KCM Introduction Manual

Kane Capital Management

October 25, 2007

1 Overview

This document provides an overview of the software tools and readings with which a KCM summer associate should be familiar. Even a non-KCM summer associate can benefit from the information here as this provides an introduction to software tools needed for professional empirical data analysis. Future summer associates must use all these tools in the Spring semester before they begin work at KCM.

2 Operating Systems

Professionals use serious operating systems. At KCM we use and prefer Linux operating systems, but Mac OS X is also acceptable. Smart users will use Linux operating systems. Examples of good Linux operating systems are Ubuntu (`http://www.ubuntu.com/`) and Fedora (`http://fedoraproject.org/`).

3 R

KCM uses `R`, an open-source statistical programming language, in most of its operations. You can learn more about and download `R` at its homepage: `http://lib.stat.cmu.edu/R/CRAN/`. A nice overview of `R` can also be found at `http://en.wikipedia.org/wiki/R_%28programming_language%29`. First, read an *An Introduction to R* (`http://lib.stat.cmu.edu/R/CRAN/doc/manuals/R-intro.pdf`) and do all the exercises. Second, read Chapters 1-3 of *Writing R Extensions* (`http://lib.stat.cmu.edu/R/CRAN/doc/manuals/R-exts.pdf`). Third, read *Classes and Methods in the S Language* (`http://www.omegahat.org/RSMethods/Intro.pdf`), *S4 Classes in 15 pages, more or less* (`http://www.stat.auckland.ac.nz/S-Workshop/Gentleman/S4Objects.pdf`), and *Programmers' Niche: A Simple Class, in S3 and S4* (`http://cran.r-project.org/doc/Rnews/Rnews_2004-1.pdf`).

4 LaTeX

LaTeX (`http://en.wikipedia.org/wiki/LaTeX`) is a powerful typesetting system used by KCM (This document was generated using LaTeX in fact). Read *The Not so Short Introduction to LaTeX2ε* (`http://www.ctan.org/tex-archive/info/lshort/english/lshort.pdf`). References in LaTeX are typically done with BibTeX (`http://en.wikipedia.org/wiki/BibTeX`). See `http://www.andy-roberts.net/misc/latex/latextutorial3.html` for a tutorial on BibTeX.

5 Sweave

`Sweave` is an `R` package that lets users execute chunks of `R` code in a LaTeX document. Read *Sweave, Part I: Mixing R and LaTeX* (`http://www.ci.tuwien.ac.at/~leisch/Sweave/Sweave-Rnews-2002-3.pdf`), *Sweave, Part II: Package Vignettes* (`http://www.ci.tuwien.ac.at/~leisch/Sweave/Sweave-Rnews-2003-2.`

pdf), and the *Sweave User Manual* (`http://www.ci.tuwien.ac.at/~leisch/Sweave/Sweave-manual-20060104.pdf`).

6   XEmacs

XEmacs (`http://en.wikipedia.org/wiki/Xemacs`) is the text editor of choice at KCM. It's a powerful tool that is useful for creating documents, coding, and running R. XEmacs can be downloaded from `http://www.xemacs.org/Download/index.html`. Read *Getting Started with XEmacs* (http://ftp.xemacs.org/pub/xemacs/docs/letter/new-users-guide-letter.pdf.gz) and do the XEmacs tutorial (which appears in XEmacs when you press Ctrl-h t).

7   ESS

You can run R in XEmacs. To do so an XEmacs package called ESS (`http://en.wikipedia.org/wiki/Emacs_Speaks_Statistics`) must be installed. ESS can be downloaded from `http://ess.r-project.org/downloads/ess/`. Read the *ESS manual* (http://ess.r-project.org/ess.pdf) and `http://socserv.mcmaster.ca/jfox/Books/Companion/ESS/ess-xemacs.pdf` for more details.

8   Subversion

Good revision control (`http://en.wikipedia.org/wiki/Version_control`) is an important part of KCM. The revision control system we use is called Subversion (`http://en.wikipedia.org/wiki/Subversion_%28software%29`). Subversion can be downloaded at `http://subversion.tigris.org/`. Read the first 2 chapters of *Version Control with Subversion* (`http://svnbook.red-bean.com/`).

9   Conclusion

Good research is reproducible. Read *Reproducible Research: A Bioinformatics Case Study*(`http://www.bioconductor.org/docs/papers/2003/Compendium/Golub.pdf`) for an introduction. KCM also employs extreme programming (`http://en.wikipedia.org/wiki/Extreme_programming`) and especially pair-programming (`http://en.wikipedia.org/wiki/Pair_programming`). If you're interested in some of the work KCM does, see our publications page at `http://kanecap.com/publications.html`. In particular, the `backtest` package provides a good example of the sort of work that KCM summer associates do.