Databasing without the database: the *indexing* package

Jeffrey A. Ryan

R/Finance 2010
Chicago, Illinois

ia insight algorithmics inc.
Statistical software, training, and consulting

indexing mmap

indexing   mmap

index vectors or columns of any object

stored in an 'indexed_db' environment

fast binary search

regular R syntax for subsetting

... fast

ia   insight algorithmics inc.
      Statistical software, training, and consulting

indexing

mmap

index vectors or columns of any object

R access to OS-level mmap call

stored in an 'indexed_db' environment

map very large files into memory on demand

fast binary search

supports single byte to 64 bit complex

regular R syntax for subsetting

regular R syntax for subsetting

... fast

... did I mention *fast?*

ia insight algorithmics inc.
Statistical software, training, and consulting

Real Life Example

End of Day Equity Options Database

January 2, 2009 to October 31, 2009

67836671 rows by 13 columns

12 GB (columns & indexes)

ia | insight algorithmics inc.
Statistical software, training, and consulting

Thanks!

jeffrey.ryan @ insightalgo.com

http://r-forge.r-project.org/projects/indexing

ia insight algorithmics inc.
Statistical software, training, and consulting