R2 - Visualization

Richard T. Watson

2020-12-27 10:00:05

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

Prepare the data

```
url <-  "http://www.richardtwatson.com/data/ATLweather.csv"
w <- read_delim(url,delim=',')
url <-  "http://www.richardtwatson.com/data/electricityprices.csv"
e <- read_delim(url,delim=',')
m <- inner_join(w,e,by=c("Timestamp" = 'timestamp'))
```

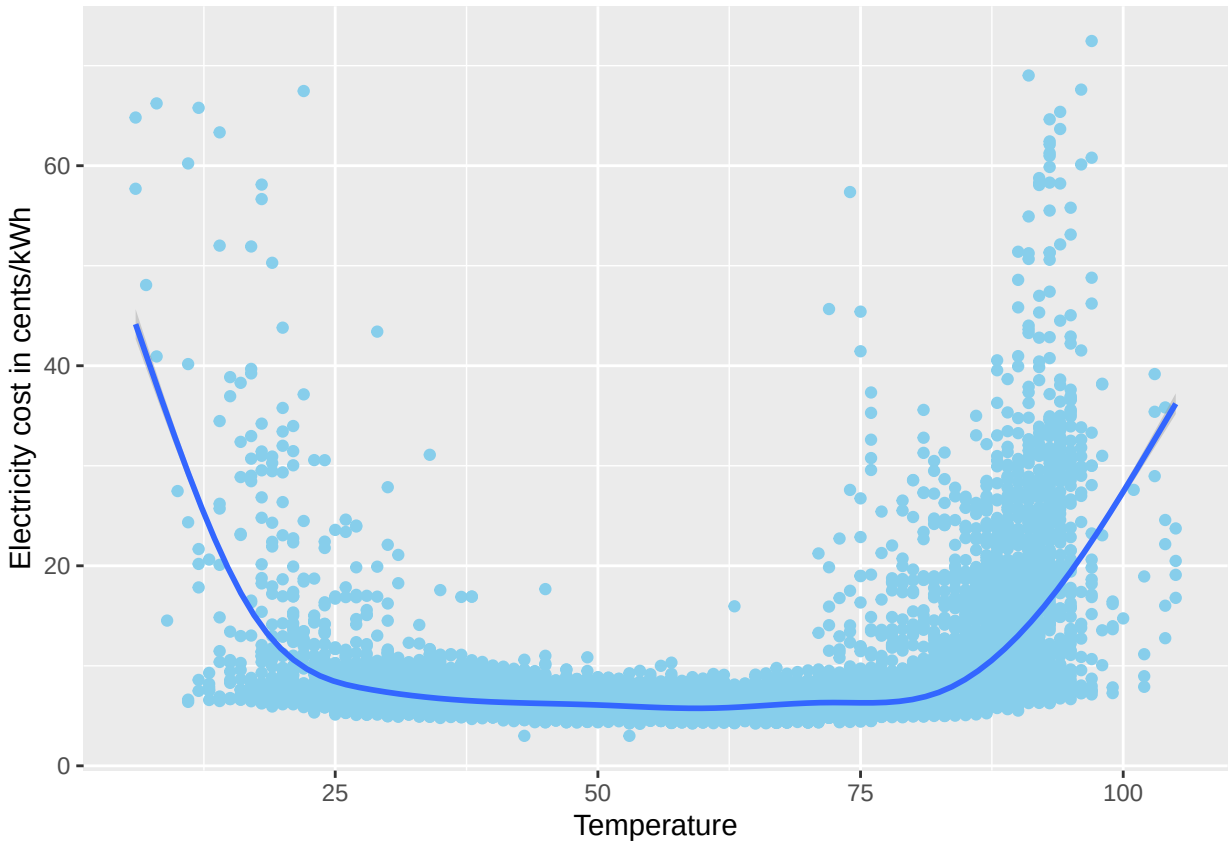
a.

