Notes Week 18

Jeff

2 mai 2016

Week 18, 2016 (week 9 of the internship)

I've finally begun this journal (weekly, actually, or hebdo en français), though as it's taken two months to get "comfortable"" with R it's late in the process. Onward quand même! The idea will be to compile todo lists and accomplishments in a running narration of the work.

This (short 3 day) week I have these goals:

- update shiny plot as function with inputs:
    - this score
    - average score
- add query types to shiny interface
- test with "seasonal" and monthly timescales
- create "interface tables" to speed interface queries (locations, etc)

```r
dbsub <- filter(tbl_scores, dateValue > "2005-01-01" & dateValue < "2005-12-31")
db2005 <- collect(dbsub)
summary(db2005)
```

```
##    row.names          locationID         scoreValue
## Length:35904       Length:35904       Min.   :  0.0000
## Class :character   Class :character   1st Qu.:  0.2861
## Mode  :character   Mode  :character   Median :  1.0170
##                                       Mean   :  3.8430
##                                       3rd Qu.:  3.0656
##                                       Max.   :448.4800
##                                       NA's   :342
##   scoreType          dateValue                LT
## Length:35904       Min.   :2005-02-01   Min.   : 1.00
## Class :character   1st Qu.:2005-04-01   1st Qu.:14.00
## Mode  :character   Median :2005-07-01   Median :39.50
##                    Mean   :2005-07-01   Mean   :40.91
##                    3rd Qu.:2005-10-01   3rd Qu.:65.00
##                    Max.   :2005-12-01   Max.   :90.00
##
```
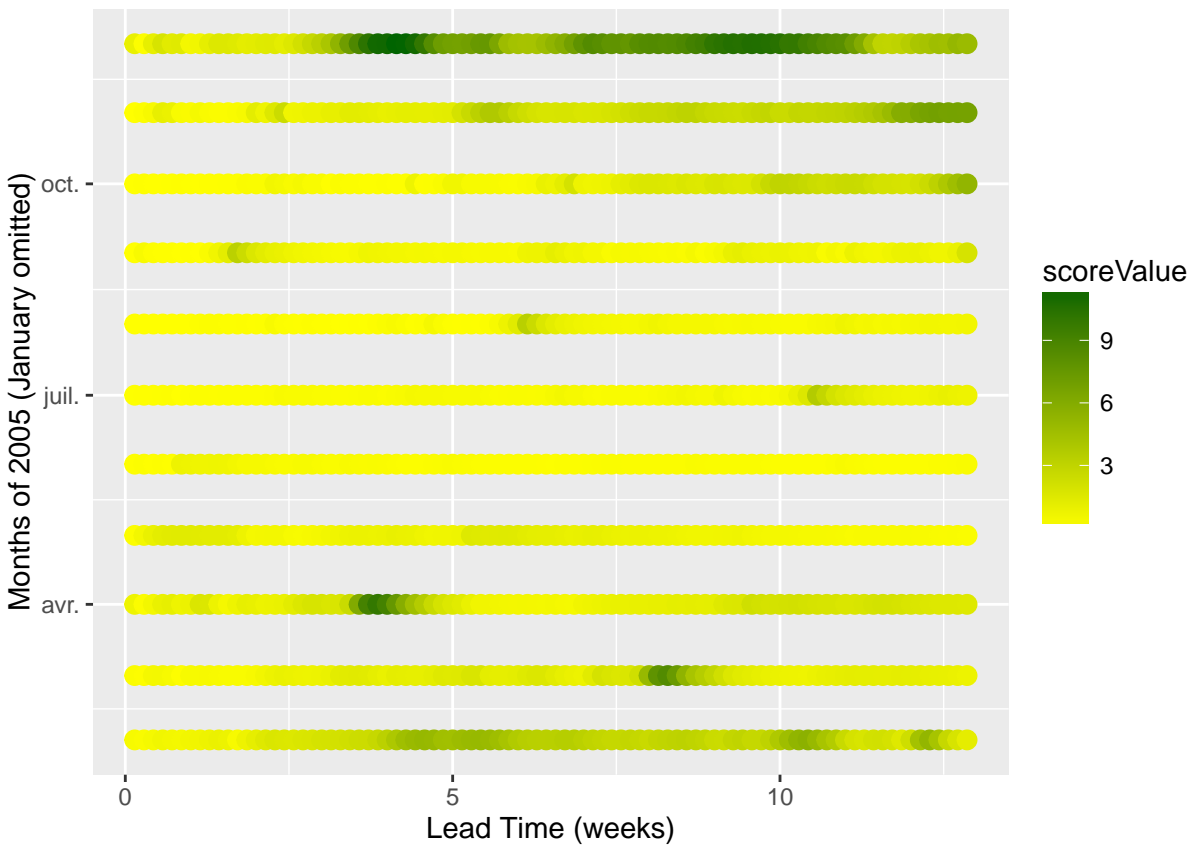
Including Plots

