BDCD Business Analytics Lab Exercise

Prof Di Cook, Steph Kobakian, Stuart Lee, Nick Spyrison

Econometrics & Bus Stat, Monash, Clayton campus, 22/2/2019

Melbourne 2050

What do you think Melbourne will be like in 2050? Can you imagine what your life will be like, if you are still living here? What will be different, what will be the same?

I imagine a city where people have control of data about them. They can download this data, and make plots to learn about their life, and what they might do to improve it.

We are going to do this with one data set that we currently have access to, your household energy use. Our goal is to be energy-wise, for a greener future. Climate change is consistently in the news. The consequences for Melbourne are more extreme high temperatures and drought. Individuals can play a big role in mitigating the extremes of the future, if they know how their actions contribute to the world as a whole. All of Melbourne households now have smart meters, which report the energy use every 30 minutes.

Maybe in the future Melbourne City will run competitions with fun prizes for the most energy efficient households, or most improved household!

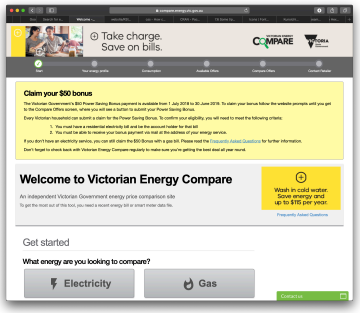
Monash Green

Task

Working with the smart meter data for Professor Di's household, we wrangle it into shape, and make plots to explore household energy use. We'll try to answer questions like: What times of day does the household use the most energy, or time of year? Is it related to the weather, with air conditioner or heater usage, or special events, like washing and drying clothes the night before a holiday trip, or dinner parties with friends. Then we will see if Professor Di is an energy hog.

All the tools used are open source software that will be available to you into the future.

Getting started



1. *Web research*

a. Point your web browser to this site: `https://compare.energy.vic.gov.au`. How can you earn $50 from your energy data?

What's a smart meter? Take a look at the web site `http://www.smartmeters.vic.gov.au/#` How many smart meters have been installed across Victoria?

b. Point your browser to Business intelligence guide to R in

https://www.computerworld.com. Read through the first page. (The second two pages its ok to scan but there are a couple of bad practices that need to be ignored.)

2. Get materials

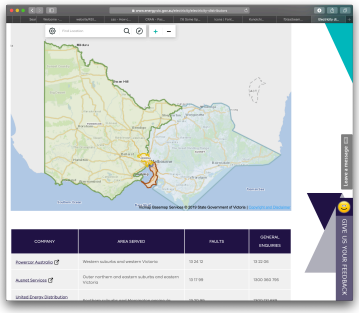
Materials for the workshop can be downloaded from `https://github.com/Monash-BDCD/energy`. It will download and you will need to **unzip onto the Desktop** of your computer with the name "energy-master", by default. **Change it to "energy".**

There are several files that will download:

- `lab.Rmd` (This is an Rmarkdown file, a special plain text document that contains code and explanations, and can compiled compile into a Word, html or pdf document.)
- `lab.html` (This has compiled document.)
- `data` is a directory containing some sample energy files
- `energy_app` a directory that contains files to get you started on making a web app
- `energy.Rproj` An R project, clicking on this will open RStudio (and R) on your computer.

3. About the data

Data collected by downloading Di's (and friend of hers) electricity usage data as recorded by the household smart meter. Details on how to do this are (**THIS IS ONLY IF YOU WANT TO COLLECT YOUR OWN HOUSEHOLD'S DATA**):



- Find your distributor
- Create an account, using your meter number which you can find on a current bill

Maybe after this workshop, you can do your own household, upload it to the compare suppliers site, claim your $50 (and get your parent to pay you this for your efforts), and possibly get a better deal on household energy costs.

Follow along with me for these exercises

Exercise 1: Background work

Read in Di's energy data. Look at the format of the data, and then rearrange it to a tidier format.

```
## # A tibble: 3 x 7
##   id    date       d1    d2    d3    d4    d5
##   <chr> <chr>   <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 200   meter1     NA    NA    NA    NA    NA
```

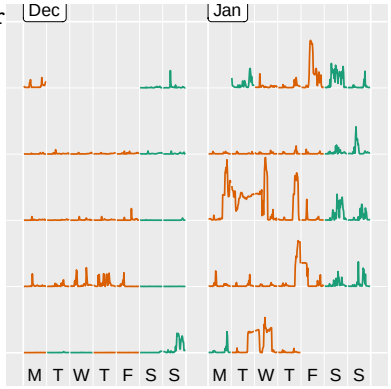
```
## 2 300   201711~     0     0     0     0     0
## 3 300   201711~     0     0     0     0     0
```