Graph the relationship between temperature and electricity price with a smoother. What do you conclude?

```
ggplot(m,aes(Temperature,cost)) +
  geom_point(color='skyblue') +
  geom_smooth() +
  xlab('Temperature') +
  ylab('Electricity cost in cents/kWh')
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```



Conclusion Electricity is more expensive on very cold and very hot days.

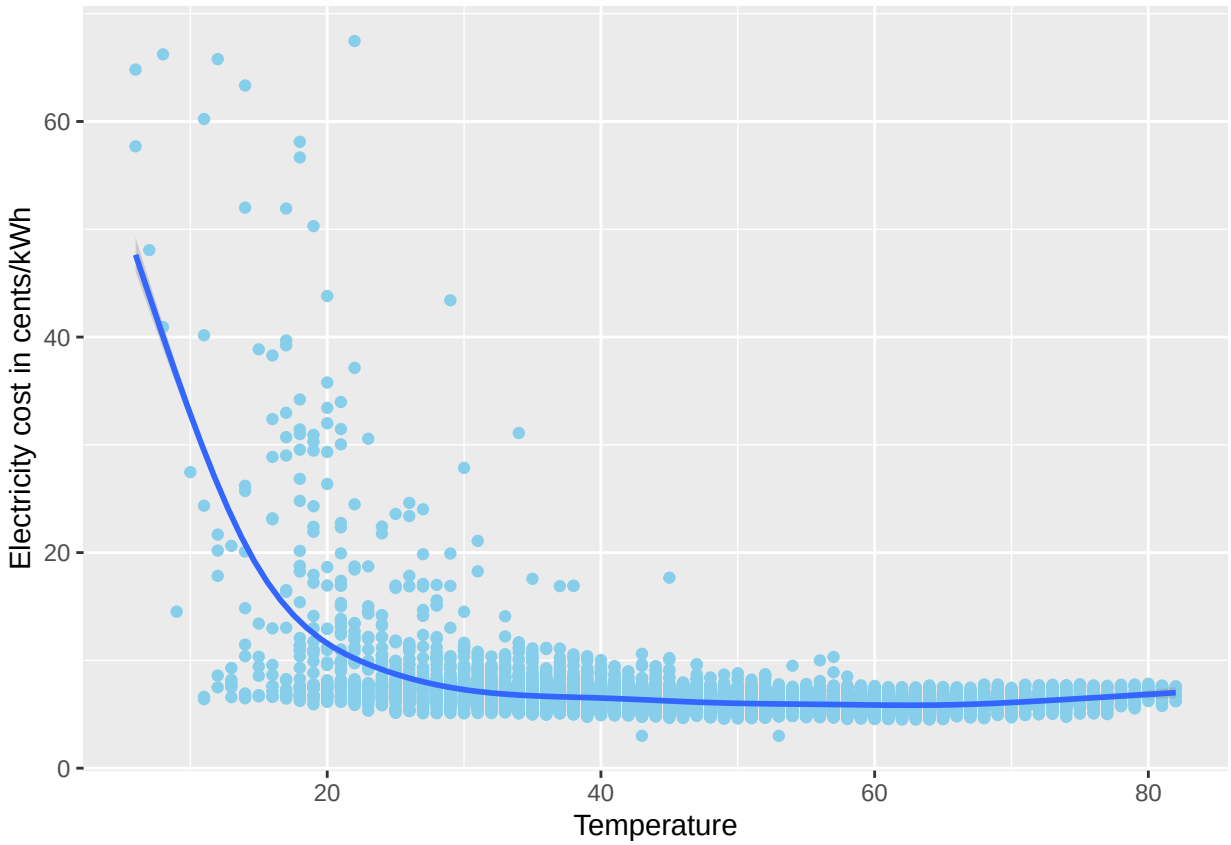
b.

Graph the relationship between the temperature and electricity price for winter (January to March) with a smoother.

