

Birthdays

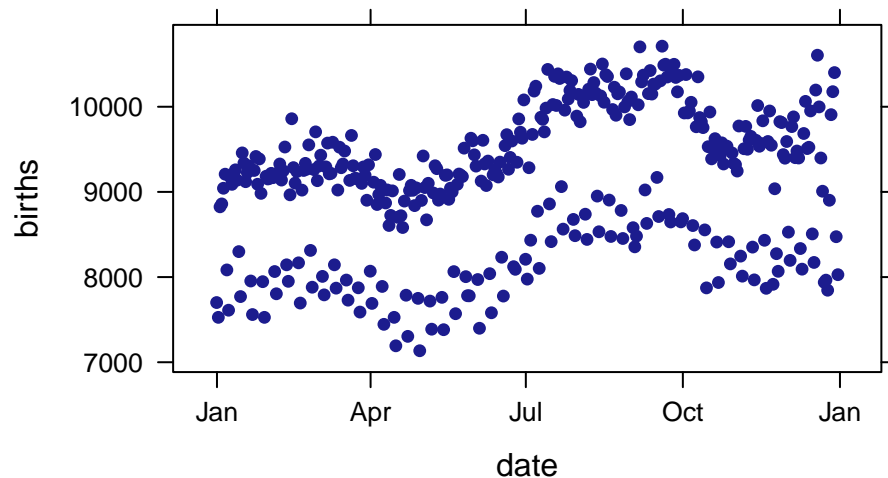
R Pruim

2014-09-08

Looking at Birthdays Again

We have already looked at this plot a couple times

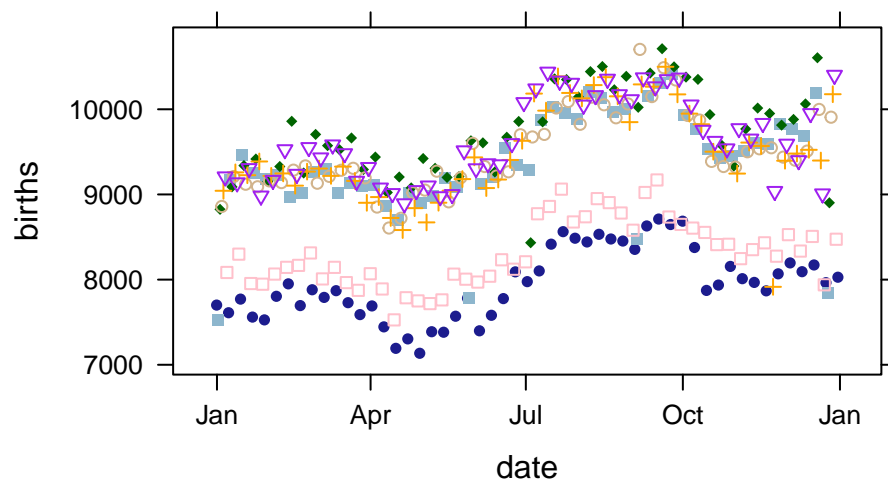
```
xyplot( births ~ date, data=Births78)
```



It was suggested that a possible explanation for the two parallel waves is a weekend/weekday affect. We can check whether this is plausible by adding some features to our plot.

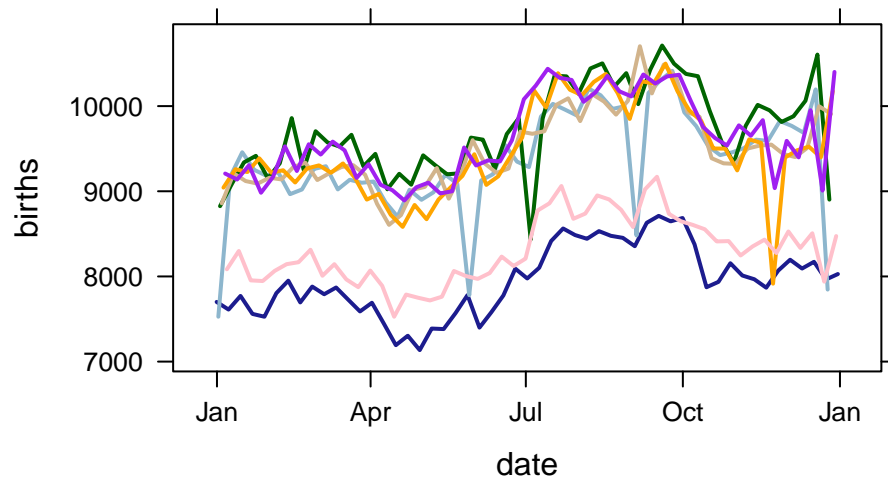
First let's color by day of week. The `wday` function in the `lubridate` package will convert dates to day of the week for us.

```
require(lubridate)
xyplot( births ~ date, data=Births78,
        groups=wday(date))
```



This looks like some good evidence to support our conjecture. There are primarily two colors in the lower wave and 5 in the upper wave – but there are some exceptions. These are easier to spot if we “connect the dots”.

