"Run 4 Course 1 Week 4 Office Hour

Jiaying Wu

22/09/2020

Step 4.3: Communicating the message

```r
# path to the data
path <- "../data/meter_01.csv"
```

```r
library(tidyverse)
library(lubridate)
library(sugrrants)
library(tsibble)
library(glue)
```

```r
elec <- read_csv(path, skip=1,
                 col_names = c("id", "date",  1:48),
                 col_types =str_c("cc", str_c(rep("d", 48), collapse = ""), str_c(rep("-", 5), collapse
)
```

```
## Warning: 90 parsing failures.
## row col expected   actual                  file
##   1   1 a double E1E2      '../data/meter_01.csv'
##   1   2 a double E1        '../data/meter_01.csv'
##   1   3 a double E1        '../data/meter_01.csv'
##   1   5 a double CZ226453 '../data/meter_01.csv'
##   1   6 a double KWH       '../data/meter_01.csv'
## ... ... ........ ........ .....................
## See problems(...) for more details.
```

```r
elec %>% select(id:`5`) %>% head(3)
```

```
## # A tibble: 3 x 7
##   id    date        `1`   `2`   `3`   `4`   `5`
##   <chr> <chr>     <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 200   meter1       NA    NA    NA    NA    NA
## 2 300   20171124      0     0     0     0     0
## 3 300   20171125      0     0     0     0     0
```

```r
vic_holidays <- holiday_aus(2017:2019, state = "VIC")
```

```r
elec <- elec %>% filter(id == 300)
```

```
elec <- elec %>%
  mutate(date = ymd(date)) %>%
  gather(halfhour, kwh, -id, -date) %>%
  mutate(halfhour = as.numeric(halfhour) / 2) %>%
  arrange(date, halfhour) %>%
  mutate(
    wday = wday(date, label = TRUE, abbr = TRUE, week_start = 1),
    month = month(date, label = TRUE, abbr = TRUE),
    year = year(date),
    dt = ymd_hm(glue("{date} 12:00"), tz = "Australia/Melbourne") + minutes(60*halfhour)
  ) %>%
  mutate(work = ifelse(wday %in% c("Mon", "Tue", "Wed", "Thu", "Fri"), "workday", "holiday")) %>%
  mutate(work = ifelse(date %in% vic_holidays$date, "holiday", work))

elec %>% head(3)
```

```
## # A tibble: 3 x 9
##   id    date       halfhour   kwh wday  month year dt                  work
##   <chr> <date>        <dbl> <dbl> <ord> <ord> <dbl> <dttm>              <chr>
## 1 300   2017-11-24      0.5     0 Fri   Nov    2017 2017-11-24 12:30:00 workday
## 2 300   2017-11-24      1       0 Fri   Nov    2017 2017-11-24 13:00:00 workday
## 3 300   2017-11-24      1.5     0 Fri   Nov    2017 2017-11-24 13:30:00 workday
```

lot (some of) the data in a calendar layout

```
p1 <- elec %>%
  filter(date < dmy("01022019"), date >= dmy("01122018")) %>%
  ggplot(aes(x = halfhour, y = kwh)) +
  geom_line(aes(colour = work)) +
  scale_colour_brewer("work", palette = "Dark2") +
  facet_calendar(~ date) +
  xlim(c(0,24)) +
  theme(legend.position="none")

p1
```

Combine with weather data

```r
library(bomrang)
library(viridis)
stations <- sweep_for_stations(latlon = c(-37.8136, 144.9631)) #Melbourne lat/long is
maxtemp <- get_historical(stationid = "086282", type = "max")
maxtemp <- maxtemp %>%
  mutate(date = paste(maxtemp$year, maxtemp$month, maxtemp$day, sep="-")) %>%
  mutate(date = ymd(date))
maxtemp_df <- maxtemp %>% filter(year > 2017, month %in% c(12, 1))
elec <- elec %>%
  left_join(maxtemp_df, by="date")
```

```r
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 = "bottom")

prettify(p1)
```
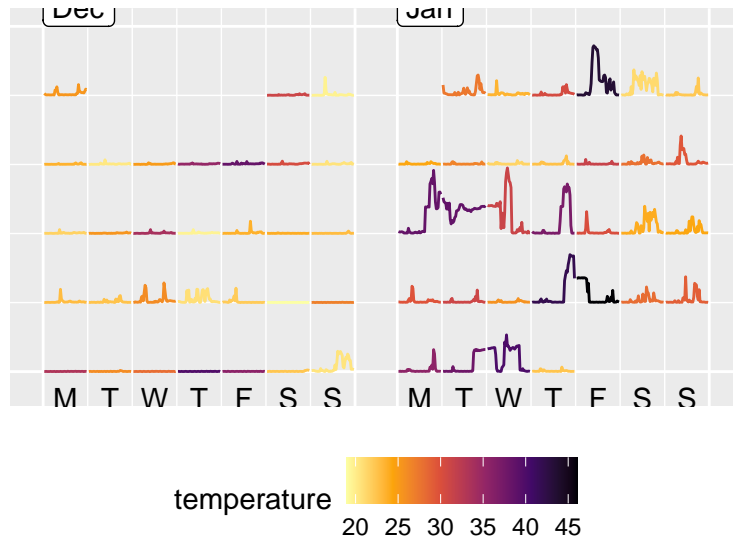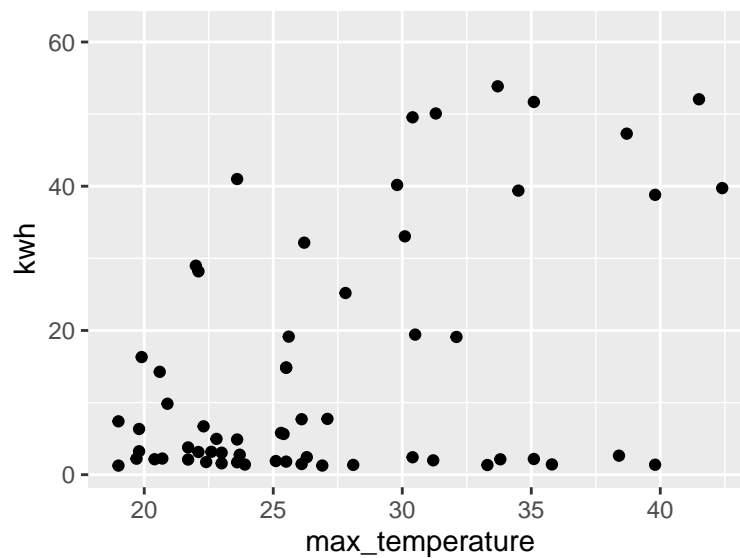
Make a simple plot

```
p <- elec %>%
  filter(year(date) == 2018) %>%
  group_by(date) %>%
  summarise(kwh = sum(kwh)) %>%
  left_join(maxtemp_df, by="date") %>%
  ggplot(aes(x=max_temperature, y=kwh, text=date)) + geom_point()
p
```

```
## Warning: Removed 303 rows containing missing values (geom_point).
```


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
```