Subset of data:

```r
sm <- subset(db2005, locationID %in% c('S2242510') & scoreType == "Seasonal_LS_month")
summary(sm)
```

```
##   row.names          locationID            scoreValue
## Length:990        Length:990         Min.   : 0.000276
## Class :character   Class :character   1st Qu.: 0.348181
## Mode  :character   Mode  :character   Median : 0.873162
##                                       Mean   : 1.681303
##                                       3rd Qu.: 1.982486
##                                       Max.   :11.444298
##   scoreType          dateValue                LT
## Length:990        Min.   :2005-02-01   Min.   : 1.0
## Class :character   1st Qu.:2005-04-01   1st Qu.:23.0
## Mode  :character   Median :2005-07-01   Median :45.5
##                    Mean   :2005-07-01   Mean   :45.5
##                    3rd Qu.:2005-10-01   3rd Qu.:68.0
##                    Max.   :2005-12-01   Max.   :90.0
```
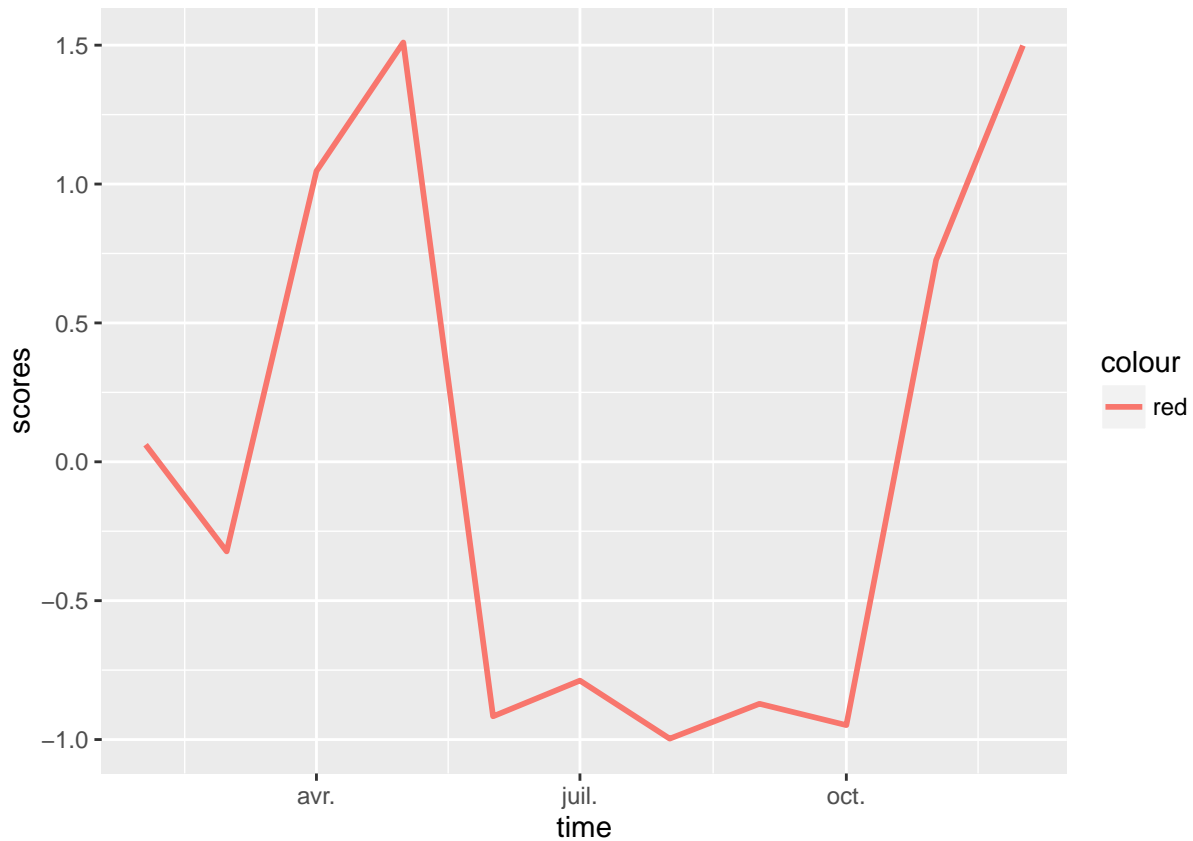
My plot de jour:

