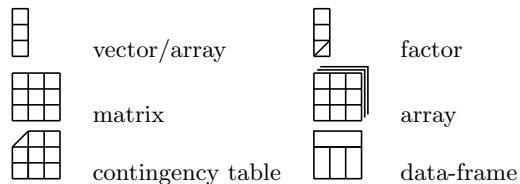


# A Visual Guide to Split-Apply-Combine in R

## Legend

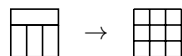
Symbols used in this document:



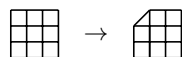
A list is denoted by square brackets:  $[\dots]$ .

## Reshaping

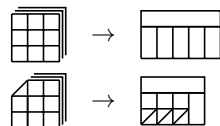
**as.matrix** Coerces a data-frame to a matrix.



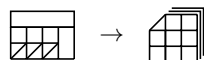
**as.table** Coerces an array to a contingency table.



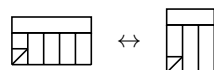
**as.data.frame** Coerces an array or contingency table to a data-frame.



**xtabs** Coerces a data-frame to a contingency table.



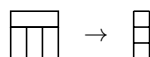
**reshape** Reshapes a data-frame between ‘wide’ and ‘long’ format.



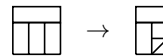
**c** Coerces an array or contingency table to a vector.



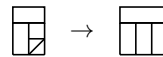
**unlist** Coerces a data-frame to a vector.



**stack** Concatenates the columns of a data-frame into a single column along with a factor. Non-vector columns are dropped.

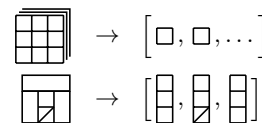


**unstack** Reverses the effects of **stack**.

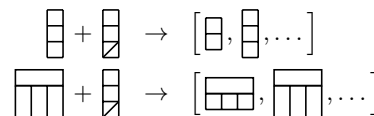


## Splitting

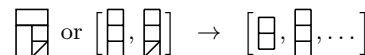
**as.list** Coerces a data-frame to a list of vectors.



**split** Divides a vector or data-frame into the groups defined by a factor.



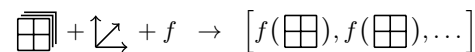
**unstack** Splits a vector or the column of a data-frame into the groups defined by a factor.



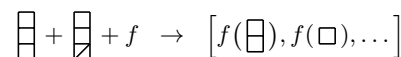
## Applying

The following functions simplify the output to an array, if possible; otherwise they return a list.

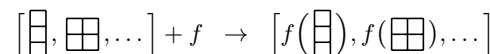
**apply** Applies a function to margins of an array.



**tapply** Applies a function to each cell of a ragged array.



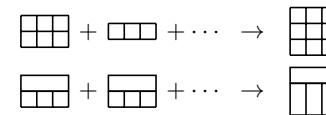
**sapply** Applies a function over a list or vector.



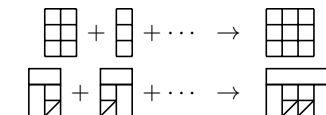
Related functions, not covered here, include **lapply** (does not simplify) and **mapply** (multivariate version).

## Combining

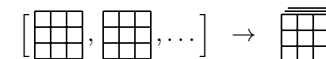
**rbind** Combines a sequence of vectors, arrays or data-frames by rows.



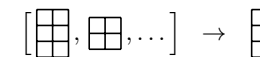
**cbind** Combines a sequence of vectors, arrays or data-frames by columns.



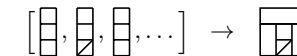
**simplify2array** Combines a list of arrays into a single array by adding more dimensions.



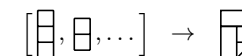
**unlist** Coerces a list of arrays to a vector.



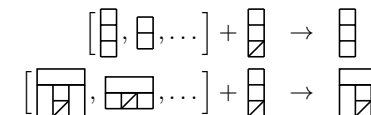
**as.data.frame** Coerces a list of vectors to a data-frame.



**stack** Coerces a sequence of vectors into a single-column data-frame along with a factor.



**unsplit** Combines a sequence of vectors or data-frames into a single vector or data-frame by interleaving rows according to a factor.



**merge** Merges two data-frames by common columns.



Version: 1.0

Copyright © 2016 Ernest Adrogué Calveras.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation.