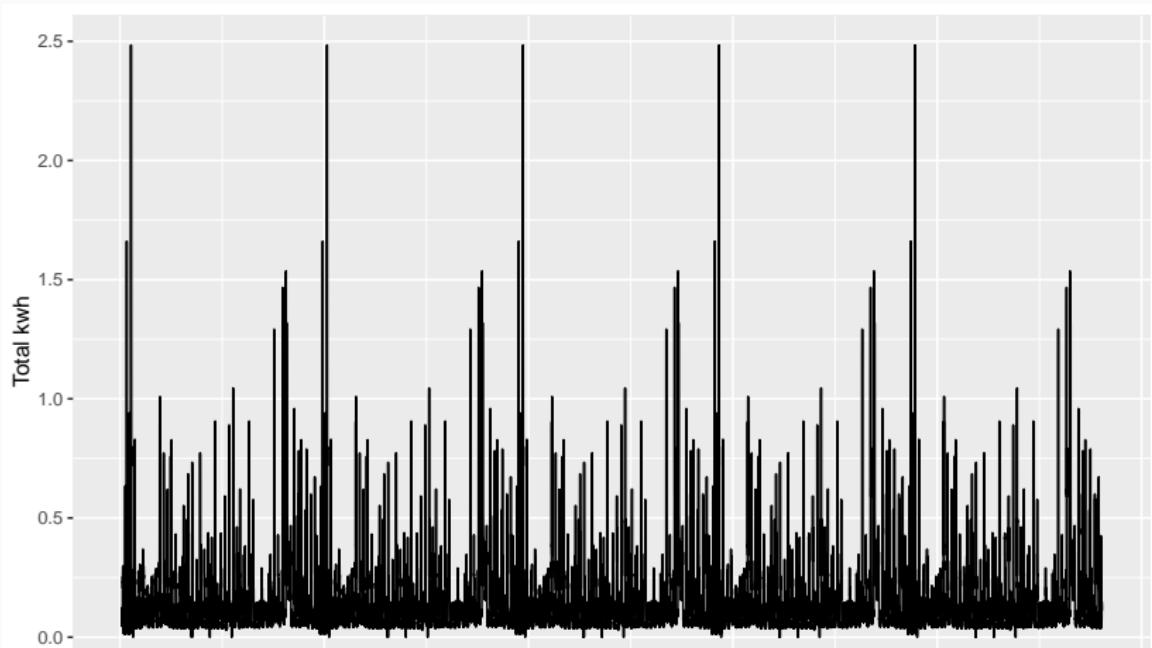


Boxplot statistics superimposed on data



Arrangement of time

- Linear
- Cyclic



Deconstruction of Time - Time Granularities

Data and associated time can be modeled and integrated at different levels of human made abstractions of time called **granularities** (e.g., months, days, hours, and seconds), enabling **display of probability distribution** along granularities

- Aggregating information across granularities can communicate completely different aspects of data

Example of time granularities

With bottom-most granularity 30 minutes and upper-most granularity as year, granularities can be thought of as

Month of the Year: 1

Week of the Month/Year: 2

Day of the Week/ Month/Year: 3

Hour of the Day/Week/Month/Year: 4

30-mins of the Hour/Day/Week/Month/Year: 5

Total number of granularities = $(1+2+3+4+5) = 15$

Questions arising when determining the exhaustive number of plots

- Can we plot all 15c2 combinations or 15c3 combinations if we want to draw two or three granularities at a time?
- Are there combinations which will misguide us for comparison and should not be plotted together?
- When should we say a plot is misguiding us? Are there any defining criterions?
- What are the attributes of the time granularities which makes a plot misleading?

