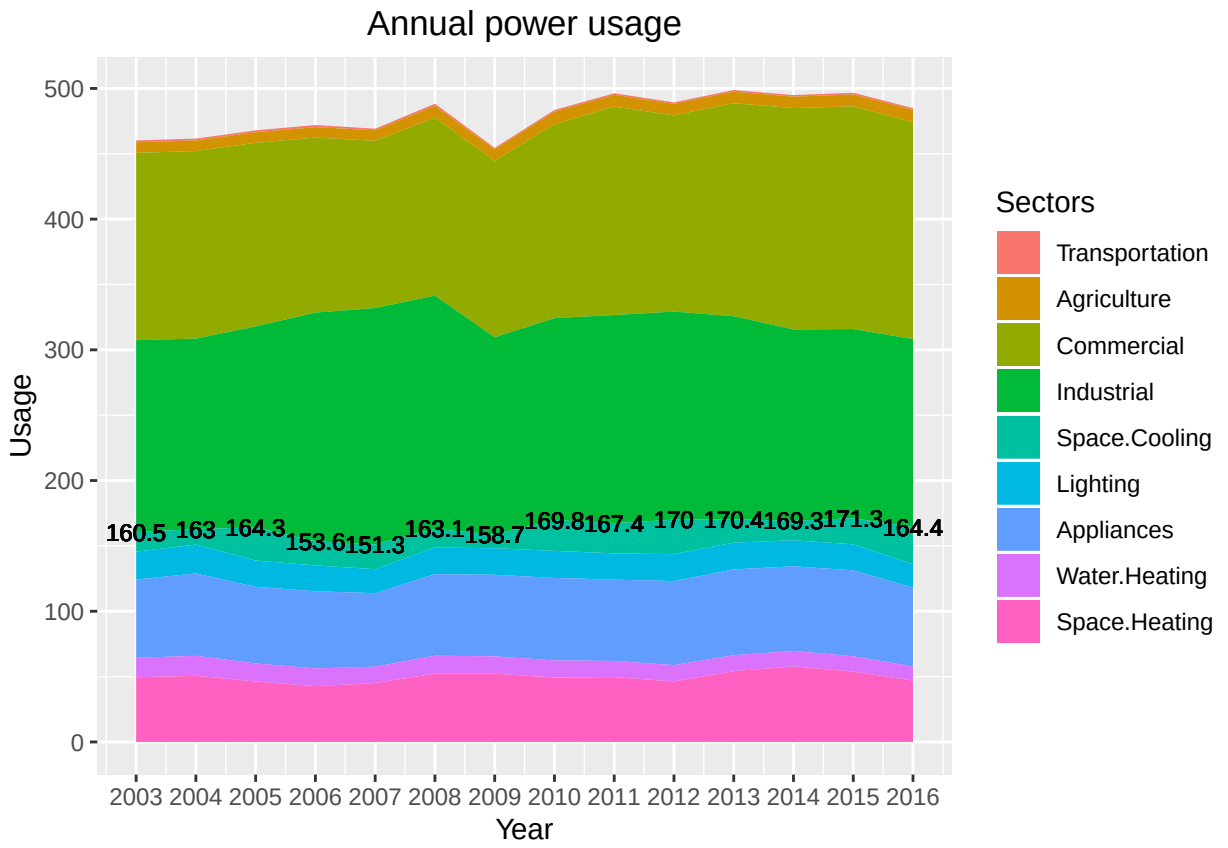
EDA, power forecasting

Boxi Lin

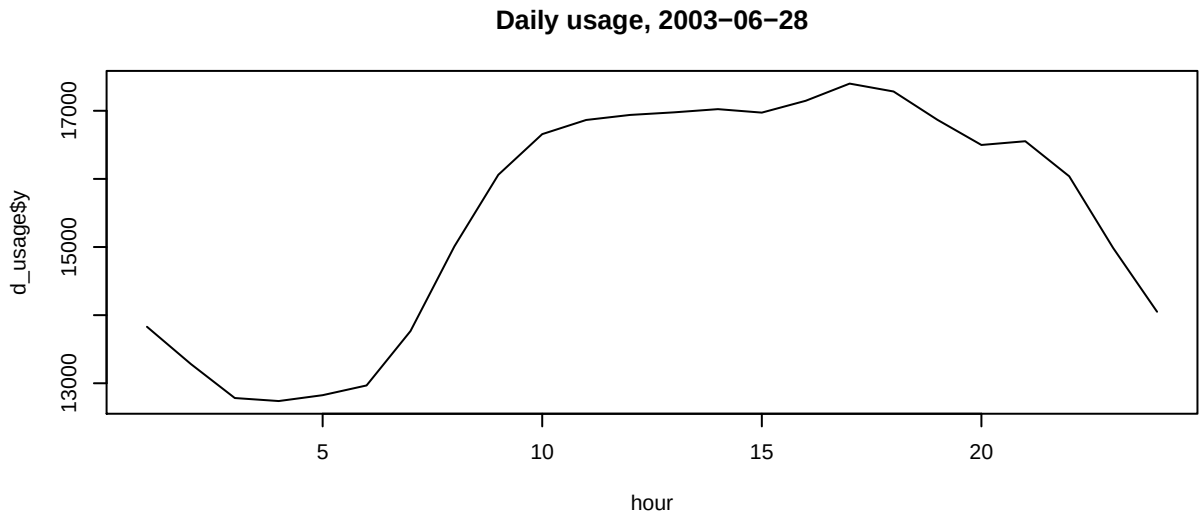
2020-02-04

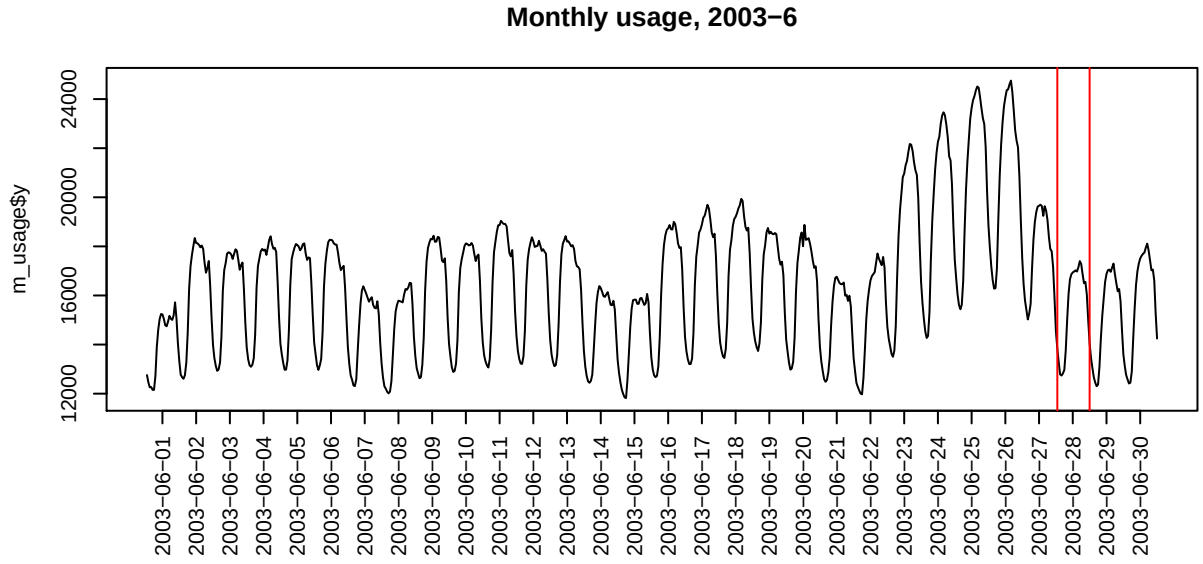
```
# Hourly: all sectors aggregated
#         (residential + industrial + commercial/institutional + agriculture + transportation)
h_usage <-read_excel("data/SSC2020_hourly_demand.xlsx",2)

# Annualy:  (residential: s.heat, w.heat, appliance, light, s.cooling
#            industrial, commercial/institutional, agriculture, transportation)
y_usage <-read.csv("data/ssc2020_annual_demand.csv")

# Hourly:
h_weather <- read_excel("data/ssc2020_hourly_weather.xlsx",2)
names(h_usage)<-c("date", "hour","y", "year","month")
n_year = 14
```

