

A Tale of Two Packages

Tim Hesterberg

Google, Inc.

TimHesterberg@gmail.com

Keywords: `dataframe`, `aggregate`

I'll talk about two packages. The **dataframe** package contains replacements for the data frame creation and subscripting functions in R, that dramatically cut the number of copies of that *R* makes of the data, e.g. from four to one for `as.data.frame(a vector)`, and from 10 to 3 for `data.frame` (a list). Naturally, this is faster. I'll give tips on how to improve your own code.

The **aggregate** package contains aggregation utilities, including improved versions of existing *R* functions, and new faster functions that drop down to *C* for speed.