

Assignment5-part1

LingzhouAo

4/27/2018

Data Clean

```
library(stringr)
suppressMessages(library(tidyverse))
library(shiny)
library(shinydashboard)

## Warning: package 'shinydashboard' was built under R version 3.4.4
##
## Attaching package: 'shinydashboard'
## The following object is masked from 'package:graphics':
##
##      box

library(dplyr)
library(ggplot2)

time_seq <- seq(1988,2017)[-26]
for (i in 1:length(time_seq)){
  filename<-paste("weather data/",time_seq[i],".txt",sep = "")
  tex<-read.table(filename,header = TRUE, fill = TRUE)
  if(length(tex[1,])==16|length(tex[1,])==17){
    sub_tex<-tex[,c(1,2,3,4,13,14)]
  }
  if(length(tex[1,])==18){
    sub_tex<-tex[,c(1,2,3,4,14,15)]
  }
  names(sub_tex)<-c('YY','MM','DD','hh','ATMP','WTMP')
  if (i == 1){
    data <- sub_tex
  }
  else{
    data<-rbind(data,sub_tex)
  }
}

temperature <- data %>%
  filter(hh==12)

temperature[temperature$YY < 99,]$YY <-temperature[temperature$YY < 99,]$YY + 1900
temperature <- temperature %>%
  filter(ATMP<99) %>%
  filter(WTMP<99)
temperature$ATMP <- as.numeric(temperature$ATMP)
temperature$WTMP <- as.numeric(temperature$WTMP)
```

```
temperature$time<-paste(temperature$YY,"-",temperature$MM,"-",temperature$DD,sep = "")
temperature$time<-as.Date(temperature$time)
```

T-test

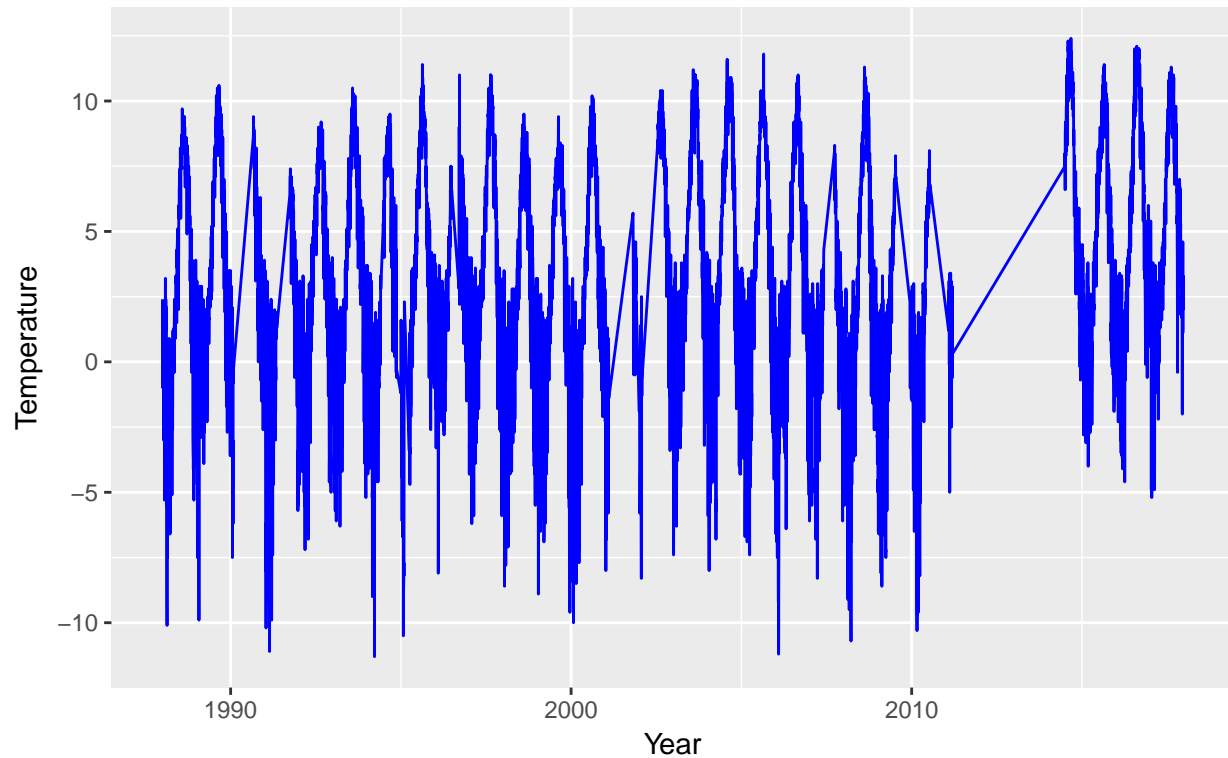
```
year_1988<-temperature[temperature$YY==1988,]
year_2017<-temperature[temperature$YY==2017,]
A_1988<-sample(year_1988$ATMP,size=200)
A_2017<-sample(year_2017$ATMP,size=200)
W_1988<-sample(year_1988$WTMP,size=200)
W_2017<-sample(year_2017$WTMP,size=200)
test_A<-t.test(A_1988,A_2017)
test_W<-t.test(W_1988,W_2017)
```

