US Stocks

by Duncan Cummings, David Kane and Andy Yu Zhu Yao

Abstract An abstract of less than 150 words.

Introductory section which may include references in parentheses (Kane, 2008), or cite a reference such as Kane (2008) in the text.

Section title in sentence case

This section may contain a figure such as Figure 1.

Figure 1: The logo of R.

Getting Started

```
install_github("yuzhuyao/USstocks")
library(USstocks)
data(stocks)
```

The US Stocks Data

Variables and Meaning	
"id"	unique security identifier (randomly generated?)
"symbol"	stock exchange ticker
"v.date"	date of observation
"price.unadj"	unadjusted price
"price"	price (adjusted)
"volume.unadj"	unadjusted volume
"volume"	volume (adjusted)
"tret"	total returns (how is it defined?)
"m.sec"	sector to which the stock belongs
"m.ind"	industry to which the stock belongs
"name"	company name
"year"	year
"cap.usd"	market capitalization (of the year)
"top.1500"	boolean indicating whether the stock is part of the top 1500 performing in that year

Cleaning Up the Original Stocks Dataset

In our examination of the original stocks dataset(ws.data from Kane Capital), we found certain outliers. Some of these outlying observations may have been due to data-entry error, and some other may be caused by external factors (e.g. company fraud). Either way, for all intents and purposes of this package, we decided that it was best for the user to perform stock analysis without having to worry about these questionable outliers. In addition, we cross-validated the outlier results with historical data. A complete and detailed audit trail could be found on our github, but here, we summarize the list of stocks we removed and the reason(s) for their removal:

Name of Company	Individual/Everything*	Reason for Removal
CHATHAM CORP-DE	Everything	Unreasonably high return, extremely low trade volume
STRATOSPHERE CORP	Individual	Unreasonably high return
MFN FINANCIAL CORP	Individual	Unreasonably high return, extremely low trade volume
CYCLELOGIC INC	b	c
PTV INC	b	c
METRICOM INC	b	c
LORAL SPACE & COMMUNICATIONS	b	c
MARCHFIRST INC	b	c
RHYTHMS NETCONNECTIONS INC	b	c
CLARENT CORP	b	c
LUMINANT WORLDWIDE CORP	b	c
CKX INC	b	c
PACIFIC GATEWAY EXCHANGE INC	b	c
USINTERNETWORKING INC	b	c
GADZOOX NETWORKS INC	b	c
ACCRUE SOFTWARE INC	b	c
WEBVAN GROUP INC	b	c
SCIENT INC	b	c
ENGAGE INC	b	c

*Individual means that only the questionable outliers on particular days were removed. Everything means that every stock from that company was removed.

Having cleaned up our stocks data, we are now ready to dive into the USstocks data and explore!

Playing with the Data(I): Exploring the Data / Plotting

Playing with the Data(II): Looking at One Stock in Detail

Playing with the Data(III): Cross Sector/Industry Volatility Comparison

Playing with the Data(IV): Let's Buy Some Stocks!

Another section

There will likely be several sections, perhaps including code snippets, such as:

```
x <- 1:10
result <- myFunction(x)
```

Summary

This file is only a basic article template. For full details of *The R Journal* style and information on how to prepare your article for submission, see the Instructions for Authors.

Bibliography

D. Kane. ws.data, 2008. [p1]

Author One
Affiliation
Address

Country
author1@work

Author Two
Affiliation
Address
Country
author2@work

Author Three
Affiliation
Address
Country
author3@work