Date formatting

Gaëlle Cordier

06/11/2014

```r
head(act)
```

Data frame:
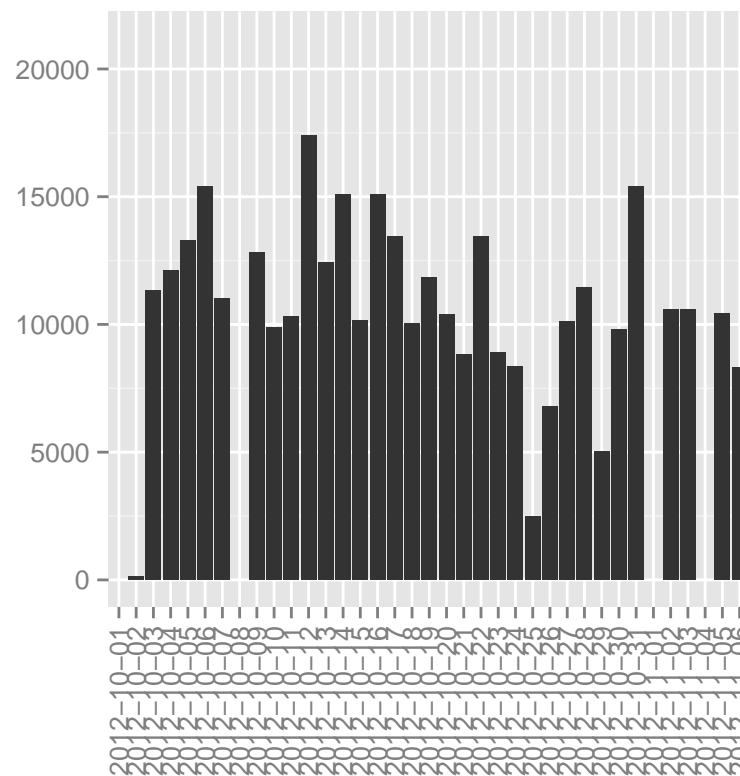
```
##   steps       date interval
## 1    NA 2012-10-01        0
## 2    NA 2012-10-01        5
## 3    NA 2012-10-01       10
## 4    NA 2012-10-01       15
## 5    NA 2012-10-01       20
## 6    NA 2012-10-01       25
```

```r
str(act)
```

```
## 'data.frame':    17568 obs. of  3 variables:
##  $ steps   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ date    : Factor w/ 61 levels "2012-10-01","2012-10-02",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ interval: int  0 5 10 15 20 25 30 35 40 45 ...
```

```r
library(ggplot2)
attach(act)

ggplot(data = act,aes(x=date,y=steps))+
    geom_histogram(stat="summary",fun.y="sum")+
    labs(y="",x="")+
    theme(axis.text.x=element_text(angle = 90, vjust = 0.5))
```

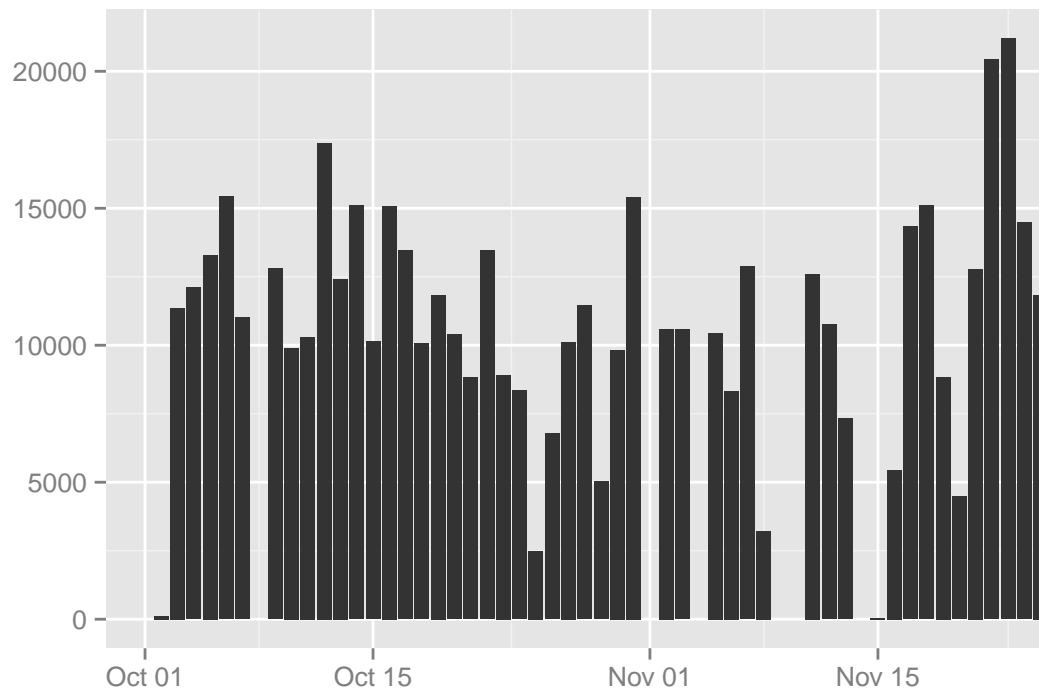**Original histogram with date as a factor:** ori-1.pdf

```
act2<-act
act2$date<-as.Date(act$date,format = "%Y-%m-%d")
str(act2)
```