Air Temperature

```
graph <- temperature %>% ggplot(aes(time, ATMP)) +
  geom_line(na.rm = TRUE, col = "blue") +
  labs(title = "Time Series of Air Temperature",
        subtitle = "Data obtained from the National Data Buoy Center",
        y = "Temperature",
        x = "Year")
plot(graph)
```

Time Series of Air Temperature

Data obtained from the National Data Buoy Center

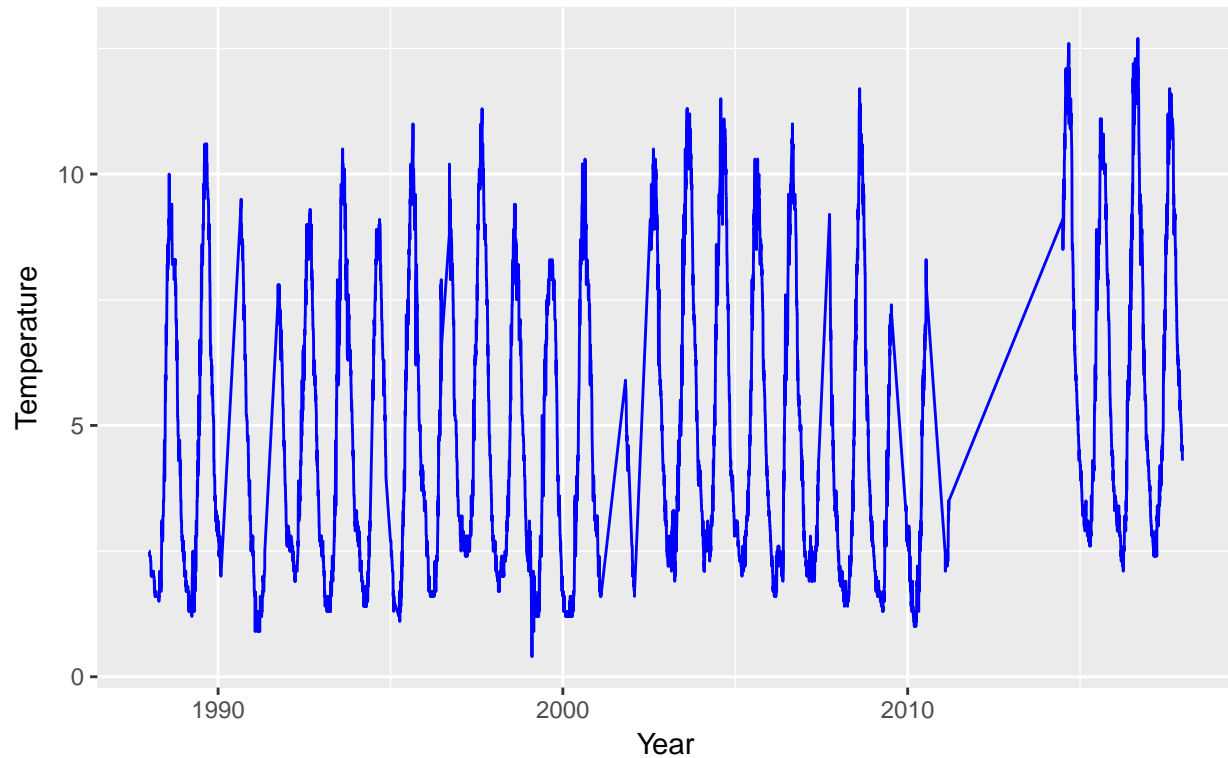


Sea Temperature

```
graph <- temperature %>% ggplot(aes(time, WTMP)) +  
  geom_line(na.rm = TRUE, col = "blue") +  
  labs(title = "Time Series of Sea Temperature",  
        subtitle = "Data obtained from the National Data Buoy Center",  
        y = "Temperature",  
        x = "Year")  
plot(graph)
```