```
winter <- m %>%
  filter(month(Timestamp) >= 1 & month(Timestamp) <= 3)
ggplot(winter,aes(Temperature,cost)) +
  geom_point(color='skyblue') +
  geom_smooth() +
  xlab('Temperature') +
  ylab('Electricity cost in cents/kWh')
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

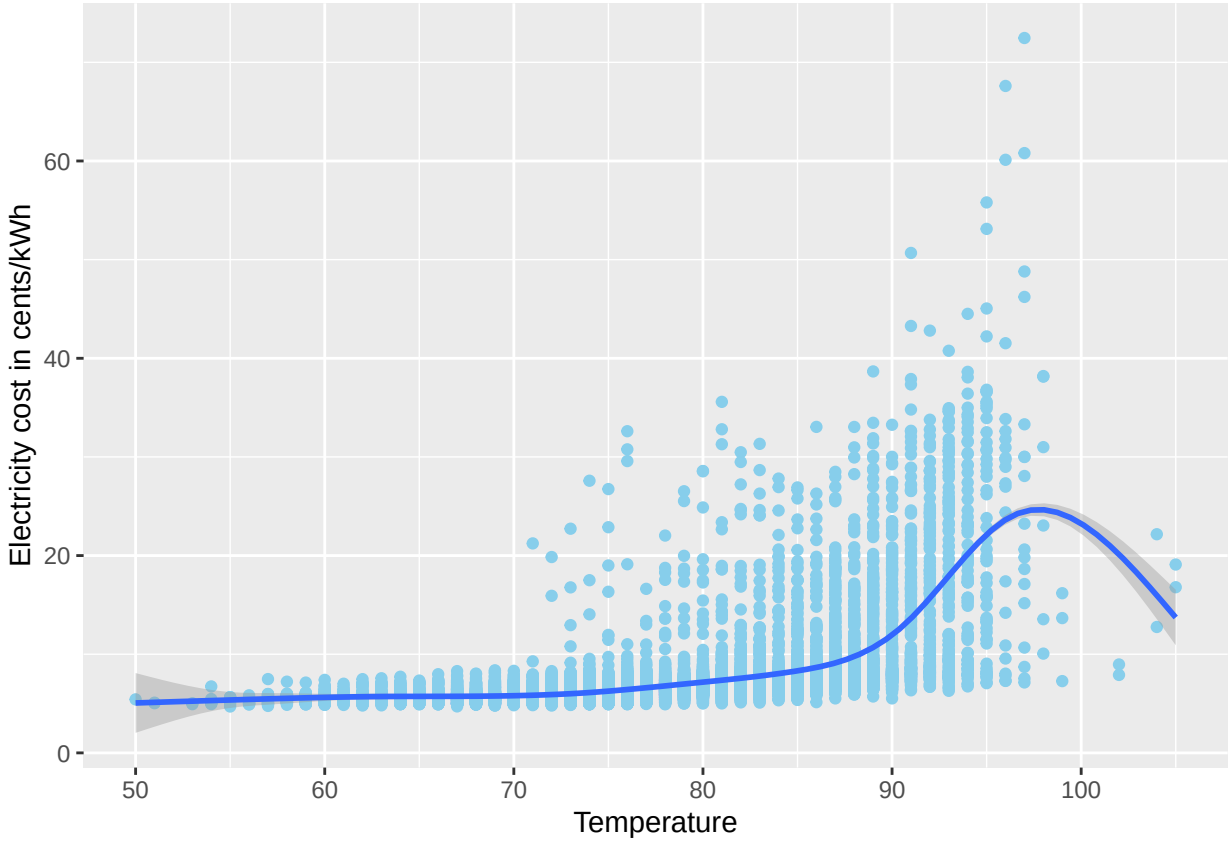
```
## Warning: Removed 1 rows containing missing values (geom_point).
```



c.

Graph the relationship between the temperature and electricity price for summer (July to September) with a smoother.

```
summer <- m %>%
  filter(month(Timestamp) >= 7 & month(Timestamp) <= 9)
ggplot(summer,aes(Temperature,cost)) +
  geom_point(color='skyblue') +
  geom_smooth() +
  xlab('Temperature') +
  ylab('Electricity cost in cents/kWh')
```



