Exploratory Data Analysis Project 1

Ying Zhou

8/4/2022

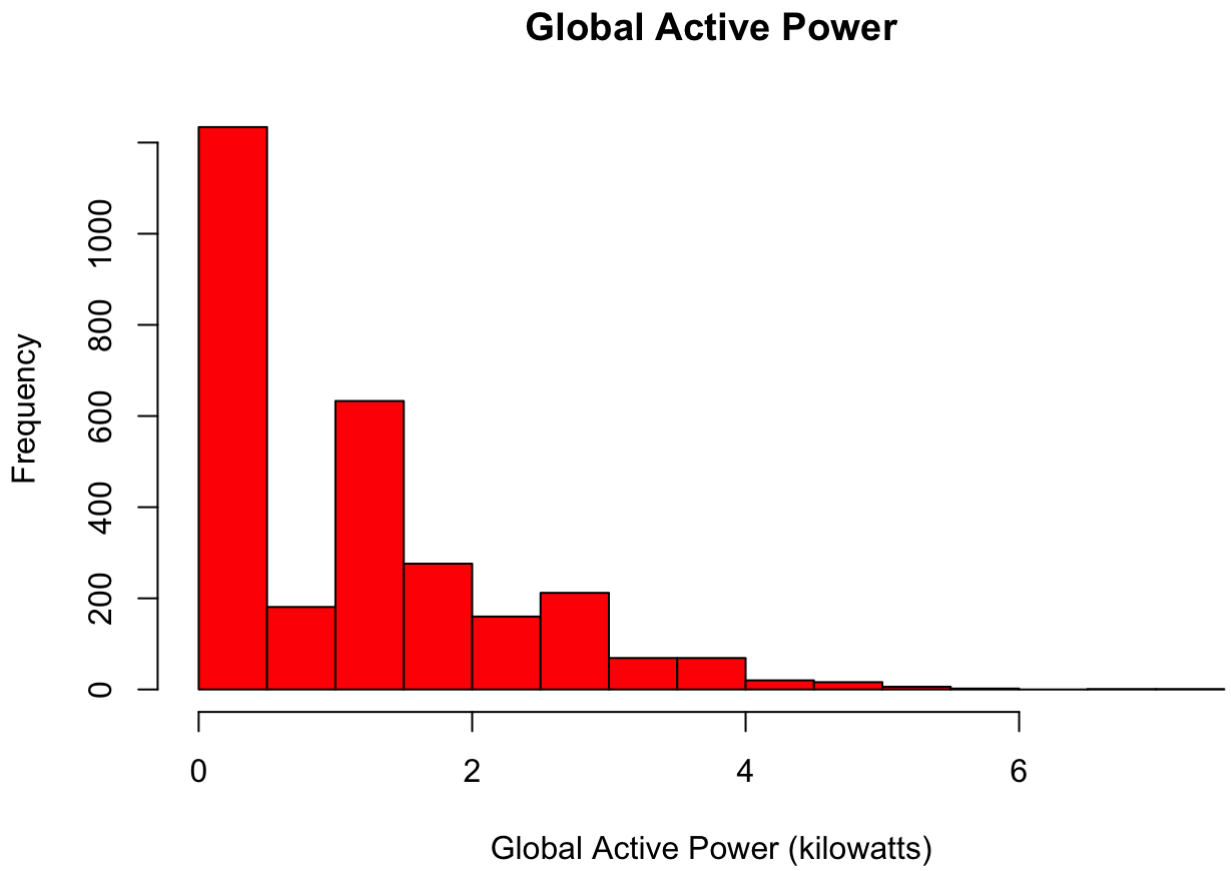
Project 1

```
setwd("~/Desktop/r")
# read table
t <- read.table("household_power_consumption.txt",
                header=TRUE, sep=";", na.strings = "?",
                colClasses = c('character','character','numeric',
                'numeric','numeric','numeric','numeric','numeric','numeric'))

# Format date
t$Date <- as.Date(t$Date, "%d/%m/%Y")

# Filter data set from Feb. 1, 2007 to Feb. 2, 2007
t <- subset(t,Date >= as.Date("2007-2-1") & Date <= as.Date("2007-2-2"))

# Remove incomplete observation
t <- t[complete.cases(t),]

# Combine Date and Time column
dateTime <- paste(t$Date, t$Time)

# Name the vector
dateTime <- setNames(dateTime, "DateTime")

# Remove Date and Time column
t <- t[ ,!(names(t) %in% c("Date","Time"))]

# Add DateTime column
t <- cbind(dateTime, t)

# Format dateTime Column
t$dateTime <- as.POSIXct(dateTime)
```