Here's the wrangling, , and new format:

```
## # A tibble: 3 x 9
##   id    date       halfhour   kwh wday  month
##   <chr> <date>        <dbl> <dbl> <ord> <ord>
## 1 300   2017-11-24      0.5     0 Fri   Nov
## 2 300   2017-11-24      1       0 Fri   Nov
## 3 300   2017-11-24      1.5     0 Fri   Nov
## # ... with 3 more variables: year <dbl>,
## #   dt <dttm>, work <chr>
```

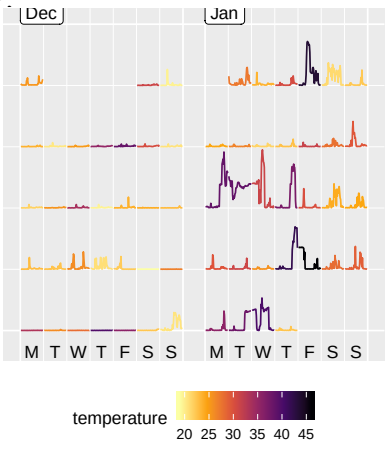
Exercise 2: Plot (some of) the data in a calendar layout

```
p1 <- elec %>% filter(date < dmy("01022019"),
    date >= dmy("01122018")) %>% frame_calendar(x = halfhour,
    y = kwh, date = date, ncol = 4) %>% ggplot(aes(x = .halfhour
    y = .kwh, group = date, colour = factor(work))) +
    geom_line() + scale_colour_brewer("work",
    palette = "Dark2") + theme(legend.position = "none")
prettify(p1)
```



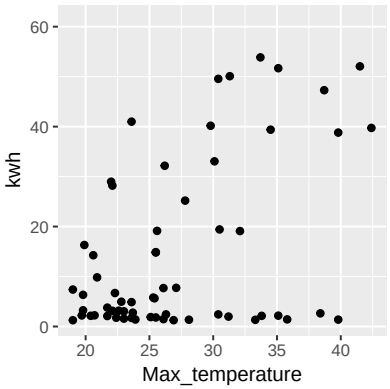
Exercise 3: Combine with weather data

```
p1 <- elec %>% filter(date < dmy("01022019"),
    date >= dmy("01122018")) %>% frame_calendar(x = halfhour,
    y = kwh, date = date, ncol = 4) %>% ggplot(aes(x = .halfhour
    y = .kwh, group = date, colour = Max_temperature)) +
    geom_line() + scale_colour_viridis_c("temperature",
    option = "inferno", direction = -1) + theme(legend.position
prettify(p1)
```



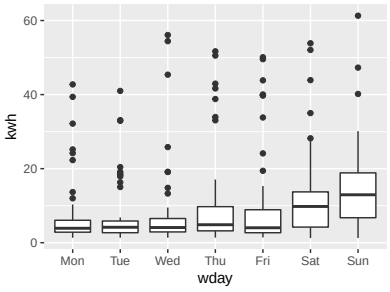
Exercise 4: Make a simple plot, and add some interactivity

```
library(plotly)
p <- elec %>% filter(year(date) == 2018) %>% group_by(date) %>%
    summarise(kwh = sum(kwh)) %>% left_join(maxtemp_DF) %>%
    ggplot(aes(x = Max_temperature, y = kwh, text = date)) +
    geom_point()
p
```



Exercise 6: Compare day of the week

```
p <- elec %>% filter(year(date) == 2018) %>% group_by(date) %>%
    summarise(kwh = sum(kwh)) %>% mutate(wday = wday(date,
    label = TRUE, week_start = 1)) %>% ggplot(aes(x = wday,
    y = kwh)) + geom_boxplot()
p
```



Your turn

1. Easy tasks:
   a. Create a test Rmarkdown document
   b. Compile it into an html document
2. Medium task: Create a plot to answer these questions:
   a. What day of the week does the household use more energy? What day of the week is there more variation in energy use?
   b. What time of day does the household wake up? And go to sleep? Is it different on a holiday vs a working day?
   c. Can you guess when Di's son was home from Uni? Was there a time when the family was away on holidays?
3. Difficult: Make a new app to study energy usage. The steps to do this are:
   a. Run the sample app, and play with changing dates
   b. Add a numerical summary to the output
   c. Make a new tab with a new display, and numerical summary, something that you think is useful to track in energy usage.

How to win the grand prize

There is a prize for the team that works together the best on the lab activities.

1. Get the best score in the interactive online quiz (after lunch)
2. Produce the most creative web app, with interesting data plots, that someone can use to explore their energy use