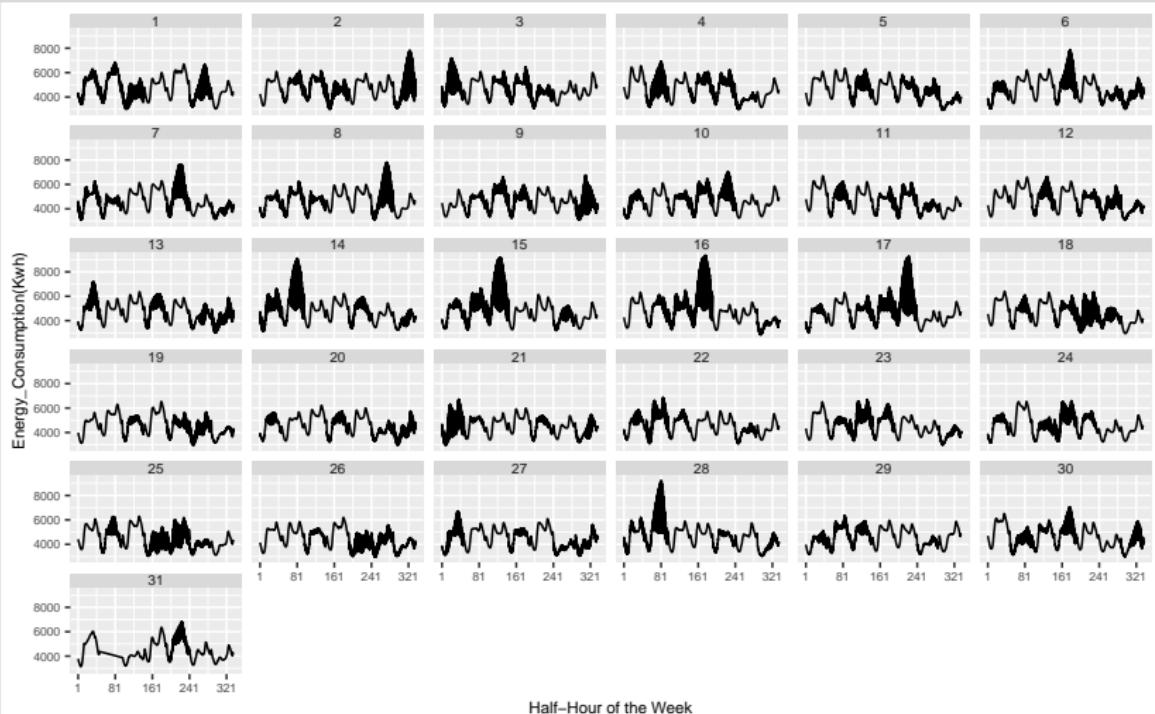
Criterion for misguidedness

Ranges of time in subsets have different sample size or too small sample size to construct probability distribution

- Choosing time granularities from the bottom of the tree amounts to losing degrees of freedom for displaying so many levels of the granularities

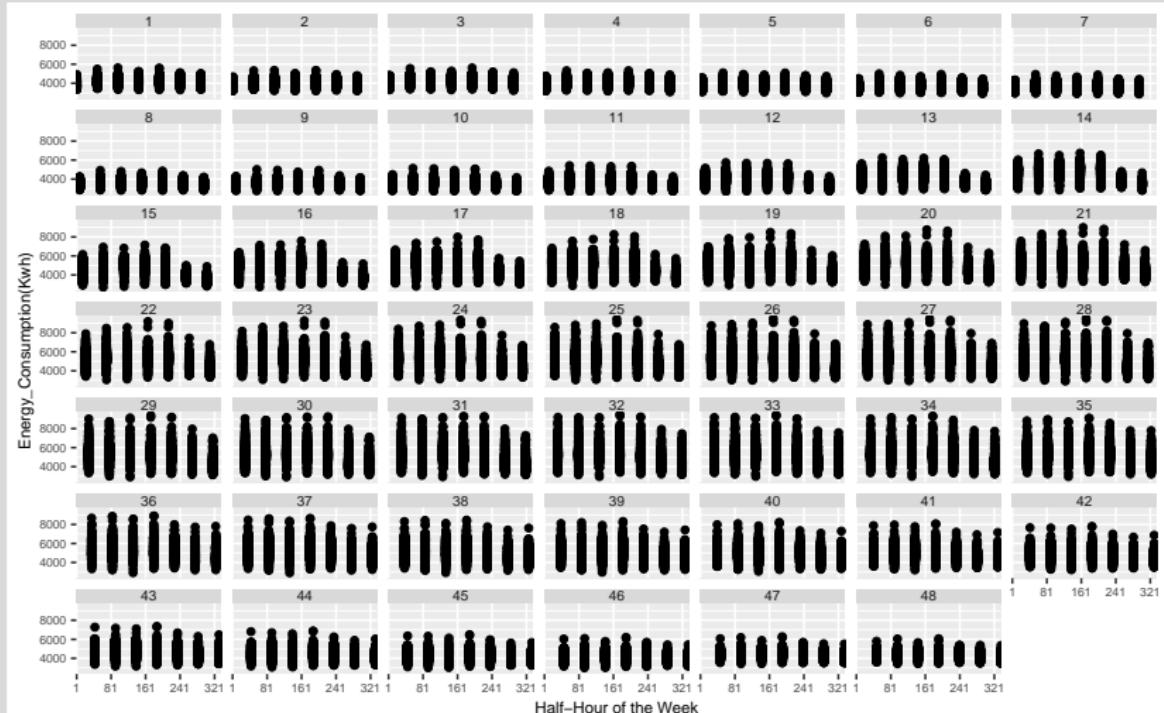
Plotting irregular time granularities

Half hours of the week facetted by days of the month for state level electricity demand of Victoria for 2014



Incomplete design introduced by superfluous gaps

Half-hour of the week facetted by half-hour of the day for electricity demand of Victoria from 2002 to 2014



Way Forward

- How to define criteria for plots that work vs plots that don't?
- How to extend the ideas to allow for spatial as well as time dimensions?
- Are there other visual representations of probability distributions that would be suitable?