Eletric power consumption

Matteo Gambera

```
url <- "https://d396qusza40orc.cloudfront.net/exdata%2Fdata%2Fhousehold_power_consumption.zip"
download.file(url, "/home/matteo/Scrivania/datasciencecoursera/Exploratory Data Analysis/Data/dataFiles
setwd("/home/matteo/Scrivania/datasciencecoursera/Exploratory Data Analysis/Data")
path = "/home/matteo/Scrivania/datasciencecoursera/Exploratory Data Analysis/Data"
unzip(zipfile = "dataFiles.zip")
```

```
powerData <- read.table(file.path(path, "household_power_consumption.txt"), sep=";", header=TRUE,na.str
head(powerData)
```

```
##         Date     Time Global_active_power Global_reactive_power Voltage
## 1 16/12/2006 17:24:00               4.216                 0.418  234.84
## 2 16/12/2006 17:25:00               5.360                 0.436  233.63
## 3 16/12/2006 17:26:00               5.374                 0.498  233.29
## 4 16/12/2006 17:27:00               5.388                 0.502  233.74
## 5 16/12/2006 17:28:00               3.666                 0.528  235.68
## 6 16/12/2006 17:29:00               3.520                 0.522  235.02
##   Global_intensity Sub_metering_1 Sub_metering_2 Sub_metering_3
## 1             18.4              0              1             17
## 2             23.0              0              1             16
## 3             23.0              0              2             17
## 4             23.0              0              1             17
## 5             15.8              0              1             17
## 6             15.0              0              2             17
```

```
colnames(powerData)[1:9] <- c("Date", "Time", "Active_power", "Reactive_power", "Voltage", "Intensity",
#powerData <- powerData[2:nrow(powerData),]
str(powerData)
```

```
## 'data.frame':    2075259 obs. of  9 variables:
##  $ Date          : chr  "16/12/2006" "16/12/2006" "16/12/2006" "16/12/2006" ...
##  $ Time          : chr  "17:24:00" "17:25:00" "17:26:00" "17:27:00" ...
##  $ Active_power  : num  4.22 5.36 5.37 5.39 3.67 ...
##  $ Reactive_power: num  0.418 0.436 0.498 0.502 0.528 0.522 0.52 0.52 0.51 0.51 ...
##  $ Voltage       : num  235 234 233 234 236 ...
##  $ Intensity     : num  18.4 23 23 23 15.8 15 15.8 15.8 15.8 15.8 ...
##  $ type1         : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ type2         : num  1 1 2 1 1 2 1 1 1 2 ...
##  $ type3         : num  17 16 17 17 17 17 17 17 17 16 ...
```

```
head(powerData)
```

```
##         Date     Time Active_power Reactive_power Voltage Intensity type1 type2
```

```
## 1 16/12/2006 17:24:00        4.216          0.418  234.84      18.4     0     1
## 2 16/12/2006 17:25:00        5.360          0.436  233.63      23.0     0     1
## 3 16/12/2006 17:26:00        5.374          0.498  233.29      23.0     0     2
## 4 16/12/2006 17:27:00        5.388          0.502  233.74      23.0     0     1
## 5 16/12/2006 17:28:00        3.666          0.528  235.68      15.8     0     1
## 6 16/12/2006 17:29:00        3.520          0.522  235.02      15.0     0     2
##   type3
## 1    17
## 2    16
## 3    17
## 4    17
## 5    17
## 6    17
```

```
tail(powerData)
```

```
##                Date     Time Active_power Reactive_power Voltage Intensity type1
## 2075254 26/11/2010 20:57:00        0.946              0  240.33       4.0     0
## 2075255 26/11/2010 20:58:00        0.946              0  240.43       4.0     0
## 2075256 26/11/2010 20:59:00        0.944              0  240.00       4.0     0
## 2075257 26/11/2010 21:00:00        0.938              0  239.82       3.8     0
## 2075258 26/11/2010 21:01:00        0.934              0  239.70       3.8     0
## 2075259 26/11/2010 21:02:00        0.932              0  239.55       3.8     0
##         type2 type3
## 2075254     0     0
## 2075255     0     0
## 2075256     0     0
## 2075257     0     0
## 2075258     0     0
## 2075259     0     0
```

```
powerData$Date <-as.Date(powerData$Date, "%d/%m/%Y")
str(powerData)
```

