Programming with financial data:
Connecting R to MIM and Bloomberg

Dirk Eddelbuettel*
edd@debian.org

Submitted to useR! 2004

Abstract

In this paper, we discuss uses of R for 'programming with data' in an investment banking trading floor environment. We outline connector packages that permit direct access to both the Market Information Server (MIM), and the Bloomberg data servers.

1 Introduction

Trading floor environments require access to historical market data. In our case, the principal market data warehouse for data on historical security prices, as well as various related time series, contains several hundred thousand series comprising several gigabytes of content.

R, an environment for 'programming with data', to use the title of one of the defining books on the S language, is an extremely approriate choice for research analysts charged with producing empirical work for support of sales and trading. R can really shine in such an environment as one of its core strengths is — by design, it should be stressed — on interactive, or exploratory, data work with a particular focus on compelling visual representation of relationships in the data. However, the traditional platform in a trading room environment, arguably more by default than choice, is based on spreadsheets of often enormous size (and hence some complexity). Users who are new to R are often frustrated with the difficulty of having to extract data from one system (e.g. a database) before transforming it in order to import it into another system (e.g. R). Direct access to the data can aid tremendously in this regard. The D(C)OM server by Thomas Baier (c.f. http://cran.r-project.org/contrib/extra/dcom/) is of some use in this area too. However, the focus of this paper is on slightly more platform-independent ways to access data directly from other data repositories.

The Market Information Server, or MIM for short, is a powerful database backend provided by Logical Information Machines (http://www.lim.com) or LIM for short. The MIM database is a hierarchical database system (as opposed to a relational database system queried by SQL) which offers very fast access to data in time series form. Another key aspect of the MIM system is the ability to query the data in a particular language which attempts to be both expressive and relatively easy for non-programmers. The potential for work in the S language with data stored in MIM has not been lost on LIM who is offering a plugin for S-PLUS as well, c.f. http://www.lim.com/partners/splus_example.htm.

Bloomberg (http://www.bloomberg.com) provides a very powerful 'terminal' with access to an incredible number of data series, as well as (financial) functions to price and analyse an immense number of different security types. Bloomberg is widely regarded as a benchmark by market particants in debt, equity, foreign exchange and commodity markets.

This paper illustrates two in-house packages currently under development. These packages already permit highly efficient connections from R to both of these data services. We discuss the basic C language interface using the R `.call` function interface, as well as portability aspects between Windows and Solaris. A data quality control and validation example details one actual application before a summary concludes.

---

*The views expressed in this paper are those of its author, and do not necessarily reflect the positions or policies of his employer. The paper describes research by the author and is issued to elicit comments and to further discussion.