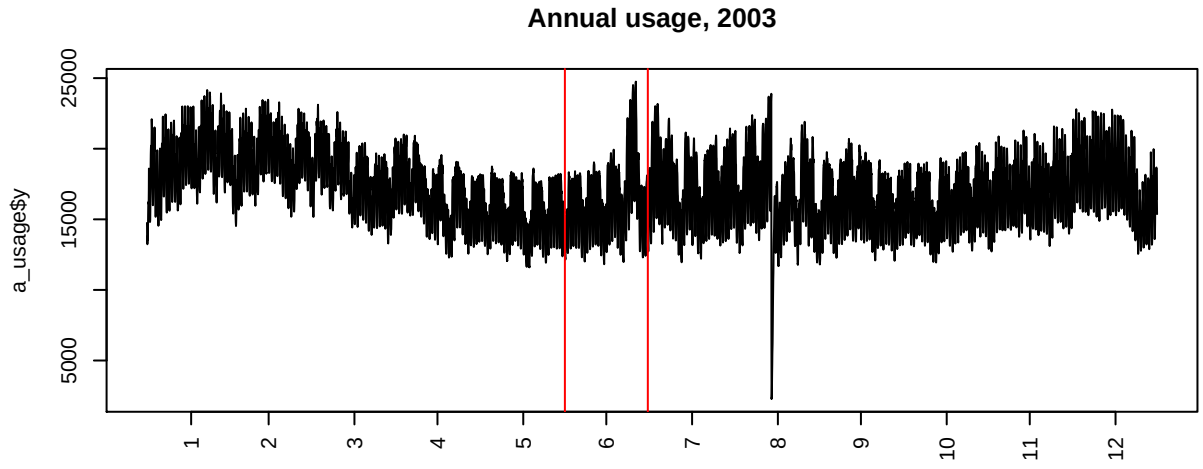
```
### Annual power usage

y_stack <- tidyr::gather(y_usage,"Sectors", "Usage",3:11) %>%
  mutate(Sectors = factor(Sectors, levels = rev(unique(Sectors))))
ggplot(y_stack, aes(x=Year, y=Usage, fill=Sectors)) +
  geom_area()+
  geom_text(data = y_stack, aes(x = Year, y = Residential), label=y_stack$Residential,size = 3)+
  scale_x_continuous("Year", labels = as.character(2003:2016), breaks = 2003:2016)+
  ggtitle("Annual power usage")+
  theme(plot.title = element_text(hjust = 0.5))
```




```
### Hourly usage data
par(mfrow = c(3,1))
d_usage <- h_usage %>% filter(date==pull(h_usage, date)[4290])
plot(d_usage$hour,d_usage$y, type = "l", main = "Daily usage, 2003-06-28", xlab = "hour")
m_usage <- h_usage %>% filter(year==d_usage$year & month==d_usage$month)
plot(1:720,m_usage$y, type = "l", main = "Monthly usage, 2003-6", xlab = "",xaxt="n")
axis(1, at= seq(0,700,24)+12,labels=m_usage$date[seq(0,700,24)+12], col.axis="black", las = 2)
abline(v = c(649,672), col = "red")
a_usage <- h_usage %>% filter(year==d_usage$year)
plot(1:8760, a_usage$y, type = "l", main = "Annual usage, 2003", xlab = "",xaxt="n")
axis(1, at=  cumsum(as.numeric(table(a_usage$month)))-15*24,labels=1:12, col.axis="black", las = 2)
abline(v = c(3625,4344), col = "red")
```