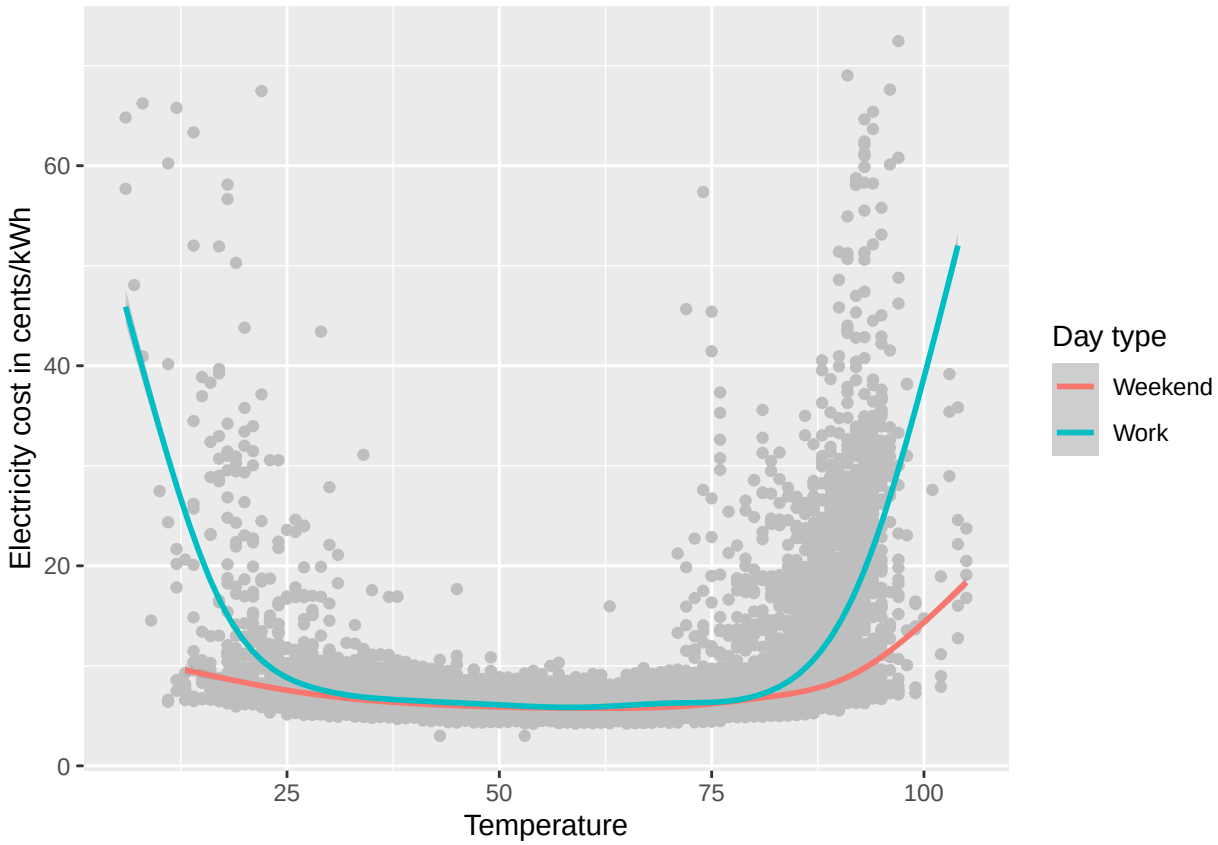
d.

