Report on: **A new tidy data structure to support exploration and modeling of temporal data**

This is a very nice article which addresses an important problem in statistical computing with temporal data. The authors are to be congratulated for their thoughtful and considered approach to the problem of designing a data class for representing temporal data in a tidy way and for providing a publicly available implementation through the tsibble R package. I enjoyed the style with which it is written, a pleasing balance of formal academic writing and blog-style colloquial.

Minor comments

P5L37 “Temporal data tools need to catch up.” This is a little casual. Could perhaps be rephrased as “This paper and the associated **tsibble** R package extends the tidy way of thinking to temporal data.”

P7L44 “A variable number of time representations can be spotted in the wild.” This is a little casual, and I’m not fond of “variable number”. Perhaps “Many different representations of temporal data can be spotted in the wild.” (The “spotted in the wild” is still a little informal – but I like it.)

P7L50 “This diversity and time zone…” suggest rephrase “The tsibble’s index column leverages the date-time data type in R to accommodate diverse temporal representations and associated time zone information.”

P8L16 “The identifying variables from which the “key” is constituted remain the same as in the original table with no further tweaks.” It’s not totally clear to me what is meant here. Is it saying that like with grouping columns (i.e. columns supplied to group_by()), key columns cannot be altered in subsequent analyses? It might be useful to make the comparison to group_by() explicit as it also allows multiple fields. The distinction between group_by() and key() is clear in that the collection of columns that constitute a key must uniquely identify observations (obviously this isn’t the case for grouping variables).

P8L55 “fit models that expect **regular** time series.” Should regular be added to the sentence.

P10L19 “Duplicates signal ~~the~~ **a** data quality…”

P10L21 “gaze at data” I initially thought this was odd phrasing, but checking the definition of gaze “**look steadily and intently**, **especially in** admiration, surprise, or **thought**” I guess it’s fine.

P12-13 Consider referencing “fable” and possibly “mable” too.

P14-15 I had to read section 5.2 a few times before I understood it… I think what threw me was the last paragraph starts with a reference to purrr, but the functions described in the

remainder of the paragraph describes functions in the tsibble package. Perhaps move the purrr sentence up to the previous paragraph when functional programming is first mentioned.

P15L46 "…table verb *achieves in collapsing* a tsibble…" perhaps "…table verb collapse a tsibble…"

P15L53 "On the other hand, these tsibble verbs can incorporate many third-party vector functions to *step out of the current tsibble zone*" perhaps "The extensible design of the tsibble package means that the table verbs can incorporate many third-party vector functions."

Fig 5 y-axis doesn't need the overall label "Delayed" or if it stays, should be renamed "Status". Similarly for the legend label.

Fig 6 Consider replacing gather() with pivot_longer() which is intended to be easier to use particularly for more complex "gathering" operations.
https://tidyr.tidyverse.org/dev/articles/pivot.html

Fig 7 y-axis label should be "Proportion of flights delayed" or similar. Legend label "Originating airport" (but legend can also be removed, because it is covered by the facet labels.

Fig 9 y-label "Departure delay (mins)". X-axis could be converted to time… e.g. ggplot2::scale_x_time() Could introduce some alpha transparency to the geom_line(). Caption text currently states *"Passengers are apt to hold up around 18 during a day, and are recommended to travel early. The variations increase substantially as the upper tails."* Both sentences need to be reworded.

P22L3 "might be joined" should it be "can be joined" or perhaps a bit more detail, "Aggregating the half-hourly electricity data to the same daily time interval as the temperature data allows us to join the two data tables to explore…"

P22L27 Suggest a reference to Wang, E, D Cook & RJ Hyndman (2018) to explicitly acknowledge that the frame_calendar() function is part of a different package (sugrrants).

P22L37 "degrees C" to "degrees Celsius" or °C

Fig 12 remove legend label "hot"

P25L19 It's not "A **missing** piece of the tsibble package" because later in the paragraph it says that custom calendars are easily embedded in the tsibble framework. I suggest rephrasing the first sentence in a more positive light… "The tsibble package can easily be extended to accommodate user-defined calendars." Suggest also referencing at least one package that provides a user-defined calendar (perhaps a market-specific business calendar).

P26L54 there’s an errant *Â*

P28L22 *Journal of Statistical Software~~, Articles~~*