Design Principles, Comparisons and Limitations

David Hugh-Jones

2020-09-18

This document briefly describes the design of huxtable, and compares it with other R packages for creating tables. A current version is on the web in HTML or PDF formats.

Design principles

I wrote this package because I wanted a simple way to create tables in my LaTeX documents. At the same time, I wanted to be able to output HTML or Markdown for use in RStudio. And, I wanted to be able to edit tables intuitively using standard R features. My typical use case is creating tables of regression outputs, but I also wanted to be able to represent arbitrary data, like a table of descriptive statistics or of plain text.

The idea behind huxtable is to store data in a normal data frame, along with properties that describe how to display the data, at cell, column, row or table level. Operations on the data frame work as normal, and they also affect the display properties. Then, the data can be output in an appropriate format. At the moment, those formats are LaTeX, HTML, Markdown, Word/Excel/Powerpoint, RTF and on-screen pretty printing. More could be added.

Another design choice was to have separate functions per feature. Many existing packages use a single function with a large number of options. For instance, `print.xtable` in the `xtable` package has 34 options, and `texreg` in the `texreg` package has 41. Having one function per feature should make life easier for the end user. It should also lead to clearer code: each function starts with a valid huxtable, changes one thing, and returns a valid huxtable.

The output formats are very different, and decisions have to be made as to what any package will support. My background is more in HTML. This is reflected in some of the huxtable properties, like per-cell borders and padding. The package tries to keep output reasonably similar between LaTeX and HTML, but there are inevitably some differences and limitations. For Markdown and on-screen output, obviously, only a few basic properties are supported.

The package makes no attempt to output beautiful HTML or LaTeX source code. In fact, in the case of LaTeX, it's pretty ugly. The approach is "do what it takes to get the job done".

Comparing Huxtable With Other Packages

When I first wrote this vignette there were many competing packages to create LaTeX and HTML tables. There still are, but in my opinion, the field of sensible modern choices has narrowed down to four:

- huxtable;
- flextable;
- gt;
- kableExtra.

Here's what I think of these:

- `huxtable` has the widest range of outputs, including HTML, LaTeX, RTF, markdown and Word (via flextable). It is very customizable. Its model is drawn from HTML, which can be tricky when outputting LaTeX. It aims to give you fine-grained control over formatting, at cell level wherever possible.
- David Gohel's `flextable` package started off as a way of producing Word and Powerpoint tables. It can also output markdown tables for use in rmarkdown documents. Recently it gained the ability to create PDFs using the pagedown package. Rather than outputting LaTeX, this uses paged HTML. In my view this is a fabulous idea, since LaTeX is a 1980s leftover that needs to die. But, if you need LaTeX, then it won't help. On the other hand, `huxtable` uses `flextable` to produce Word output, so if you only need Word, you might as well go to the source.
- RStudio's `gt` package is powerful and has a big name behind it. As yet it only produces HTML, but PDFs are on the agenda. It aims for more convenience than control, with included functions for many standard formatting options.
- `kableExtra` is a simple set of tweaks for the `knitr::kable()` function. I am not a huge fan of the internal design, but it produces nice output and has a helpful website.

The original table of competing packages is below for historical reference, but it has not been updated.

Table 1: Comparison table, part 1

	huxtable	xtable	formattable	ascii	kableExtra	condformat	htmlTable
HTML output	Y	Y	Y	Y	Y	Y	Y
Background color	Y		Y	Y		Y	Y
Width and height	Y		Y	Y	(Y)		
Text formatting	Y		Y	Y	(Y)		
Borders	Y						Y
Rotation	Y						
Multicolumn/multirow	Y			Y			Y
Alignment	Y			Y			Y
Numeric formatting	Y		Y	Y			Y
Captions	Y	Y		Y			
Format NA	Y			Y			
LaTeX output	Y	Y		(Y)	Y	Y	?
Background color	Y			Y		Y	
Width and height	Y	Y		Y	(Y)		
Text formatting	Y			Y	(Y)		
Borders	Y	(Y)					
Rotation	Y	Y					
Multicolumn/multirow	Y			Y			
Alignment	Y	Y		Y			
Numeric formatting	Y	Y		Y			
Captions	Y	Y		Y			
Format NA	Y			Y			
Long tables		Y					
Labels	Y	Y					
Other features							
Conditional formats	Y		Y			Y	
Autocreate summary statistics							
Dynamic HTML widgets			Y				
Create tables from model objects	Y	Y		Y			
dplyr compatible	Y						
Add footnotes	(Y)				Y		
Other formats							
Markdown	Y						
Screen	Y						
Excel	Y					Y	
Powerpoint	Y						
Word	Y						
DocBook				Y			
RTF	Y						
Notes						Via markdown	

A (Y) means that there is limited support for the feature.

For example, multirow cells may only be supported in headers,[3] or only horizontal

border lines may work.

Table 2: Comparison table, part 2

	Hmisc::latex	ztable	kable	texreg	stargazer	tables	DT	pander	flextable	pixiedust
HTML output	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Background color		Y							Y	Y
Width and height									Y	Y
Text formatting		(Y)		(Y)	Y			Y	Y	Y
Borders		Y							Y	Y
Rotation									Y	Y
Multicolumn/multirow	(Y)	Y			(Y)	(Y)			Y	Y
Alignment	Y	Y	Y		Y	Y		Y	Y	Y
Numeric formatting	Y	Y	Y		Y	Y		Y	Y	Y
Captions	Y	Y	Y				Y	Y	Y	Y
Format NA	Y					Y		Y	Y	Y
LaTeX output	Y	Y	Y	Y	Y	Y		Y		Y
Background color										Y
Width and height										Y
Text formatting		(Y)		(Y)	Y			Y		Y
Borders	(Y)	Y				(Y)				Y
Rotation										Y
Multicolumn/multirow	(Y)	Y			(Y)					Y
Alignment	Y	Y	Y		Y	Y		Y		Y
Numeric formatting	Y	Y	Y		Y	Y		Y		Y
Captions	Y	Y	Y					Y		Y
Format NA	Y					Y		Y		Y
Long tables	Y	Y				Y				Y
Labels	Y	Y			Y	Y				Y
Other features										
Conditional formats									Y	
Autocreate summary statistics						Y				
Dynamic HTML widgets							Y			
Create tables from model objects		Y		Y	Y				Y	
dplyr compatible										
Add footnotes									Y	
Other formats										
Markdown			Y					Y		Y
Screen					Y					Y
Excel										
Powerpoint									Y	
Word									Y	(Y)
DocBook										
RTF										
Notes	HTML via external translator			4				Via Javascript	Via markdown	

A (Y) means that there is limited support for the feature.

For example, multirow cells may only be supported in headers, or only horizontal

border lines may work.