Create a column that categorize a day as either 'week' or 'weekend'. Create a visualization showing separate line graphs for the relationship between the temperature and electricity price for the week and the weekend. What do you conclude?

```
# Separating work and weekend days
m$dayType <- case_when (
  wday(m$Timestamp) > 1 & wday(m$Timestamp) < 7 ~ 'Work',
  wday(m$Timestamp) == 1 | wday(m$Timestamp) == 7 ~ 'Weekend'
)
ggplot(m,aes(Temperature, cost, color = dayType)) +
  geom_point(color='grey') +
  geom_smooth() +
  labs(
  x = 'Temperature',
  y = 'Electricity cost in cents/kWh',
  color = 'Day type')
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```



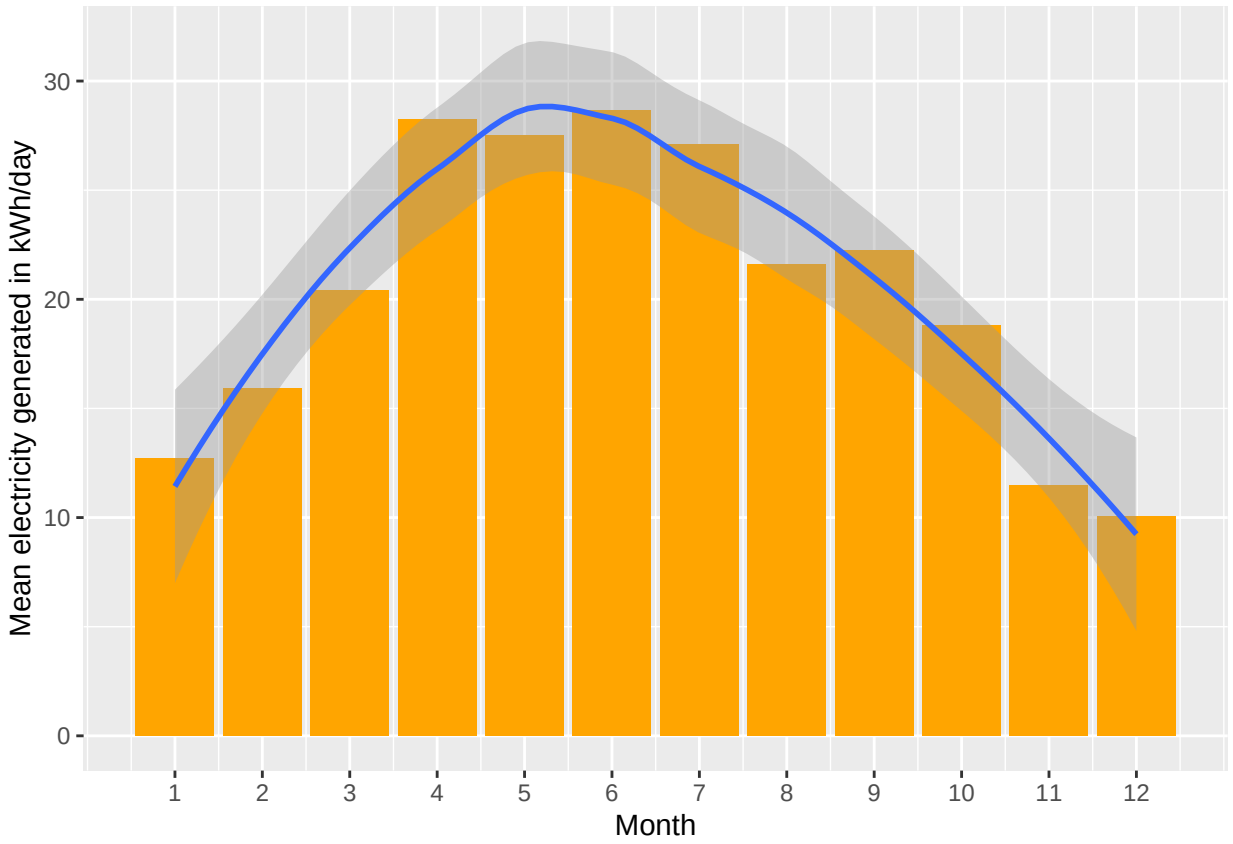
Conclusion

On very cold and very hot days, electricity is cheaper on the weekend, probably because the demand is lower with many office and factories closed. The effect is more pronounced in summer because electricity is used for cooling, and in winter gas and electricity are used for heating.

e.

Create a column (geom_col) chart showing average electricity generated for each day of the month in kWh for a house with 20m2 of solar panels of 22% efficiency. Label the months 1 through 12 consecutively.

```
url <-  "http://www.richardtwatson.com/data/SolarRadiationAthens.csv"
s <- read_delim(url,delim=',')
s2 <- s %>%
  group_by(month = month(TimeStamp)) %>%
  summarize(Mean_kWh = mean(SolarWatt*24/1000*20*.22))
ggplot(s2,aes(month,Mean_kWh)) +
  geom_col(fill='orange') +
  xlab('Month') +
  ylab('Mean electricity generated in kWh/day') +
  geom_smooth() +
  scale_x_continuous(breaks = seq(1, 12, by = 1))
```


Mean electricity generated in kWh/day
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
12
Month