Setting date as a date:

```
## 'data.frame':    17568 obs. of  3 variables:
##  $ steps   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ date    : Date, format: "2012-10-01" "2012-10-01" ...
##  $ interval: int  0 5 10 15 20 25 30 35 40 45 ...
```

```
ggplot(data = act2,aes(x = act2$date,y = act2$steps))+
    geom_histogram(stat="summary",fun.y="sum")+
    labs(title="Default options\n",y="",x="")
```
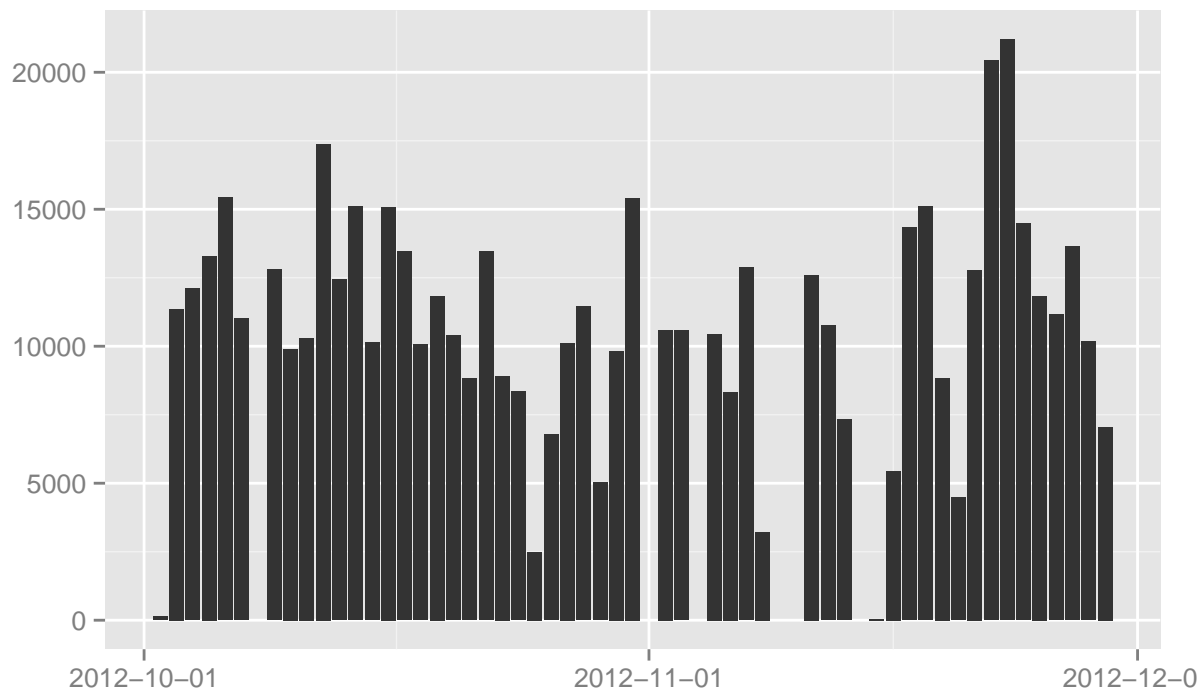
Default options