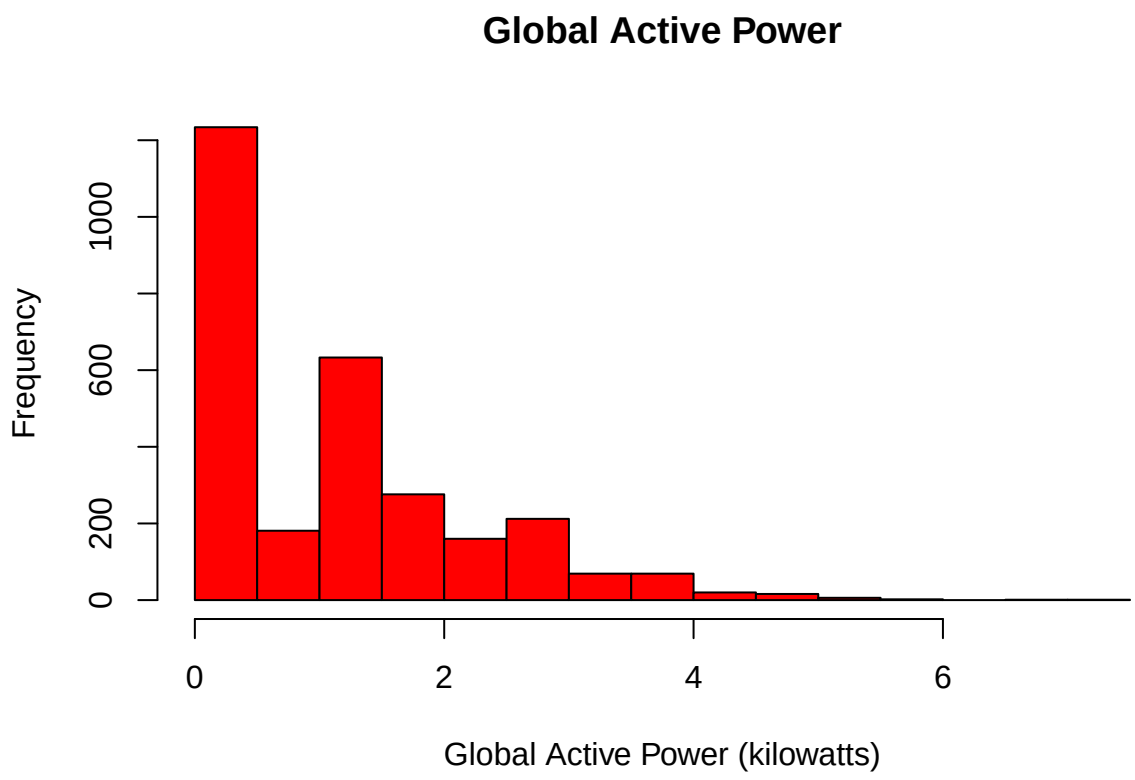
```
## 'data.frame':    2075259 obs. of  9 variables:
##  $ Date          : Date, format: "2006-12-16" "2006-12-16" ...
##  $ Time          : chr  "17:24:00" "17:25:00" "17:26:00" "17:27:00" ...
##  $ Active_power  : num  4.22 5.36 5.37 5.39 3.67 ...
##  $ Reactive_power: num  0.418 0.436 0.498 0.502 0.528 0.522 0.52 0.52 0.51 0.51 ...
##  $ Voltage       : num  235 234 233 234 236 ...
##  $ Intensity     : num  18.4 23 23 23 15.8 15 15.8 15.8 15.8 15.8 ...
##  $ type1         : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ type2         : num  1 1 2 1 1 2 1 1 1 2 ...