Sub-subset of data (lead time = 5 days):

```
smsub <- filter(sm, LT==5)
summary(smsub)
```

```
##   row.names          locationID            scoreValue
## Length:11         Length:11          Min.   :0.001751
## Class :character   Class :character   1st Qu.:0.058582
## Mode  :character   Mode  :character   Median :0.373556
```

```
##                                       Mean   :0.551504
##                                       3rd Qu.:1.040464
##                                       Max.   :1.384107
##    scoreType          dateValue              LT
##  Length:11         Min.   :2005-02-01   Min.   :5
##  Class :character  1st Qu.:2005-04-16   1st Qu.:5
##  Mode  :character  Median :2005-07-01   Median :5
##                    Mean   :2005-07-01   Mean   :5
##                    3rd Qu.:2005-09-16   3rd Qu.:5
##                    Max.   :2005-12-01   Max.   :5
```
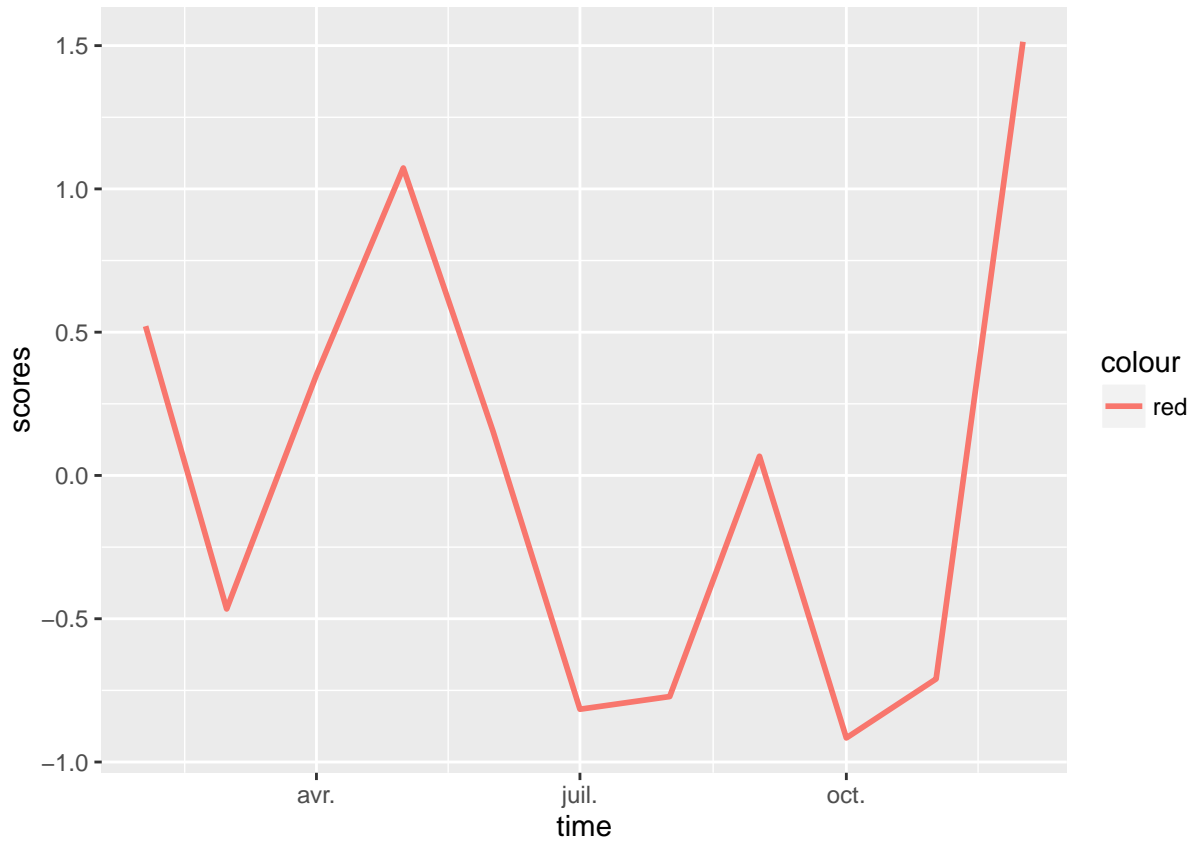

Sub-subset of data (lead time = 10 days):

```
smsub <- filter(sm, LT==10)
summary(smsub)
```

```
##    row.names          locationID          scoreValue        scoreType
##  Length:11         Length:11          Min.   :0.0537   Length:11
##  Class :character  Class :character   1st Qu.:0.1658   Class :character
##  Mode  :character  Mode  :character   Median :0.6824   Mode  :character
##                                       Mean   :0.6395
##                                       3rd Qu.:0.9182
##                                       Max.   :1.6073
##    dateValue               LT
##  Min.   :2005-02-01   Min.   :10
##  1st Qu.:2005-04-16   1st Qu.:10
```

```
##  Median :2005-07-01    Median :10
##  Mean   :2005-07-01    Mean   :10
##  3rd Qu.:2005-09-16    3rd Qu.:10
##  Max.   :2005-12-01    Max.   :10
```


Note - perhaps x axis should be time from start rather than an actual date ?