Setup:

1. sudo Rscript -e 'install.packages(c("rJava", "Rcpp", "RJSONIO", "bitops", "digest", "functional", "stringr", "plyr", "reshape2", "caTools"), repos = "http://cran.r-project.org/")'
2. sudo wget <https://github.com/RevolutionAnalytics/rmr2/releases/download/3.3.1/rmr2_3.3.1.tar.gz>
3. sudo R CMD INSTALL rmr2\_3.3.1.tar.gz
4. **sudo Rscript -e 'install.packages(c("dplyr", "R.methodsS3", "Hmisc", "memoise", "lazyeval", "rjson"), repos = "http://cran.r-project.org/")'**
5. sudo wget <https://github.com/RevolutionAnalytics/plyrmr/releases/download/0.6.0/plyrmr_0.6.0.tar.gz>
6. sudo R CMD INSTALL plyrmr\_0.6.0.tar.gz
7. **sudo R**
8. hcmd <-system("which hadoop", intern = TRUE)
9. Sys.setenv(HADOOP\_CMD=hcmd)
10. hstreaming <- system("find /usr -name hadoop-streaming\*jar", intern=TRUE)
11. Sys.setenv(HADOOP\_STREAMING = hstreaming[1])
12. Sys.getenv("HADOOP\_CMD")
13. Sys.getenv("HADOOP\_STREAMING")
14. system("wget --no-check-certificate <http://github.com/RevolutionAnalytics/rhdfs/blob/master/build/rhdfs_1.0.8.tar.gz?raw=true>")
15. install.packages("rhdfs\_1.0.8.tar.gz?raw=true", repos = NULL, type="source")
16. q()