Time Series of Sea Temperature

Data obtained from the National Data Buoy Center

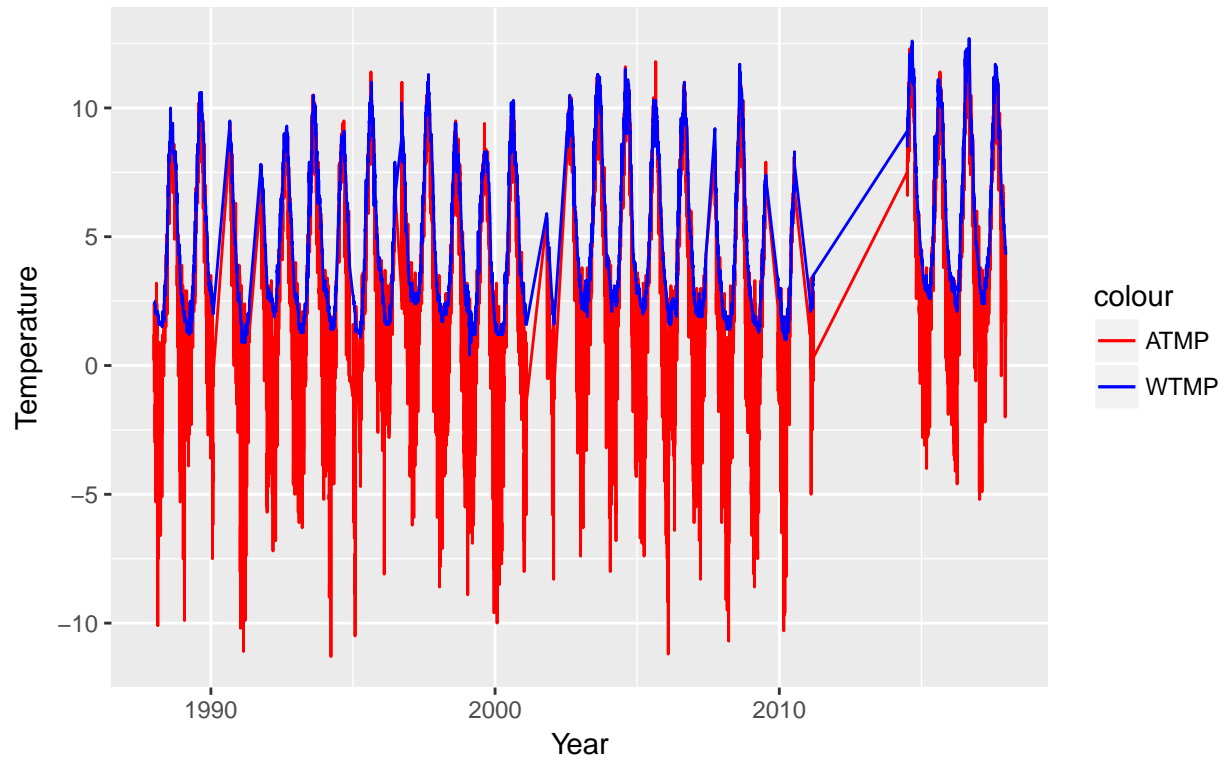


Air and Sea Temperatures

```
graph <- ggplot(temperature, aes(time)) +  
  geom_line(aes(y = ATMP, col = "ATMP")) +  
  geom_line(aes(y = WTMP, col = "WTMP")) +  
  scale_colour_manual(values=c("red", "blue")) +  
  labs(x = "Year", y = "Temperature",  
        title = "Time Series of Air & Sea Temperature",  
        subtitle = "Data obtained from the National Data Buoy Center")  
plot(graph)
```

