STA_445_Assignment_6

Myleen

2024-04-02

```
library(tidyverse)
library(lubridate)
library(mosaicData)
```

Problem 1

Convert the following to date or date/time objects.

a. September 13, 2010.

```
p1a <- mdy("September 13, 2010")
p1a
```

```
## [1] "2010-09-13"
```

```
class(p1a)
```

```
## [1] "Date"
```

b. Sept 13, 2010.

```
p1b <- mdy("September 13, 2010")
p1b
```

```
## [1] "2010-09-13"
```

```
class(p1b)
```

```
## [1] "Date"
```

c. Sep 13, 2010.

```
p1c <- mdy("September 13, 2010")
p1c
```

```
## [1] "2010-09-13"
```

```
class(p1c)
```

```
## [1] "Date"
```

d. S 13, 2010. Comment on the month abbreviation needs.

```
#incorrect
#p1d <- mdy("S 13, 2010")
#p1d
#class(p1d)

#corrected
```

```
p1d <- mdy("Sep 13, 2010")
p1d
```

```
## [1] "2010-09-13"
```

```
class(p1d)
```

```
## [1] "Date"
```

```
#instead of S it needs to be Sep
```

e. 07-Dec-1941.

```
p1e <- dmy("07-Dec-1941.")
p1e
```

```
## [1] "1941-12-07"
```

```
class(p1e)
```

```
## [1] "Date"
```

f. 1-5-1998. Comment on why you might be wrong.

```
p1f <- mdy("1-5-1998.")
p1f
```

```
## [1] "1998-01-05"
```

```
class(p1f)
```

```
## [1] "Date"
```

```
#I might be wrong because both month/day should be double digit and there is overlap between 1 and 12 f
```

g. 21-5-1998. Comment on why you know you are correct.

```
p1g <- dmy("21-5-1998.")
p1g
```

```
## [1] "1998-05-21"
```

```
class(p1g)
```

```
## [1] "Date"
```

```
# I know I am right because there is no month that is known as 21, therefore I know that it this goes d
```

h. 2020-May-5 10:30 am

```
p1h <- ymd_hm("2020-May-5 10:30 am")
p1h
```

```
## [1] "2020-05-05 10:30:00 UTC"
```

```
class(p1h)
```

```
## [1] "POSIXct" "POSIXt"
```

i. 2020-May-5 10:30 am PDT (ex Seattle)

```
p1i<- ymd_hm("2020-May-5 10:30 am",tz= "America/Los_Angeles")
p1i
```

```
## [1] "2020-05-05 10:30:00 PDT"
```

```
class(p1i)
```

```
## [1] "POSIXct" "POSIXt"
```

j. 2020-May-5 10:30 am AST (ex Puerto Rico)

```
p1j<- ymd_hm("2020-May-5 10:30 am",tz= "America/Puerto_Rico")
p1j
```

```
## [1] "2020-05-05 10:30:00 AST"
```

```
class(p1j)
```

```
## [1] "POSIXct" "POSIXt"
```

Problem 2

Using just your date of birth (ex Sep 7, 1998) and today's date calculate the following:

a. Calculate the date of your 64th birthday.

```
bday <- mdy("09/30/2004", tz="US/Arizona")
bday64 <- bday + years(64)
bday64
```

```
## [1] "2068-09-30 MST"
```

b. Calculate your current age (in years).

```
bday %--% now()
```

```
## [1] 2004-09-30 MST--2024-04-02 16:11:25 MST
```

```
as.period(bday %--% now())
```

```
## [1] "19y 6m 3d 16H 11M 25.912269115448S"
```

```
as.duration(bday %--% now())
```

```
## [1] "615571885.915607s (~19.51 years)"
```

c. Using your result in part (b), calculate the date of your next birthday.

```
nextBday <- bday + years(20)
nextBday
```

```
## [1] "2024-09-30 MST"
```

d. The number of *days* until your next birthday.

```
daysUntilBday <- as.period(now() %--% nextBday, unit="day" )
daysUntilBday
```

```
## [1] "180d 7H 48M 34.0710608959198S"
```

e. The number of *months* and *days* until your next birthday.

```
monDayUntilBday <- as.period(now()%--%nextBday)

monDayUntilBday
```

```
## [1] "5m 27d 7H 48M 34.0631520748138S"
```

Problem 3

Suppose you have arranged for a phone call to be at 3 pm on May 8, 2015 at Arizona time. However, the recipient will be in Auckland, NZ. What time will it be there?

```
phonecall <- mdy_hm("May 8, 2015 3:00pm", tz="US/Arizona")
phonecall
```

```
## [1] "2015-05-08 15:00:00 MST"
```

```
#with_tz(phonecall, tz="UTC")
with_tz(phonecall, tz="Pacific/Auckland")
```

```
## [1] "2015-05-09 10:00:00 NZST"
```

Problem 4

It turns out there is some interesting periodicity regarding the number of births on particular days of the year.