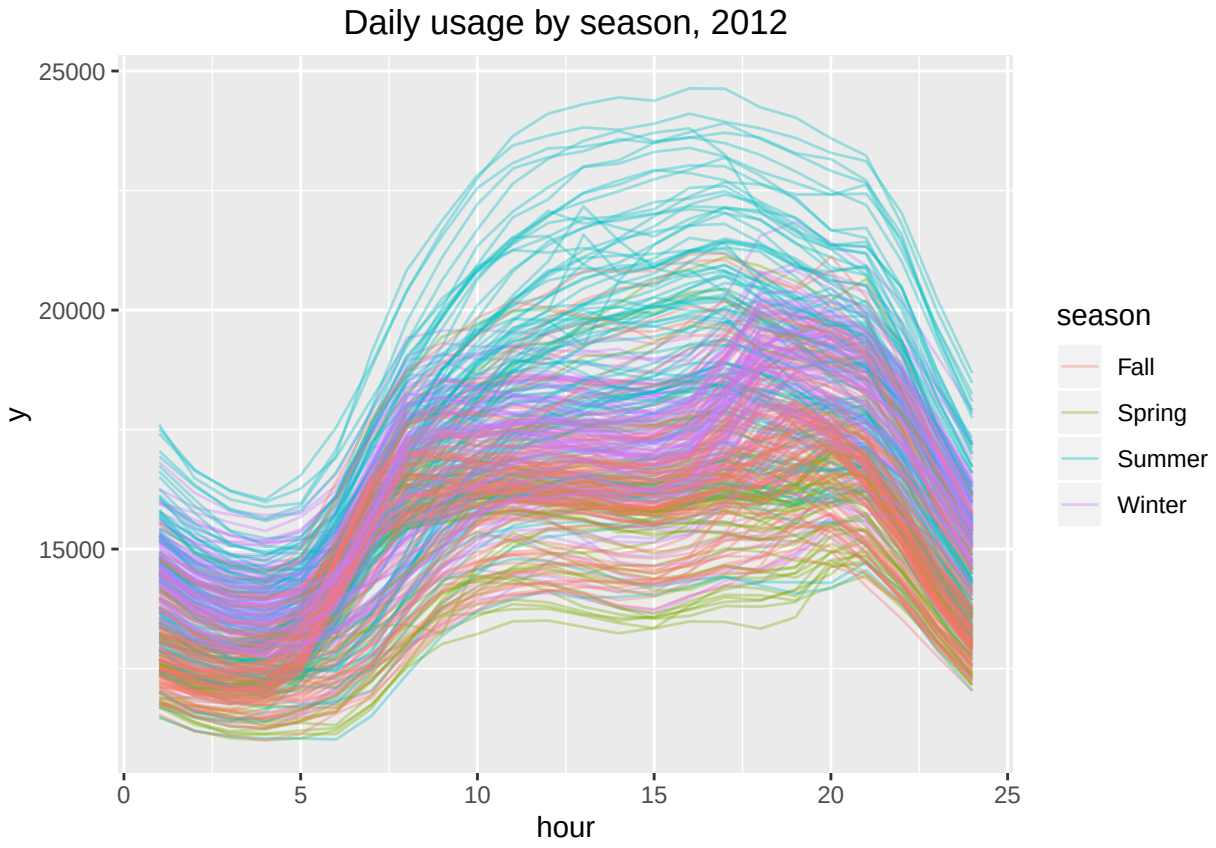

Daily usage, 2003−06−28
d_usage$y
13000
15000
17000
5
10
15
20
hour


Monthly usage, 2003−6
m_usage$y
12000
16000
20000
24000
2003−06−01
2003−06−02
2003−06−03
2003−06−04
2003−06−05
2003−06−06
2003−06−07
2003−06−08
2003−06−09
2003−06−10
2003−06−11
2003−06−12
2003−06−13
2003−06−14
2003−06−15
2003−06−16
2003−06−17
2003−06−18
2003−06−19
2003−06−20
2003−06−21
2003−06−22
2003−06−23
2003−06−24
2003−06−25
2003−06−26
2003−06−27
2003−06−28
2003−06−29
2003−06−30