Create the histogram

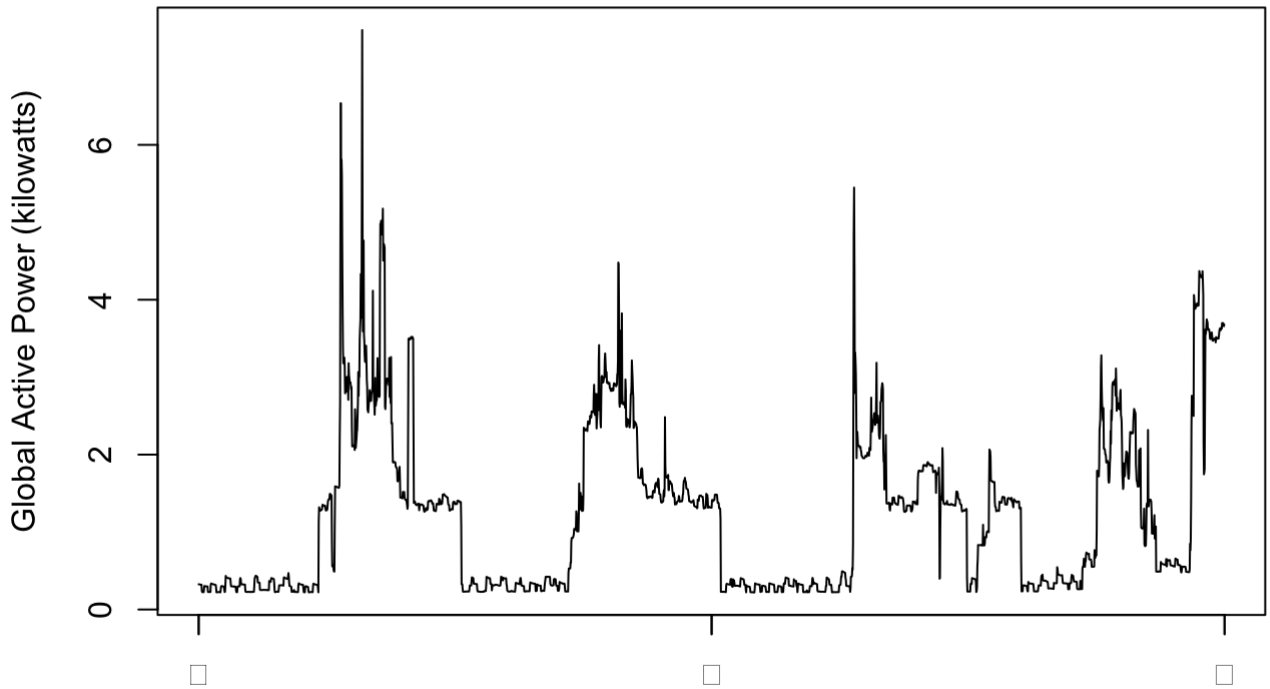
Plot 1

```
hist(t$Global_active_power, main="Global Active Power",
     xlab = "Global Active Power (kilowatts)", col="red")
```