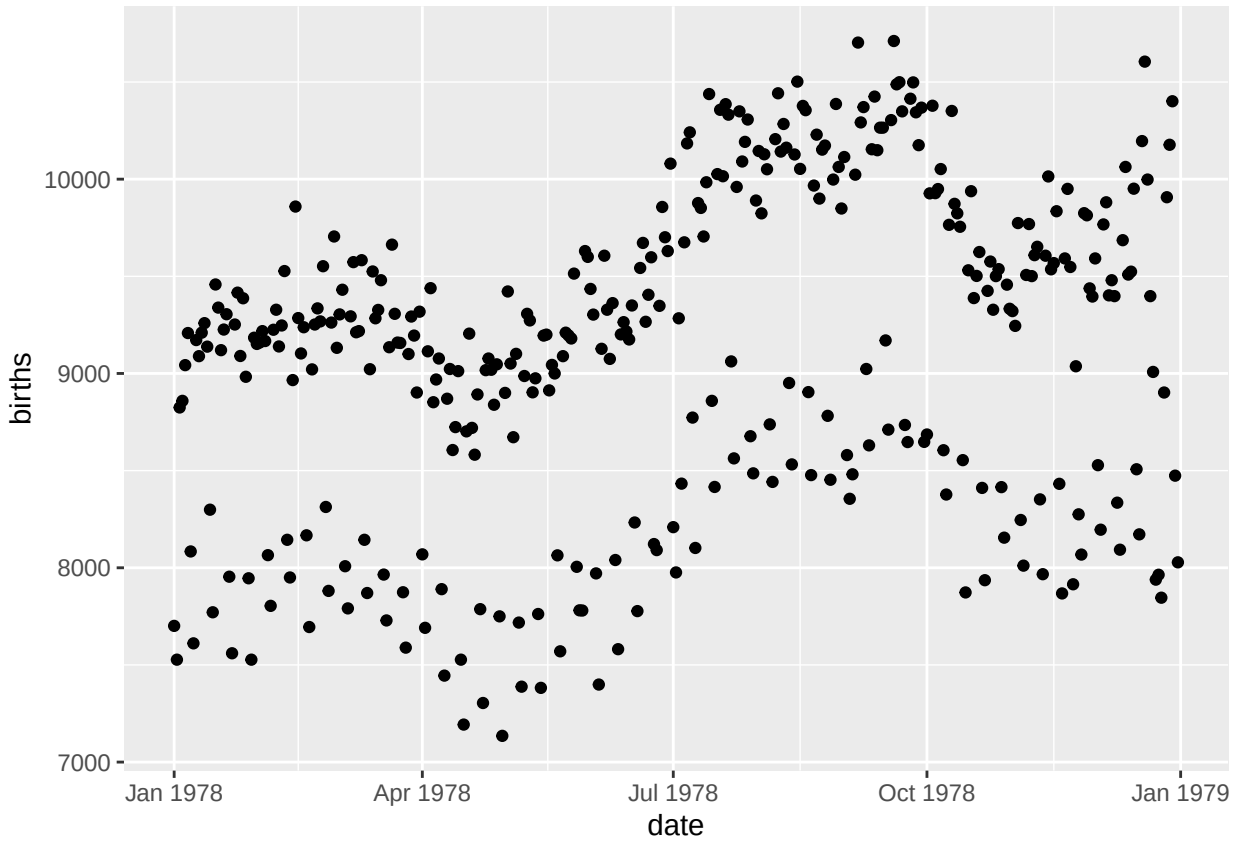
a. Using the `mosaicData` package, load the data set `Births78` which records the number of children born on each day in the United States in 1978. Because this problem is intended to show how to calculate the information using the `date`, remove all the columns *except* `date` and `births`.

```
#view(Births78)
#load("Births78")
births78 <- select(Births78, "date", "births")
head(births78)
```

```
##         date births
## 1 1978-01-01   7701
## 2 1978-01-02   7527
## 3 1978-01-03   8825
## 4 1978-01-04   8859
## 5 1978-01-05   9043
## 6 1978-01-06   9208
```

b. Graph the number of `births` vs the `date` with date on the x-axis. What stands out to you? Why do you think we have this trend?

```
ggplot(births78, aes(x=date, y=births))+
  geom_point()
```



