case study in getting data from BOM

Nicholas Tierney

7 March 2016

These are the steps I go through to get data from BOM

- I go to the website: http://www.bom.gov.au/climate/data/?ref=ftr
- "select using text"
- Select Rainfall
- Select Daily Observations
- Select a weather station in the area of interest: "Brisbane"
- Choose the matching towns: "Brisbane, QLD, 27.47ºS, 153.03ºE"
- Choose Nearest Bureau stations: "040245 Toowong Bowls Club QLD (4.4km away)
- "Get data"
- select "All years of data"
- Open ZIP file
- read into R

Then I read in the data, get a sense of what the data looks like

```
library(readr)

dat_toowong <- read_csv("IDCjAC0009_040245_1800_Data.csv")

dim(dat_toowong)
```

```
## [1] 42641     8
```

```
library(dplyr)

glimpse(dat_toowong)
```

```
## Observations: 42,641
## Variables: 8
## $ Product code                                    (chr) "IDCJAC0009", "...
## $ Bureau of Meteorology station number            (chr) "040245", "0402...
## $ Year                                            (int) 1898, 1898, 189...
## $ Month                                           (chr) "01", "01", "01...
## $ Day                                             (chr) "01", "02", "03...
## $ Rainfall amount (millimetres)                   (dbl) NA, NA, NA, NA,...
## $ Period over which rainfall was measured (days) (int) NA, NA, NA, NA,...
## $ Quality                                         (chr) NA, NA, NA, NA,...
```

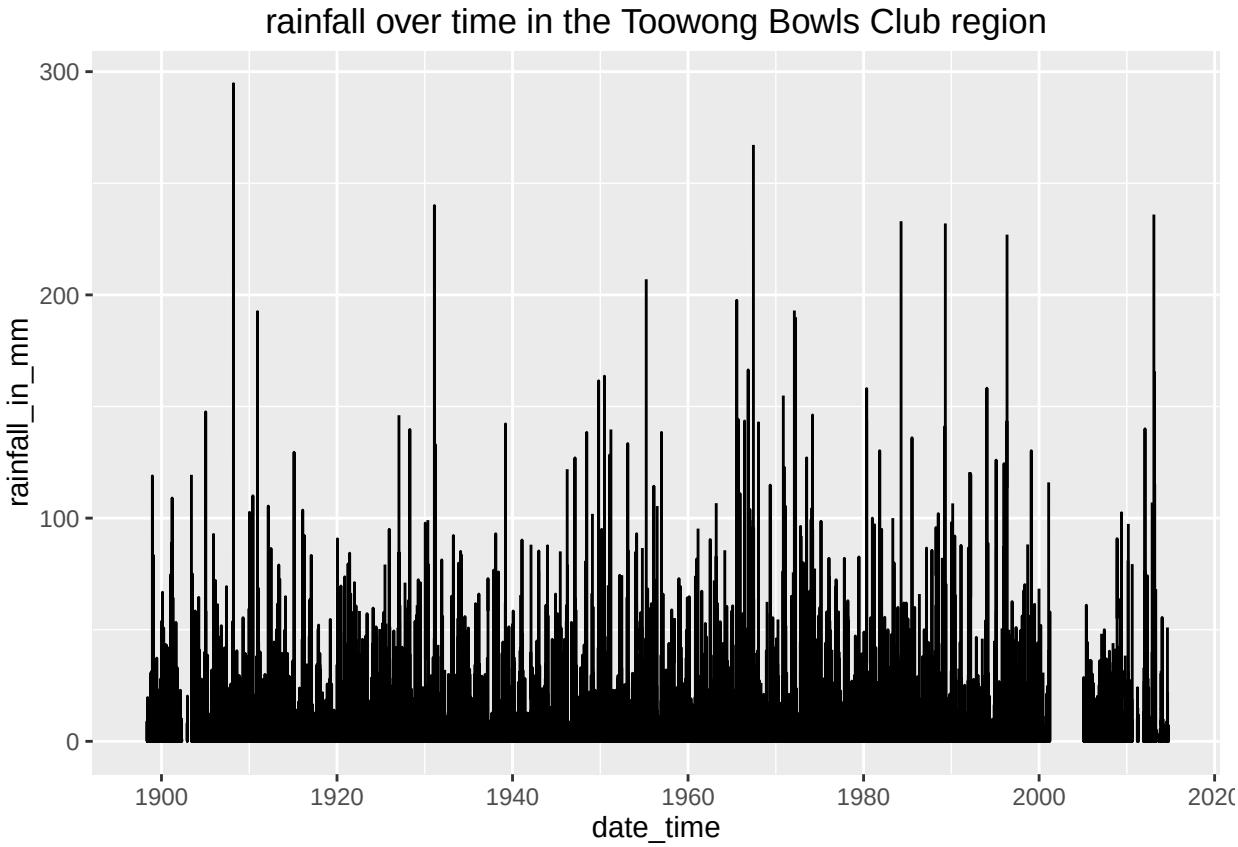
Then collect the date information into a format that R understands

```
# combine year-month-day into date string

library(lubridate)
```

```
dat_clean <-
  dat_toowong %>%
  mutate(date_time = paste(Year, Month, Day, sep = "-")) %>%
  mutate(date_time = ymd(date_time, tz = "Australia/Brisbane")) %>%
  rename(rainfall_in_mm = `Rainfall amount (millimetres)`)
```

then plot this in R

```
library(ggplot2)

ggplot(data = dat_clean,
       aes(x = date_time,
           y = rainfall_in_mm)) +
  geom_line() +
  labs(title = "rainfall over time in the Toowong Bowls Club region")
```



The good things are that the data from BOM is in a nice format, and required very little adjustment to get a nice plot.

However I think it would be nice to automate this from within R, so you could instead write something (roughly) like below:

```
library(bomr)

bomr_get_locations(location = "brisbane",
                   type = "rainfall",
```

```
                    frequency = "daily")

# list of locations around Brisbane.
#

data <-
bomr_get_data(long = "27.47S",
              lat = "153.03E",
              type = "rainfall",
              frequency = "daily",
              date_start = "01/01/2016",
              date_end = "01/29/2016")
```

It could also be possibel to merge your won dataset with date and long/lat info with BOM data, for example:

```
new_data <-  bomr_merge_data(new_data = my_data,
                             type = "rainfall",
                             frequency = "daily")
```

Which would avoid the user needing to do their own date matching on their own computer.