Homework #4

Zhejin Dong(zd2221)

11/14/2018

```
Anually_data<-read.csv('/Users/zhejindong/Downloads/__public__data__castnet__download__Castnet_4B31B9F8
Season_data<-read.csv("/Users/zhejindong/Downloads/__public__data__castnet__download__Castnet_4B31B9F865
Week_data<-read.csv("/Users/zhejindong/Downloads/__public__data__castnet__download__Castnet_4B31B9F8659
```

```
library(tidyverse)
```

```
## -- Attaching packages ----------------------------------------------------------------------------- t
```

```
## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  1.4.2     v dplyr   0.7.8
## v tidyr   0.8.1     v stringr 1.3.1
## v readr   1.1.1     v forcats 0.3.0
```

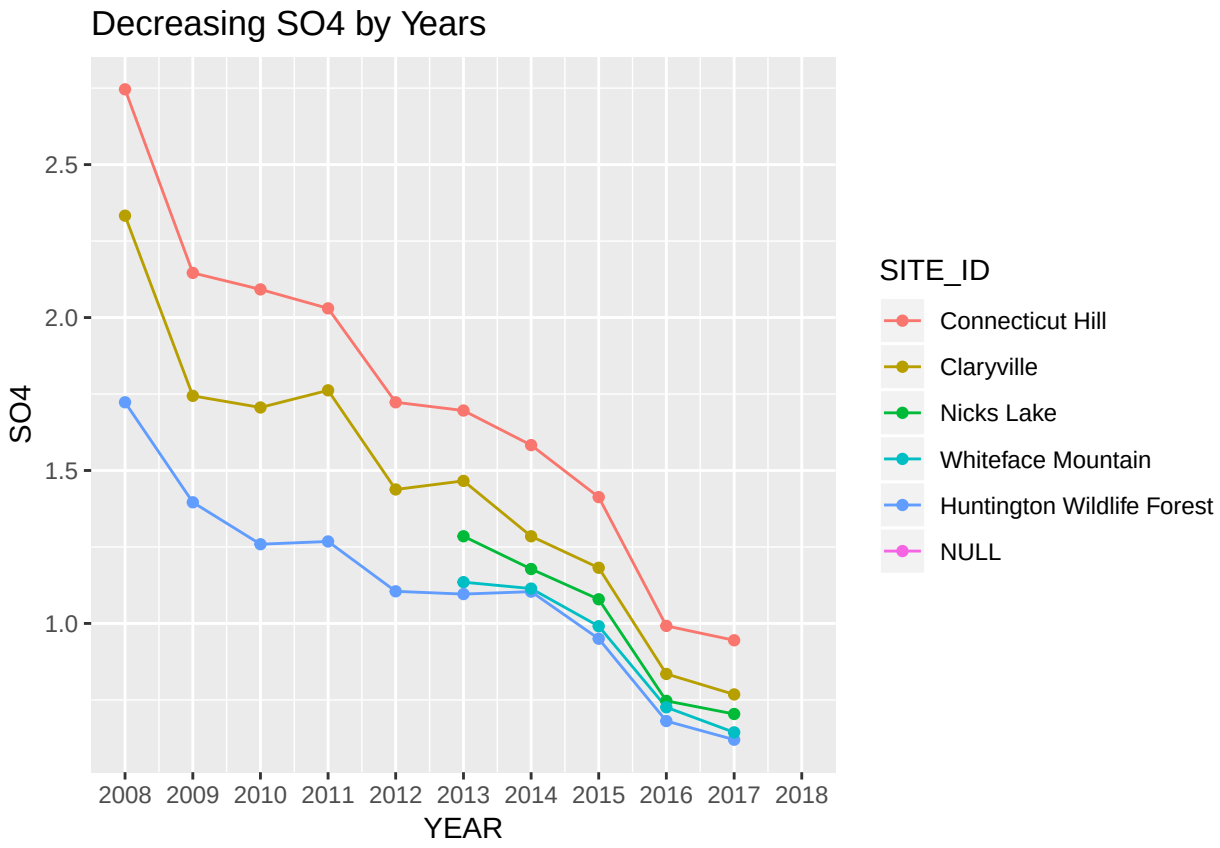
```
## -- Conflicts -------------------------------------------------------------------------------- tidyver
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
Anually_data$YEAR<-as.character(Anually_data$YEAR)
Anually_data$YEAR<-paste0(Anually_data$YEAR,"-01-01")
Anually_data$YEAR<-as.Date(Anually_data$YEAR)
data1<-filter(Anually_data,YEAR>='2008-01-01')
data1$SITE_ID<-factor(data1$SITE_ID,levels=c('CTH110','CAT175','NIC001','WFM105','HWF187','WFM007'),
                      labels=c('Connecticut Hill','Claryville','Nicks Lake','Whiteface Mountain','Hunti
```

```
ggplot(data1,aes(x=YEAR,y=data1$SO4_CONC,group=data1$SITE_ID,color=SITE_ID)
       )+geom_line()+scale_x_date(date_labels = "%Y",date_breaks = '1 year'
                                  )+geom_point()+ggtitle('Decreasing SO4 by Years')+ylab('SO4')
```

```
## Warning: Removed 10 rows containing missing values (geom_path).
```

```
## Warning: Removed 10 rows containing missing values (geom_point).
```