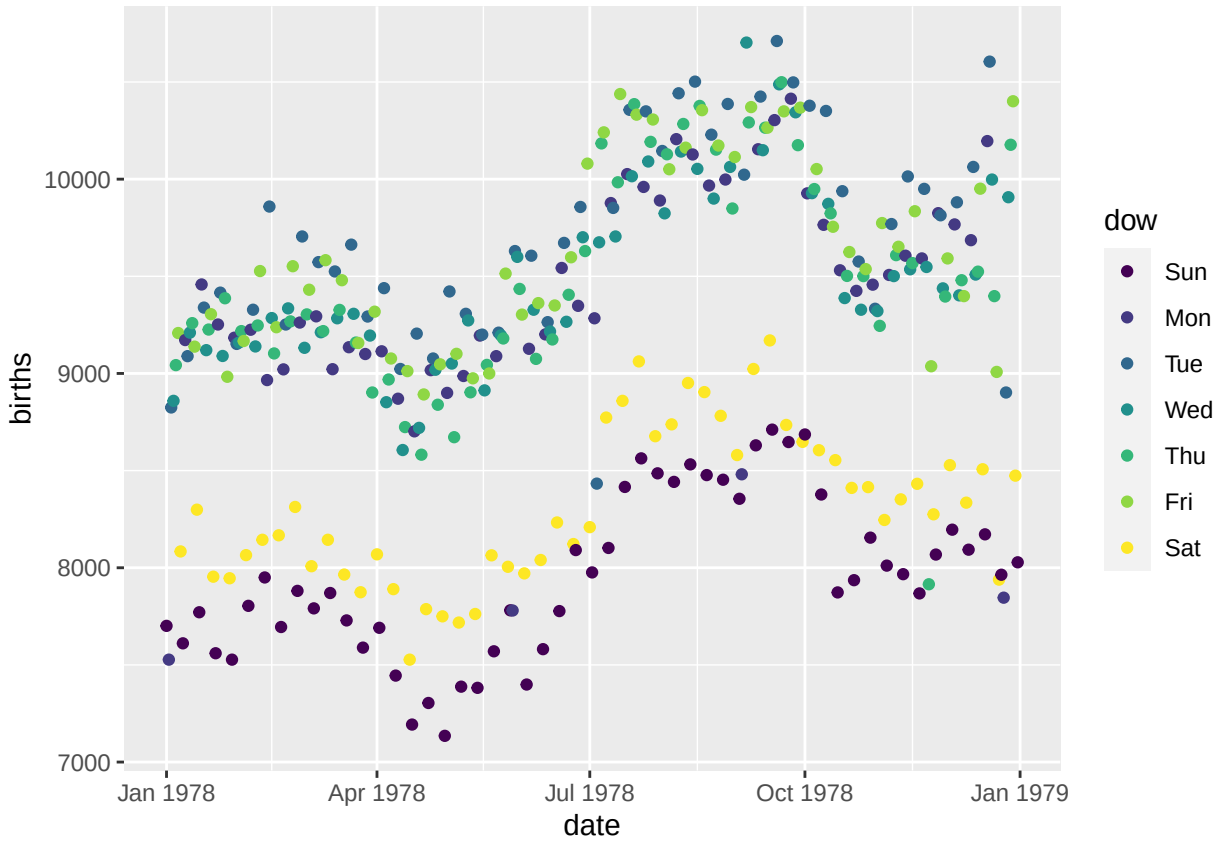
c. To test your assumption, we need to figure out the what day of the week each observation is. Use `dplyr::mutate` to add a new column named `dow` that is the day of the week (Monday, Tuesday, etc). This calculation will involve some function in the `lubridate` package and the `date` column.

```
new<-births78%>%
 mutate(dow = wday(births78$date, label=TRUE))
head(new)
```

```
##         date births dow
## 1 1978-01-01   7701 Sun
## 2 1978-01-02   7527 Mon
## 3 1978-01-03   8825 Tue
## 4 1978-01-04   8859 Wed
## 5 1978-01-05   9043 Thu
## 6 1978-01-06   9208 Fri
```

d. Plot the data with the point color being determined by the day of the week variable.

```
ggplot(new, aes(x=date, y=births, color=dow)) +
  geom_point()
```


births
10000
9000
8000
7000
Jan 1978
Apr 1978
Jul 1978
Oct 1978
Jan 1979
date
dow
Sun
Mon
Tue
Wed
Thu
Fri
Sat