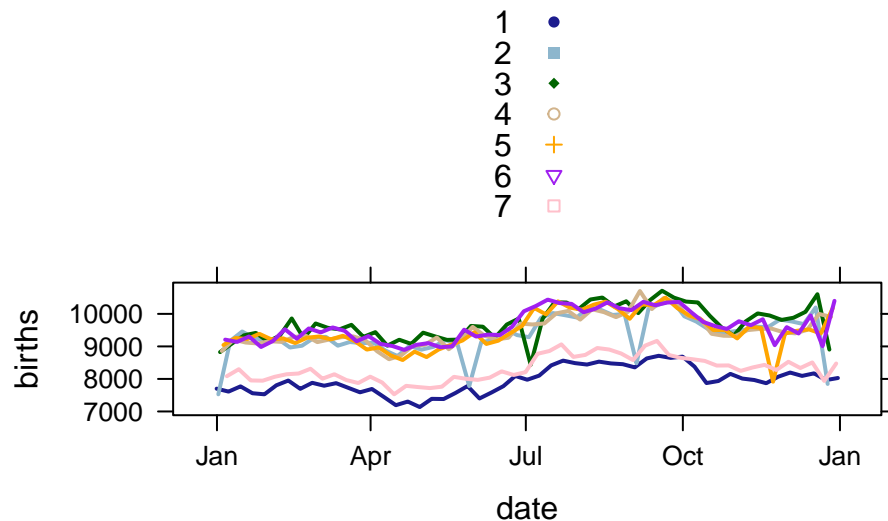
```
xyplot( births ~ date, data=Births78,
        groups=wday(date), type='l' )
```



Now it is easier to see that there are a few days that are “in the wrong group”. (These are the holidays.)

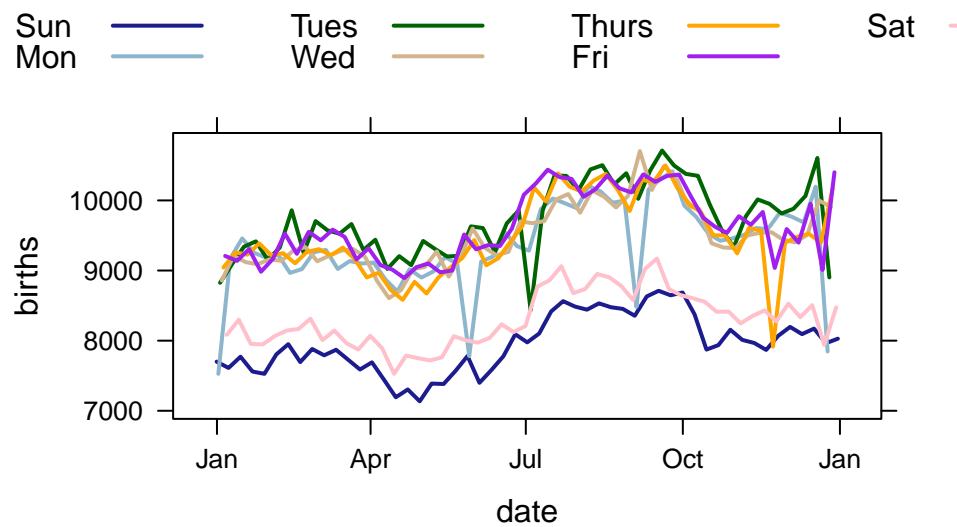
Finally, let’s add a key so we know which day is which.

```
xyplot( births ~ date, data=Births78,
        groups=wday(date), type='l',
        auto.key=TRUE)
```



Here’s a much nicer version (fancier than you need to know how to do just now).

```
xyplot( births ~ date, data=Births78,
        groups=wday(date, label=TRUE, abbr=TRUE), type='l',
        auto.key=list(columns=4, points=FALSE, lines=TRUE)
        )
```



We can also compute numbers and put them right into our text. For example, the mean number of births is 9132.1616. That's more digits than we need. We could choose to round this to 9132.