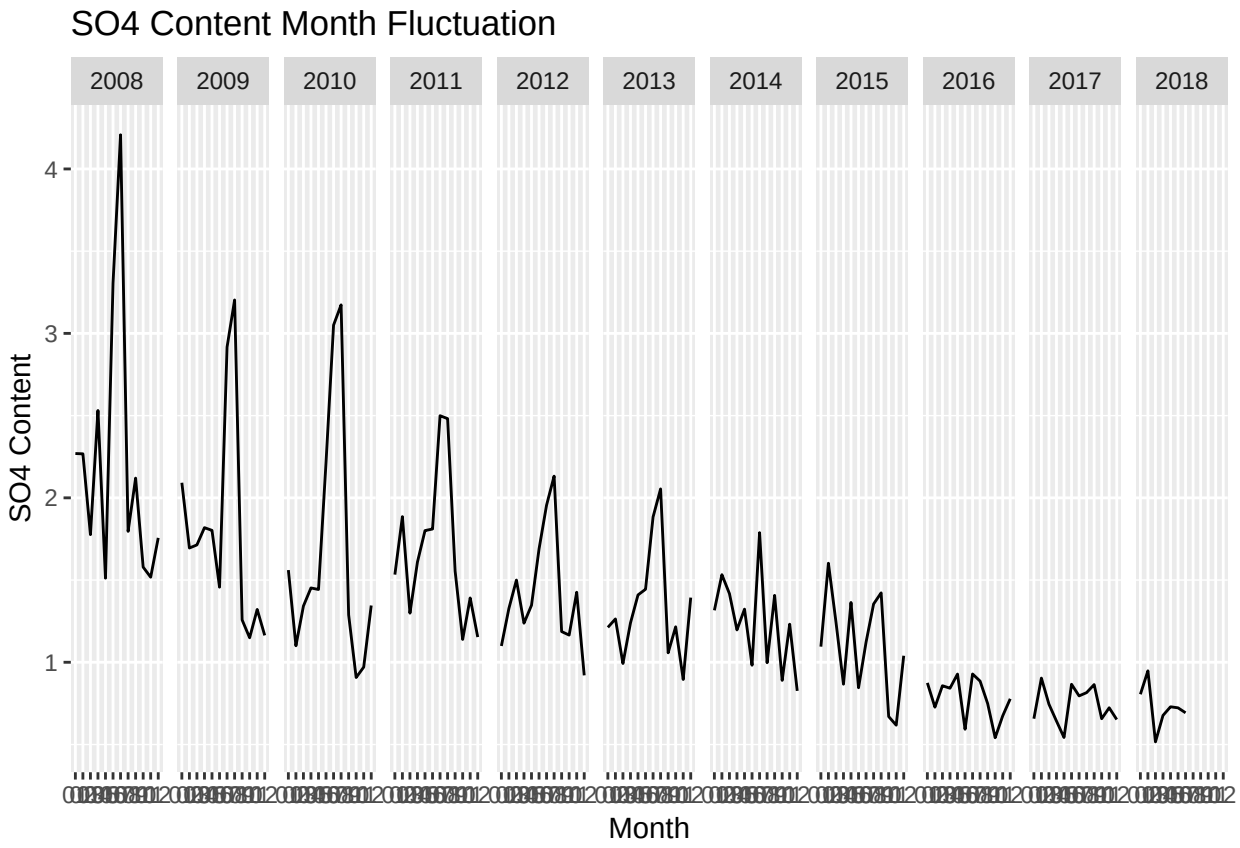
The picture above shows that SO4 content in 5 districts in NYC decreases by years from 2009 to 2017.

```
Week_data$DATE<-substr(Week_data$DATEON, 1, 10)
Week_data$DATE<-as.Date(Week_data$DATE,'%m/%d/%Y')
Week_data$Month<-substr(Week_data$DATE,6,7)
Week_data$YEAR<-factor(Week_data$YEAR)
```

```
data2<-select(Week_data,Month,YEAR,SO4_CONC)
data2<-filter(data2,SO4_CONC!='NA')

data3<-data2 %>% group_by(Month,YEAR) %>% summarise_at(vars(SO4_CONC),mean)

ggplot(data3,aes(x=Month,y=SO4_CONC,group=1))+geom_line()+facet_grid(~YEAR)+ggtitle('SO4 Content Month
```

