Exploring probability distributions for bivariate temporal granularities

Sayani Gupta – Department of Econometrics and Business Statistics, Monash University

Smart meters measure energy usage at fine temporal scales, and are now installed in many households around the world. We propose some new tools to explore this type of data, which deconstruct time in many different ways. There are several classes of time deconstructions including linear time granularities, circular time granularities and aperiodic calendar categorizations. Linear time granularities respect the linear progression of time such as hours, days, weeks and months. Circular time granularities accommodate periodicities in time such as hour of the day, and day of the week. Aperiodic calendar categorizations are neither linear nor circular, such as day of the month or public holidays.

Visualizing data across these various granularities helps us to understand periodicities, pattern and anomalies in the data. Because of the large volume of data available, using displays of probability distributions conditional on one or more granularities is a potentially useful approach. This work provides tools for creating granularities and exploring the associated within the tidy workflow, so that probability distributions can be examined using the range of graphics available in the ggplot2 package. In particular, this work provides the following tools.

- Functions to create multiple-order-up time granularities. For example, week of the month, hour of the week, or quarter hour of the day. This is an extension to the lubridate package, which allows for the creation of some calendar categorizations, usually one-order-up. [Needs defining]

- Checks on the feasibility of creating plots or drawing inferences from two granularities together. Pairs of granularities can be categorized as either a *harmony* or *clash*, where harmonies are pairs of granularities that aid exploratory data analysis, and clashes are pairs that are incompatible with each other for exploratory analysis.