Histograms with date as date:

```r
ggplot(data = act2,aes(x = act2$date,y = act2$steps))+
    geom_histogram(stat="summary",fun.y="sum")+
    labs(title="Breaks = months\n",y="",x="")+
    scale_x_date(breaks = "months")
```
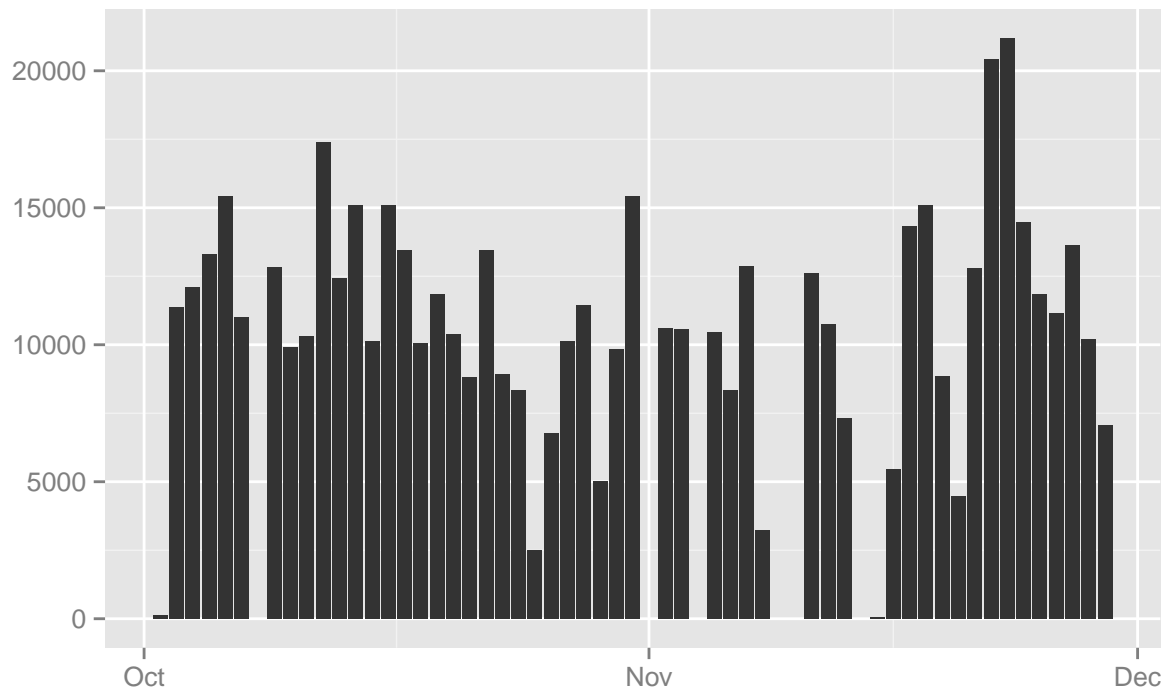
Breaks = months

```
library(scales) # to access breaks/formatting functions

ggplot(data = act2,aes(x = act2$date,y = act2$steps))+
    geom_histogram(stat="summary",fun.y="sum")+
    labs(title="Breaks = months, labels = month names\n",y="",x="")+
    scale_x_date(breaks = "months",labels=date_format("%b"))
```
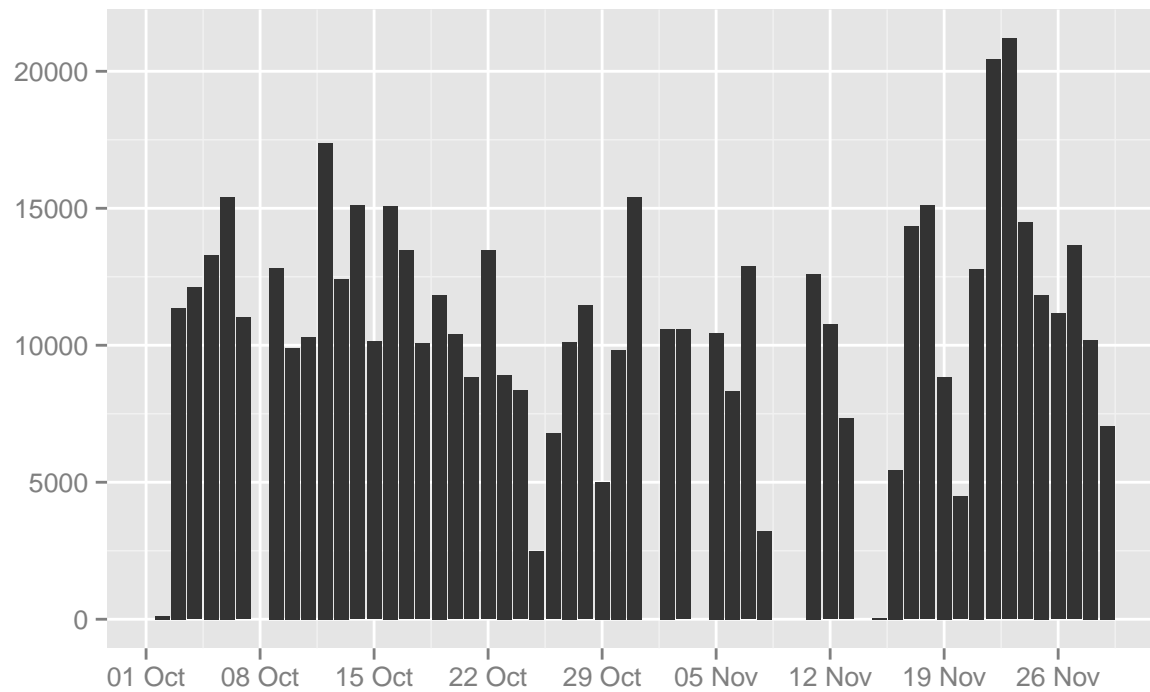
Breaks = months, labels = month names


```
ggplot(data = act2,aes(x = act2$date,y = act2$steps))+
    geom_histogram(stat="summary",fun.y="sum")+
    labs(title="Breaks = weeks, labels = day and month name\n",y="",x="")+
    scale_x_date(breaks = "weeks",labels=date_format("%d %b"))
```

Breaks = weeks, labels = day and month name

For more formatting options: scale_x_date