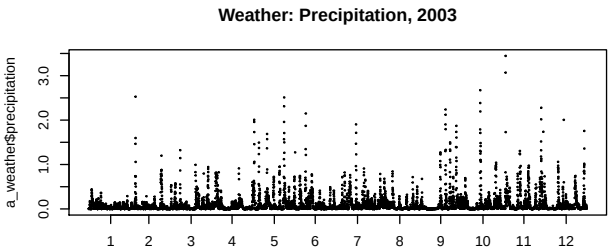

Annual usage, 2003
a_usage$y
5000
15000
25000
1
2
3
4
5
6
7
8
9
10
11
12

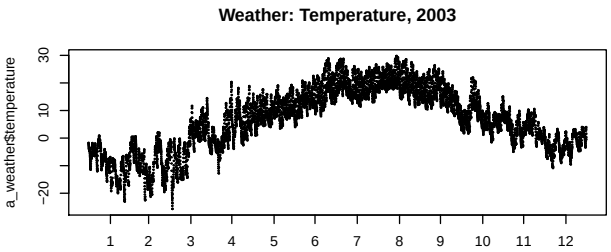
```
### Daily usage
```

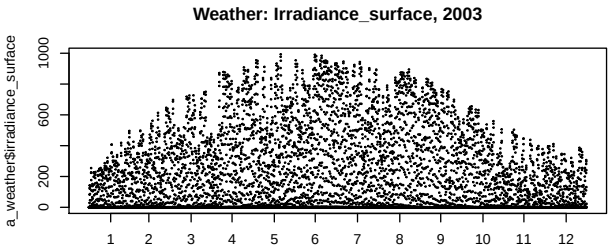
```
h_usage <- h_usage %>% mutate(season = ifelse(month %in% 9:11, "Fall", ifelse(month %in% 6:8, "Summer",
ggplot(data=h_usage[h_usage$year=="2012",],aes(hour,y,color=season,group=date)) +
  geom_line(alpha=0.35)+ ggtitle("Daily usage by season, 2012")+
  theme(plot.title = element_text(hjust = 0.5))
```

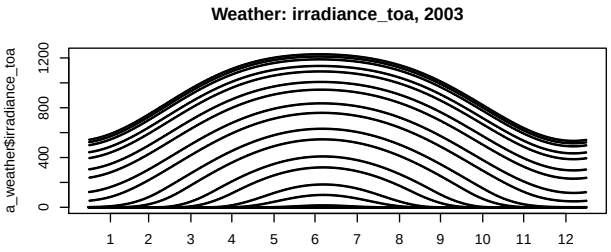


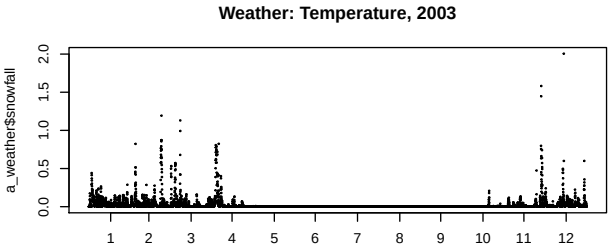

```
### Hourly Weather
par(mfrow = c(4,2))
a_weather <- h_weather %>% filter( year(h_weather$time)==d_usage$year)
plot(1:8760, a_weather$precipitation, cex = .2, main = "Weather: Precipitation, 2003", xlab = "",xaxt="
axis(1, at=  cumsum(as.numeric(table(a_usage$month)))-15*24,labels=1:12, col.axis="black")
plot(1:8760, a_weather$temperature, cex = .2, main = "Weather: Temperature, 2003", xlab = "",xaxt="n")
axis(1, at=  cumsum(as.numeric(table(a_usage$month)))-15*24,labels=1:12, col.axis="black")
plot(1:8760, a_weather$irradiance_surface, cex = .2, main = "Weather: Irradiance_surface, 2003", xlab =
axis(1, at=  cumsum(as.numeric(table(a_usage$month)))-15*24,labels=1:12, col.axis="black")
plot(1:8760, a_weather$irradiance_toa, cex= .2, main = "Weather: irradiance_toa, 2003", xlab = "",xaxt=
axis(1, at=  cumsum(as.numeric(table(a_usage$month)))-15*24,labels=1:12, col.axis="black")

plot(1:8760, a_weather$snowfall, cex = .2, main = "Weather: Temperature, 2003", xlab = "",xaxt="n")
axis(1, at=  cumsum(as.numeric(table(a_usage$month)))-15*24,labels=1:12, col.axis="black")
plot(1:8760, a_weather$snow_depth, cex = .2, main = "Weather: snowfall, 2003", xlab = "",xaxt="n")
axis(1, at=  cumsum(as.numeric(table(a_usage$month)))-15*24,labels=1:12, col.axis="black")
plot(1:8760, a_weather$cloud_cover, cex = .2, main = "Weather: cloud_cover, 2003", xlab = "",xaxt="n")
axis(1, at=  cumsum(as.numeric(table(a_usage$month)))-15*24,labels=1:12, col.axis="black")
plot(1:8760, a_weather$air_density, cex = .2, main = "Weather: air_density, 2003", xlab = "",xaxt="n")
axis(1, at=  cumsum(as.numeric(table(a_usage$month)))-15*24,labels=1:12, col.axis="black")
```