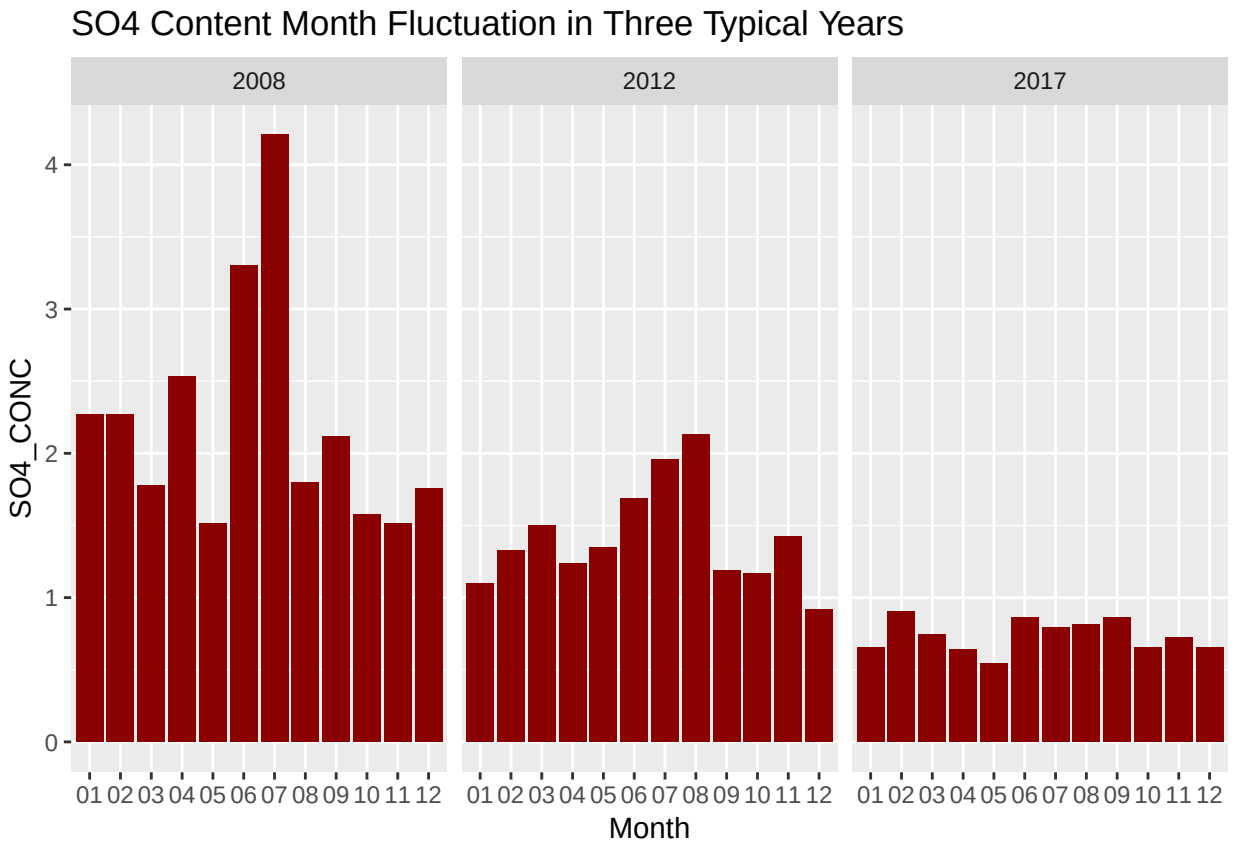


From the picture above, we can found two changes from 2008 to 2018:

1 The total annually SO4 content is decreasing.

2 The variance of SO4 in a year is decreasing.

```
data7<-filter(data2,YEAR %in% c('2008','2012','2017'))
data6<-data7 %>% group_by(Month,YEAR) %>% summarise_at(vars(SO4_CONC),mean)
ggplot(data6, aes(Month,SO4_CONC)) + geom_col(fill='darkred')+ xlab("SO4")+facet_wrap(YEAR~.)+xlab('Mon
```



I selected three typical years to display in detail the fluctuation of SO4 in 12 months. I found, before 2012, SO4 content is quite high in from June to July, but this situation weakens by years.