Review on Visualizing probability distributions across bivariate cyclic temporal granularities

Comments to the Author Short Summary: The paper introduces a calendar algebra approach to specify classes of deconstructing time into granularities and relate pairs of granularities to each other. The definitions and computations allow to construct linear and cyclic time granularities. For these granularity pairs, the paper introduces ways to visualise the descriptive statistics in different plots and illustrates with two example datasets how the visualisation is constructed, explored, interpreted and analysed. The authors also provide an implementation of the introduced computations and visualisations in an R package.

In the following I will give an overview on the major strengths and weaknesses, as well as a more detailed discussion on some shortcomings and give pointers to possible ways to improve the paper.

Strengths: - Introduces a more generalised definition and characterisation of granularities to allow single- and multiple-order-up combinations - allows granularities for different time structures, like linear and different periodic structures, like cyclic (circular, quasi-circular, and aperiodic). - Defines relativities for pairs of granularities, the necessary calendar algebra to do the computation, and relations between granularities. - Provides an R package for applying and computing the proposed ideas and visualise the results in order to analyse the data (missing only the computation of cyclic aperiodic granularities) - Illustrated the applicability and usefulness using two example applications and datasets.

Weaknesses: - Cricket sport argumentation a bit difficult to follow for people not knowing the sport at all. - Embedding the visualisation into the grammar of graphics is not convincing or may not be necessary at all. Either extend or remove. - Not clear which visualisations for descriptive statistics known from literature are implemented and supported. - Letter-value plots are widely used in the paper, but not explained in the necessary detail to understand the visualisation adequately.

Overall the paper is very well written, all necessary details provided. The paper introduces an interesting and challenging problem of formalising more general hierarchical structures, like granularities, for time and introduce the necessary computations to allow the application of these concepts to different data. The visualisation and provided discussion allows to comprehend arguments provided and how it turns out in practical applications. The provided R package is useful for applying granularities to ones own data and use for visually analyse the data on different combinations of granularity levels. Below there are some more detailed remarks to improve the identified weaknesses in the paper.

1. *I am not convinced that it is required to introduce the concept of grammar of graphics and embed the used visualization in this framework on such a level of detail like in Section 5. Of course some elements of the grammar of graphics are used in Section 5.2 to argue about facets, levels, aesthetic variables, and mappings, but if the reader does not know about this framework, the introduction and explanations about this concept/framework provided in Section 5 is not enough to follow the later argumentation. I would like to see some more details on the concept/framework on the grammar of graphics, or not using the concepts in the parts of Section 5.2 and paraphrasing it without using the terms from the grammar of graphics. I think that it is doable to explain the construction of the visualization without using this concept/framework, because the mappings are quite straight forward. I would like to see (a) either extend the introduction to grammar of graphics and the basics needed to argue in more detail about the embedding of your visualisations using this concept/framework, or (b) remove it.*

2. *It is not clear from the paragraph in Section 5.1 what basic plot choices are implemented and used and supported for visualisation. There is only the list of plots to display descriptive statistics provided. It would be important to describe which are implemented and possible to use. The later examples use*

primarily letter-value plots, then also box-plots, violin plots and line charts with area quantile plots. The letter-value plots are used in major examples of the paper, therefore I would argue to provide some more details on this visual representation and what the 5 colours and letters mean (e.g. LV=letter value, M = median, F = fourths, E = eighths, D = sixteenths, C = thirty-seconds), as it is not as widely known like standard box-plots. This would help greatly in better following the argumentation and explanations.

3. *I have to confess that I do not know anything about the sport cricket and it may be unthinkable for people from countries where cricket is a national sport, but I really had difficulties to follow the example in Section 6.2. The basic hierarchical concept of over—inning–match–season was understandable, but I think that some parts need to be revisited and clarified for also reaching readers not knowledgeable in cricket. Especially the discussion on the interesting questions raised are a bit difficult to follow without knowing much about cricket. For example some terms should be at least briefly explained or removed if not necessary for the explanation, e.g., Twenty20 format, fielding, bowling, dismissal, wicket. I do know a lot about different sports and also the basic idea and structure in many sports, but cricket was never on my radar. Apologise my ignorance about cricket.*