Subset & merge time series

Petri Fast *@petrifast*

Time series 1

2014-02-18 18:00:09 9.04
2014-02-18 18:00:16 10.23
…

XTS

Time series 2

2014-02-18 18:00:02 19.74
2014-02-18 18:01:04 20.18
2014-02-18 18:01:16 19.81
2014-02-18 18:01:40 20.31
2014-02-18 18:01:41 20.87
…

Chicago R User Group *Feb 19, 2014*

XTS time series

- *`library(xts)`* from CRAN — read the vignette
- *`X = xts( myMatrix, order.by= nicePosixTime)`*
  - *`index(X)`* is the time index (microseconds)
  - Single type elements so no factors
- *`X`* looks like a *`data.frame`* to many tools (*`lm`*)
- Excellent support for subsetting & merging time

Subset time

x['2007-03']

```
                 Open    High     Low       Close
2007-03-01 50.81620 50.81620 50.56451 50.57075
2007-03-02 50.60980 50.72061 50.50808 50.61559
                        …
```

On a date?

x['/2007-01-07']

```
2007-01-02 50.03978 50.11778 49.95041 50.11778
2007-01-03 50.23050 50.42188 50.23050 50.39767
2007-01-04 50.42096 50.42096 50.26414 50.33236
…
```

Before a date?

price2['T18:01:00/T18:01:40']

~~2014-02-18 18:00:02 19.74~~
2014-02-18 18:01:04 20.18
2014-02-18 18:01:16 19.81
2014-02-18 18:01:40 20.31
~~2014-02-18 18:01:41 20.87~~

Time range?

Merge timeseries

price1

2014-02-18 18:00:09 9.04
2014-02-18 18:00:16 10.23

price2

2014-02-18 18:00:02 19.74
2014-02-18 18:01:04 20.18
2014-02-18 18:01:16 19.81
2014-02-18 18:01:40 20.31
2014-02-18 18:01:41 20.87

prices <- merge(price1,price2)

```
                     price1 price2
2014-02-18 18:00:02     NA  19.74
2014-02-18 18:00:09   9.04     NA
2014-02-18 18:00:16  10.23     NA
2014-02-18 18:01:04     NA  20.18
2014-02-18 18:01:16     NA  19.81
2014-02-18 18:01:40     NA  20.31
2014-02-18 18:01:41     NA  20.87
```

Merge & fill in timeseries

```
filledPrices <- na.locf(prices)
                     price1 price2
2014-02-18 18:00:02     NA  19.74
2014-02-18 18:00:09   9.04  19.74
2014-02-18 18:00:16  10.23  19.74
2014-02-18 18:01:04  10.23  20.18
…
```

```
prices <- merge(price1,price2)

                     price1 price2
2014-02-18 18:00:02     NA  19.74
2014-02-18 18:00:09   9.04     NA
2014-02-18 18:00:16  10.23     NA
2014-02-18 18:01:04     NA  20.18
…
```

Merging=joining

To all fans of relational algebra…

```
(x <- xts(4:10, Sys.Date()+4:10))
(y <- xts(1:6, Sys.Date()+1:6))
merge(x,y)                  # either
merge(x,y, join='inner') # both
merge(x,y, join='left')  # merge to left
merge(x,y, join='right') # merge to right
```

Joins are useful

Getting started: reading time series

- Given a CSV file
- (1) Read into data.frame, e.g., myData
- (2) Convert the time field (e.g., myData[,1])
  - ti =as.POSIXct(strptime(myData[,1],myFormat))
  - X = xts(myData[,-1],order.by=ti)
- Now X is an XTS matrix