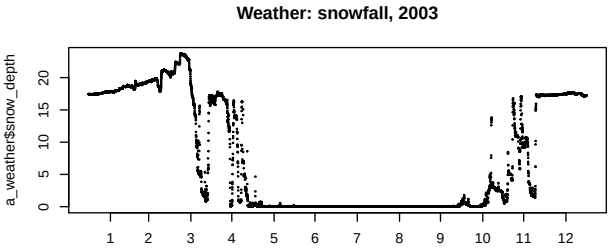

Weather: Precipitation, 2003
a_weather$precipitation
0.0
1.0
2.0
3.0
1 2 3 4 5 6 7 8 9 10 11 12

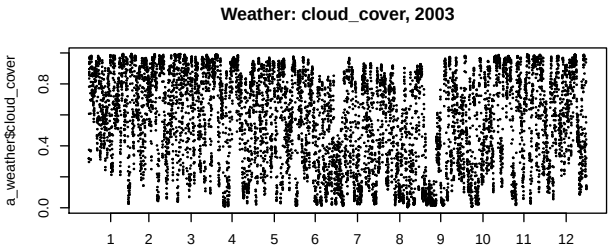

Weather: Temperature, 2003
a_weather$temperature
−20
0
10
30
1 2 3 4 5 6 7 8 9 10 11 12

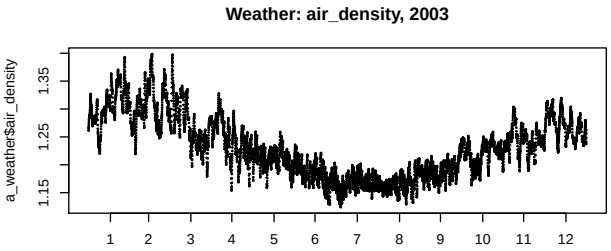

Weather: Irradiance_surface, 2003
a_weather$irradiance_surface
0
200
600
1000
1 2 3 4 5 6 7 8 9 10 11 12


Weather: irradiance_toa, 2003
a_weather$irradiance_toa
0
400
800
1200
1 2 3 4 5 6 7 8 9 10 11 12


Weather: Temperature, 2003
a_weather$snowfall
0.0
0.5
1.0
1.5
2.0
1 2 3 4 5 6 7 8 9 10 11 12


Weather: snowfall, 2003
a_weather$snow_depth
0
5
10
15
20
1 2 3 4 5 6 7 8 9 10 11 12


Weather: cloud_cover, 2003
a_weather$cloud_cover
0.0
0.4
0.8
1 2 3 4 5 6 7 8 9 10 11 12


Weather: air_density, 2003
a_weather$air_density
1.15
1.25
1.35
1 2 3 4 5 6 7 8 9 10 11 12