Paper Analysis- 3

**Cluster and Calendar based Visualization of Time Series Data**

**-Paul Sampson Ledala**

The combined representation of clusters and daily patterns using an integrated calendar and graph is such a clever visualization. It breaks the need for multiple visualizations to show different patterns within a day, over a week and holistically across a year. It also eliminates the need for assuming that there is a fixed relation between distribution of patterns over the months and weekdays. Within one visualization all the patterns over a day, week, month, and year are made clear by using the proposed solution on univariate time series data presented. For me, this brings a novel way of thinking about visualizations- where you can display so much information using just a couple of integrated displays.

This visualization is unique and simple enough so that there is no need for alternatives! I don’t see how this can be improved apart from making the right decisions on how to cluster the data. The closest alternative I could get to finding how the patterns can be seen over different periods of time on univariate time-series data is an integration of a ‘Rectangular View’ and ‘Parallel Coordinate View’ of clusters. This idea is introduced in “Methods for the Visualization of Clustered Climate Data-Thomas Nocke, Heidrun Schumann and Uwe Bohm” in Figure 8, page 15. Although the naming is different, in spirit these two visualizations are the same. They provide 2 views, one of which show clustered data on a rectangular (calendar) view and another as a graph (parallel coordinate view).

Representing univariate time series data in multiple views is also effective as the user can go directly to the view that he wishes to find patterns in (may it be daily, weekly, monthly, seasonally, or yearly and beyond that). If he wants to find patterns in the data- then clustering would give a nice solution along with dendrogram representation. It is only when we need all 3 views that we could opt for the cluster and calendar-based visualization. But at the same time, this has wide varieties of application in the energy consumption industry, meteorological departments, sales, marketing, trade and even in internal employee analysis too- as described by Figure 4 in the paper. Anything that generates data over a year!

Although this technique works best with univariate time series data, I can integrate the essence of it into my Baltimore Crime data analysis project. In my visualization (Nightingale Rose Chart), I’ve averaged the data in the dataset (decades of data) to show me crimes across the hours of the day, days of the week and months of the year. I had the issue of toggling between these 3 views. If all 3 views were to be shown in a single visualization, I could use the calendar and cluster visualization to make that possible.

Finally, the selected paper differs from the other paper (Parallel Sets: Visual Analysis of Categorical Data 3B) by applying to different data. While cluster and calendar-based visualization is on time series data. Parallel sets are used for visual analysis of categorical data.