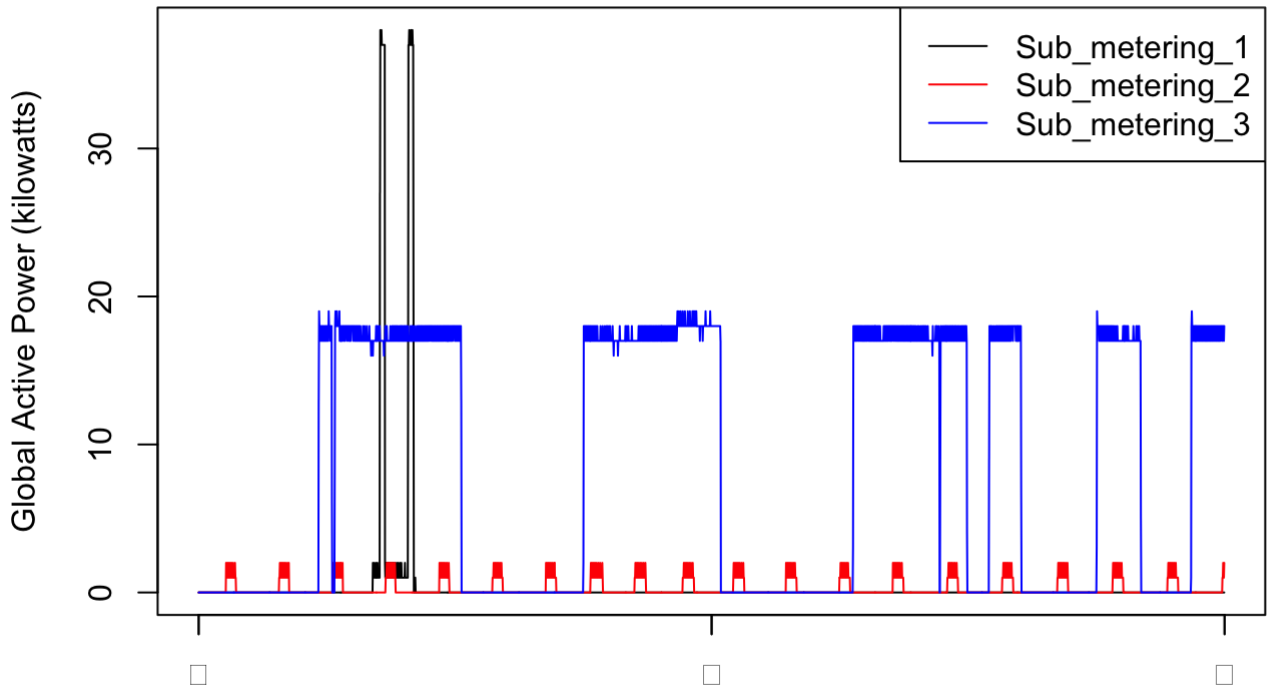
Plot 2

```
plot(t$Global_active_power~t$dateTime, type="l",
     ylab="Global Active Power (kilowatts)", xlab="")
```