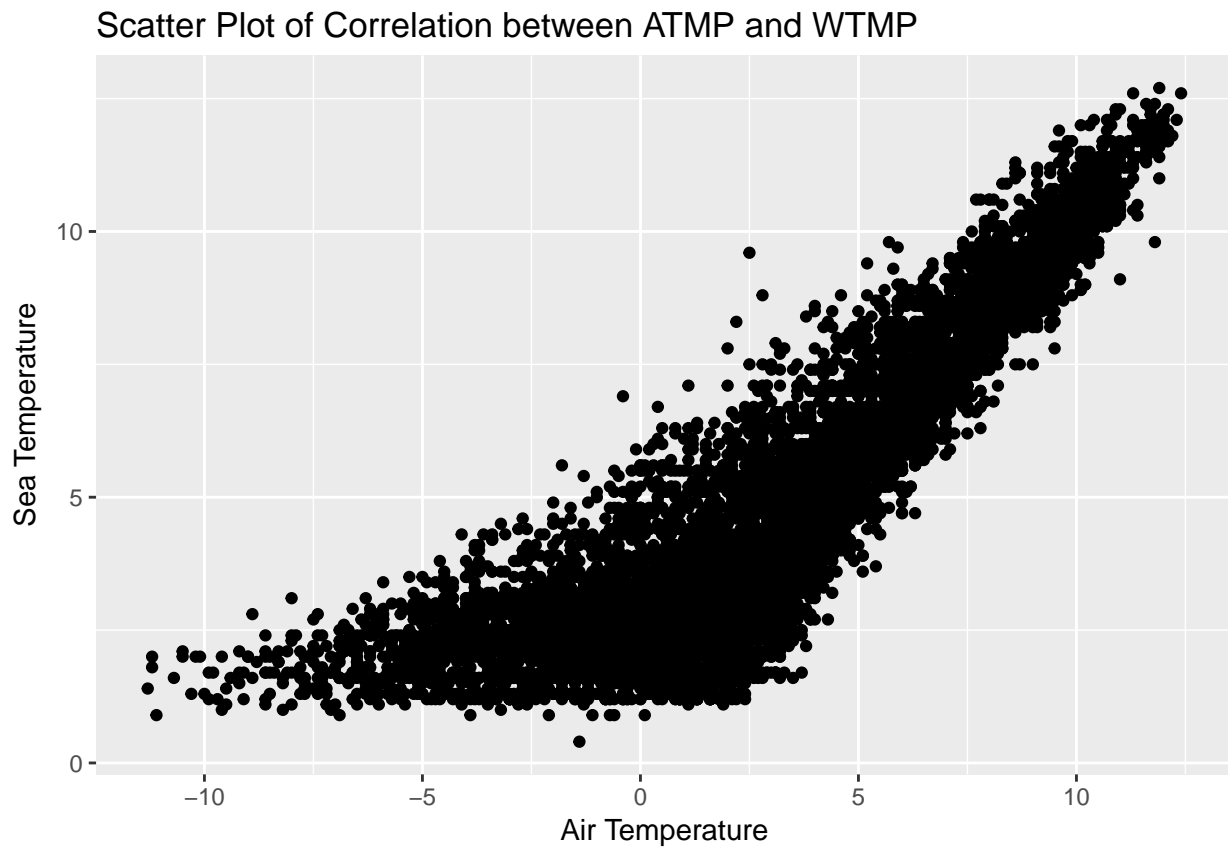
Time Series of Air & Sea Temperature

Data obtained from the National Data Buoy Center



Scatter Plot

```
graph <- ggplot(temperature) +  
  geom_point(mapping = aes(x = ATMP, y = WTMP)) +  
  labs(x = "Air Temperature",  
       y = "Sea Temperature",  
       title = "Scatter Plot of Correlation between ATMP and WTMP")  
plot(graph)
```



Smooth Line

```
graph <- ggplot(temperature) +  
  geom_smooth(mapping = aes(x = ATMP, y = WTMP)) +  
  labs(x = "Air Temperature",  
       y = "Sea Temperature",  
       title = "Smooth Line of Correlation between ATMP and WTMP")  
plot(graph)
```

```
## `geom_smooth()` using method = 'gam'
```

Smooth Line of Correlation between ATMP and WTMP

