The tsibbledata packages contains the vic\_elec data set, containing half-hourly electricity demand for the state of Victoria, along with corresponding temperatures from the capital city, Melbourne. These data cover the period 2012-2014.

Other similar data sets are also available, and these may be of interest to researchers in the area.

**Australian state-level demand**

Anyone can still load and use the data with the following code.

library(tidyverse)

library(lubridate)

library(tsibble)

repo <- "<https://raw.githubusercontent.com/tidyverts/tsibbledata/master/data-raw/vic_elec/>"

states <- c("NSW","QLD","SA","TAS","VIC")

dirs <- paste0(repo, states, "2015")

# Read holidays data

holidays <- paste0(dirs,"/holidays.txt") %>%

as.list() %>%

map\_dfr(read\_csv, col\_names=FALSE, .id="State") %>%

transmute(

State = states[as.numeric(State)],

Date = dmy(X1),

Holiday = TRUE

)

# Read temperature data

temperatures <- paste0(dirs,"/temperature.csv") %>%

as.list() %>%

map\_dfr(read\_csv, .id = "State") %>%

mutate(

State = states[as.numeric(State)],

Date = as\_date(Date, origin = ymd("1899-12-30"))

)

# Read demand data

demands <- paste0(dirs,"/demand.csv") %>%

as.list() %>%

map\_dfr(read\_csv, .id = "State") %>%

mutate(

State = states[as.numeric(State)],

Date = as\_date(Date, origin = ymd("1899-12-30"))

)

# Join demand, temperatures and holidays

aus\_elec <- demands %>%

left\_join(temperatures, by = c("State", "Date", "Period")) %>%

transmute(

State,

Time = as.POSIXct(Date + minutes((Period-1) \* 30)),

Period,

Date = as\_date(Time),

DOW = wday(Date, label=TRUE),

Demand = OperationalLessIndustrial,

Temperature = Temp,

) %>%

left\_join(holidays, by = c("State", "Date")) %>%

replace\_na(list(Holiday = FALSE))

# Remove duplicates and create a tsibble

aus\_elec <- aus\_elec %>%

filter(!are\_duplicated(aus\_elec, index=Time, key=State)) %>%

as\_tsibble(index = Time, key=State)

This block of code reads in raw data files containing holiday information, temperatures and electricity demand for each state, and then joins them into a single tsibble. For some reason, there are duplicated rows from South Australia, so the last few lines removes the duplicates before forming a tsibble, keyed by State.

aus\_elec

## # A tsibble: 1,155,408 x 8 [30m] <UTC>

## # Key: State [5]

## State Time Period Date DOW Demand Temperature

## <chr> <dttm> <dbl> <date> <ord> <dbl> <dbl>

## 1 NSW 2002-01-01 00:00:00 1 2002-01-01 Tue 5714. 26.3

## 2 NSW 2002-01-01 00:30:00 2 2002-01-01 Tue 5360. 26.3

## 3 NSW 2002-01-01 01:00:00 3 2002-01-01 Tue 5015. 26.3

## 4 NSW 2002-01-01 01:30:00 4 2002-01-01 Tue 4603. 26.3

## 5 NSW 2002-01-01 02:00:00 5 2002-01-01 Tue 4285. 26.3

## 6 NSW 2002-01-01 02:30:00 6 2002-01-01 Tue 4075. 26.3

## 7 NSW 2002-01-01 03:00:00 7 2002-01-01 Tue 3943. 26.3

## 8 NSW 2002-01-01 03:30:00 8 2002-01-01 Tue 3884. 26.3

## 9 NSW 2002-01-01 04:00:00 9 2002-01-01 Tue 3878. 26.3

## 10 NSW 2002-01-01 04:30:00 10 2002-01-01 Tue 3838. 26.3

## # … with 1,155,398 more rows, and 1 more variable: Holiday <lgl>

This data set contains half-hourly data from all states from 1 January 2002 - 1 March 2015 (and in the case of Queensland to 1 April 2015). The temperature variable is from a weather station in the capital city of each state.

**GEFCOM 2017**

The [Global Energy Forecasting Competition in 2017](http://www.drhongtao.com/gefcom/2017) involved data on hourly zonal loads of ISO New England from March 2003 to April 2017. The data have already been packaged into tibble format by Cameron Roach. So it is relatively easy to convert this to a tsibble.

devtools::install\_github("camroach87/gefcom2017data")

library(gefcom2017data)

gefcom2017 <- gefcom %>%

ungroup() %>%

as\_tsibble(key=zone, index=ts)

gefcom2017

## # A tsibble: 1,241,710 x 15 [1h] <UTC>

## # Key: zone [10]

## ts zone demand drybulb dewpnt date year month

## <dttm> <chr> <dbl> <dbl> <dbl> <date> <dbl> <fct>

## 1 2003-03-01 00:00:00 CT 3386 25 19 2003-03-01 2003 Mar

## 2 2003-03-01 01:00:00 CT 3258 23 18 2003-03-01 2003 Mar

## 3 2003-03-01 02:00:00 CT 3189 22 18 2003-03-01 2003 Mar

## 4 2003-03-01 03:00:00 CT 3157 22 19 2003-03-01 2003 Mar

## 5 2003-03-01 04:00:00 CT 3166 23 19 2003-03-01 2003 Mar

## 6 2003-03-01 05:00:00 CT 3255 23 20 2003-03-01 2003 Mar

## 7 2003-03-01 06:00:00 CT 3430 24 20 2003-03-01 2003 Mar

## 8 2003-03-01 07:00:00 CT 3684 24 20 2003-03-01 2003 Mar

## 9 2003-03-01 08:00:00 CT 3977 25 21 2003-03-01 2003 Mar

## 10 2003-03-01 09:00:00 CT 4129 27 22 2003-03-01 2003 Mar

## # … with 1,241,700 more rows, and 7 more variables: hour <dbl>,

## # day\_of\_week <fct>, day\_of\_year <dbl>, weekend <lgl>,

## # holiday\_name <chr>, holiday <lgl>, trend <dbl>