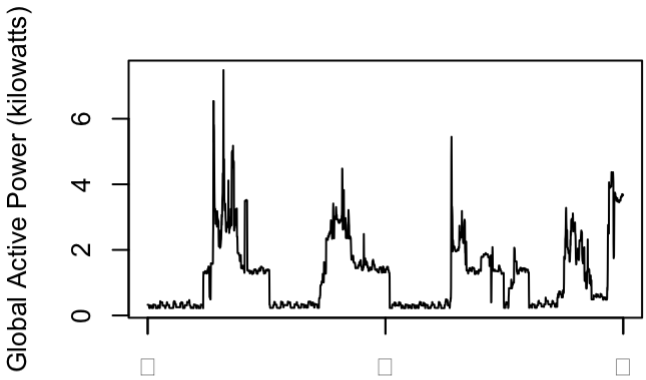
Plot 3

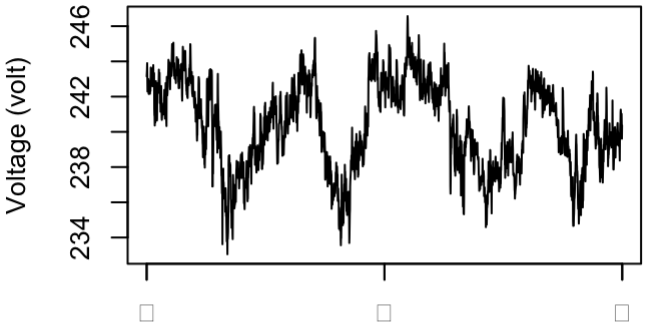
```
with(t, {
    plot(Sub_metering_1~dateTime, type="l",
         ylab="Global Active Power (kilowatts)", xlab="")
    lines(Sub_metering_2~dateTime,col='Red')
    lines(Sub_metering_3~dateTime,col='Blue')
  })
  legend("topright", col=c("black", "red", "blue"), lwd=c(1,1,1),
         c("Sub_metering_1", "Sub_metering_2", "Sub_metering_3"))
```

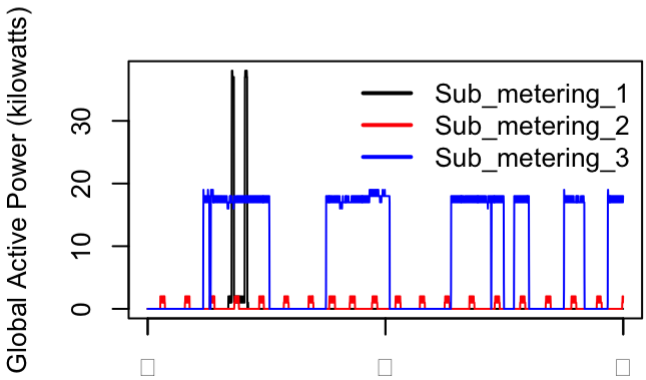


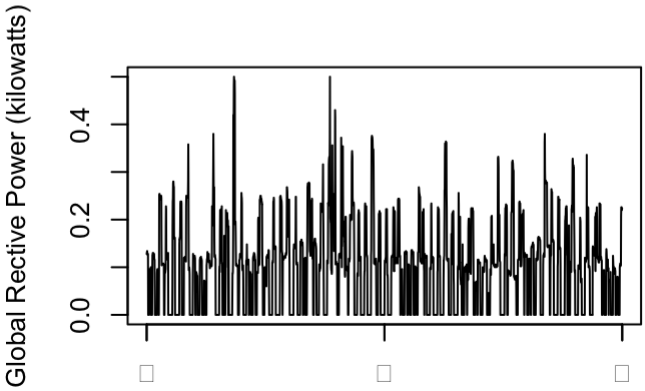
Plot 4

```
par(mfrow=c(2,2), mar=c(4,4,2,1), oma=c(0,0,2,0))
with(t, {
    plot(Global_active_power~dateTime, type="l",
         ylab="Global Active Power (kilowatts)", xlab="")
    plot(Voltage~dateTime, type="l",
         ylab="Voltage (volt)", xlab="")
    plot(Sub_metering_1~dateTime, type="l",
         ylab="Global Active Power (kilowatts)", xlab="")
    lines(Sub_metering_2~dateTime,col='Red')
    lines(Sub_metering_3~dateTime,col='Blue')
    legend("topright", col=c("black", "red", "blue"), lty=1, lwd=2, bty="n",
           legend=c("Sub_metering_1", "Sub_metering_2", "Sub_metering_3"))
    plot(Global_reactive_power~dateTime, type="l",
         ylab="Global Rective Power (kilowatts)",xlab="")
  })
```


Global Active Power (kilowatts)
0
2
4
6


Voltage (volt)
234
238
242
246


Global Active Power (kilowatts)
0
10
20
30
Sub_metering_1
Sub_metering_2
Sub_metering_3


Global Rective Power (kilowatts)
0.0
0.2
0.4