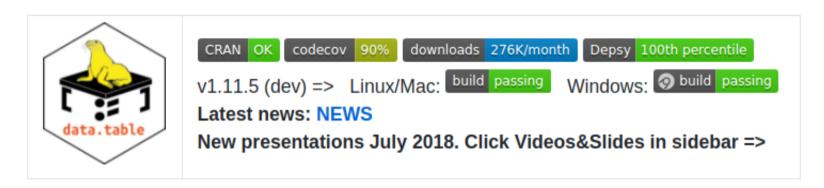


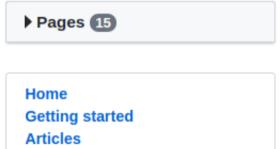
Home

Matt Dowle edited this page 6 days ago · 144 revisions



data.table is one of the 13,000 add-on packages for the programming language R which is popular in these fields. It provides a high-performance version of base R's data.frame with syntax and feature enhancements for ease of use, convenience and programming speed. As of Sep 2018, data.table was the 4th largest Stack Overflow tag about an R package with over 7,000 questions, the 10th most starred R package on GitHub and had over 650 CRAN and Bioconductor packages using it.

We have updated the 2014 grouping benchmarks comparing data.table to pandas and dplyr, and included Spark and pydatatable. The benchmark is automated and runs regularly against the latest versions of these packages. It is a work in progress.



Videos & Slides

Installation

Contributing

?data.table ?fread ?fwrite

Benchmarks: Grouping

fread for small data

Do's and Don'ts

🄰 @arun sriniv

data.table

#rdatatable
@MattDowle

Support