##  $ type3         : num  17 16 17 17 17 17 17 17 17 16 ...
```

```
t <- subset(powerData,Date >= as.Date("2007-2-1") & Date <= as.Date("2007-2-2"))
```

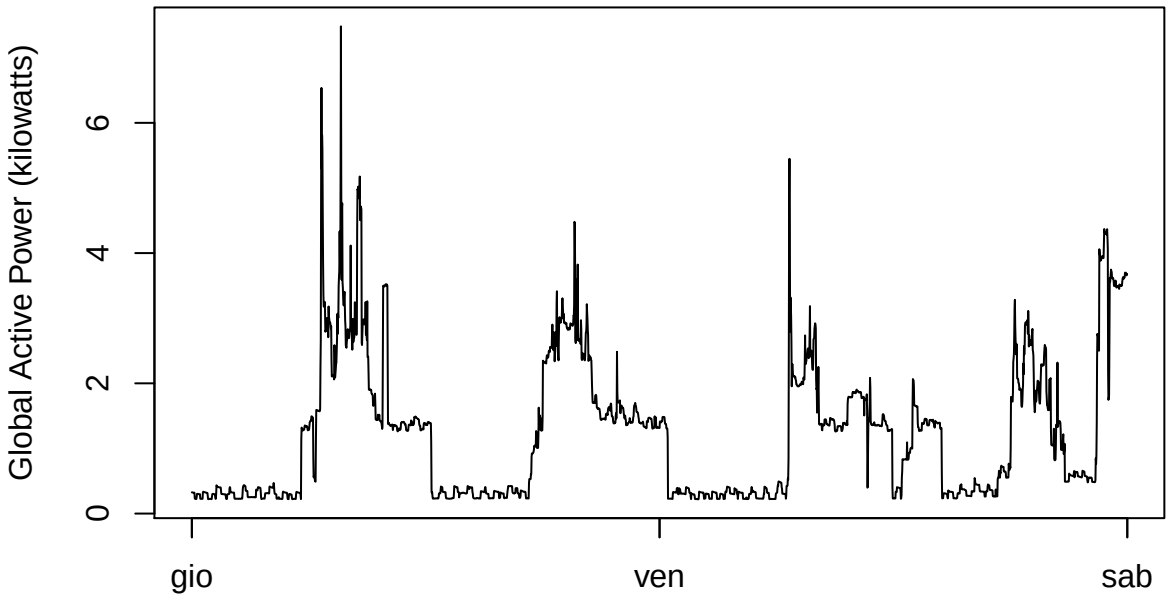
```
## Remove incomplete observation
t <- t[complete.cases(t),]
## Combine Date and Time column
dateTime <- paste(t$Date, t$Time)
## Name the vector
```

```
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)
```

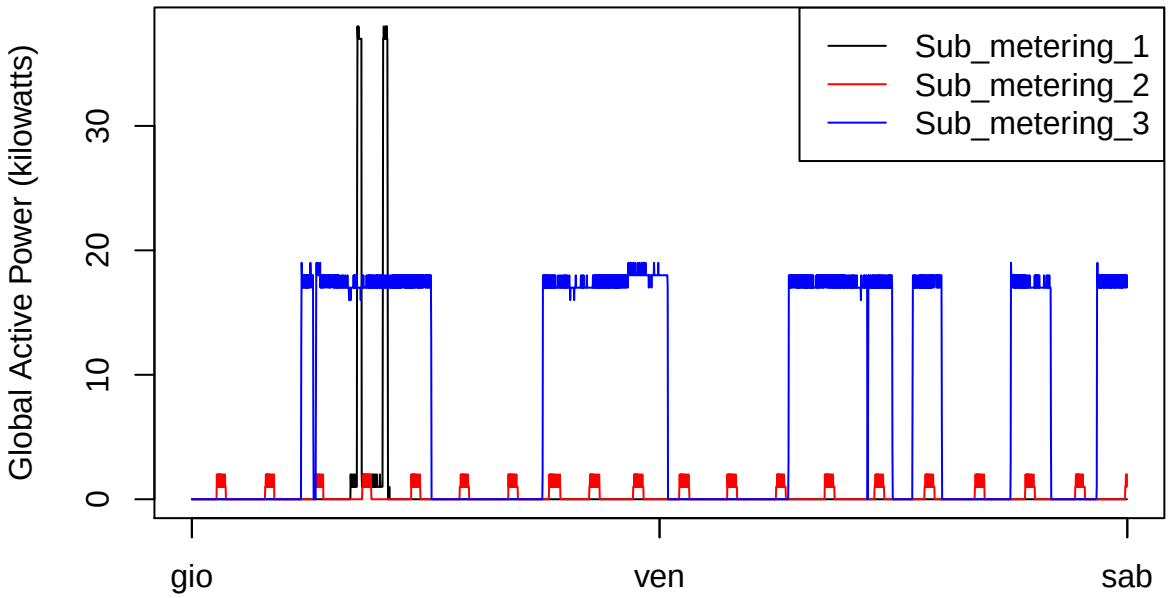
```
hist(t$Active_power, main="Global Active Power", xlab = "Global Active Power (kilowatts)", col="red")
```



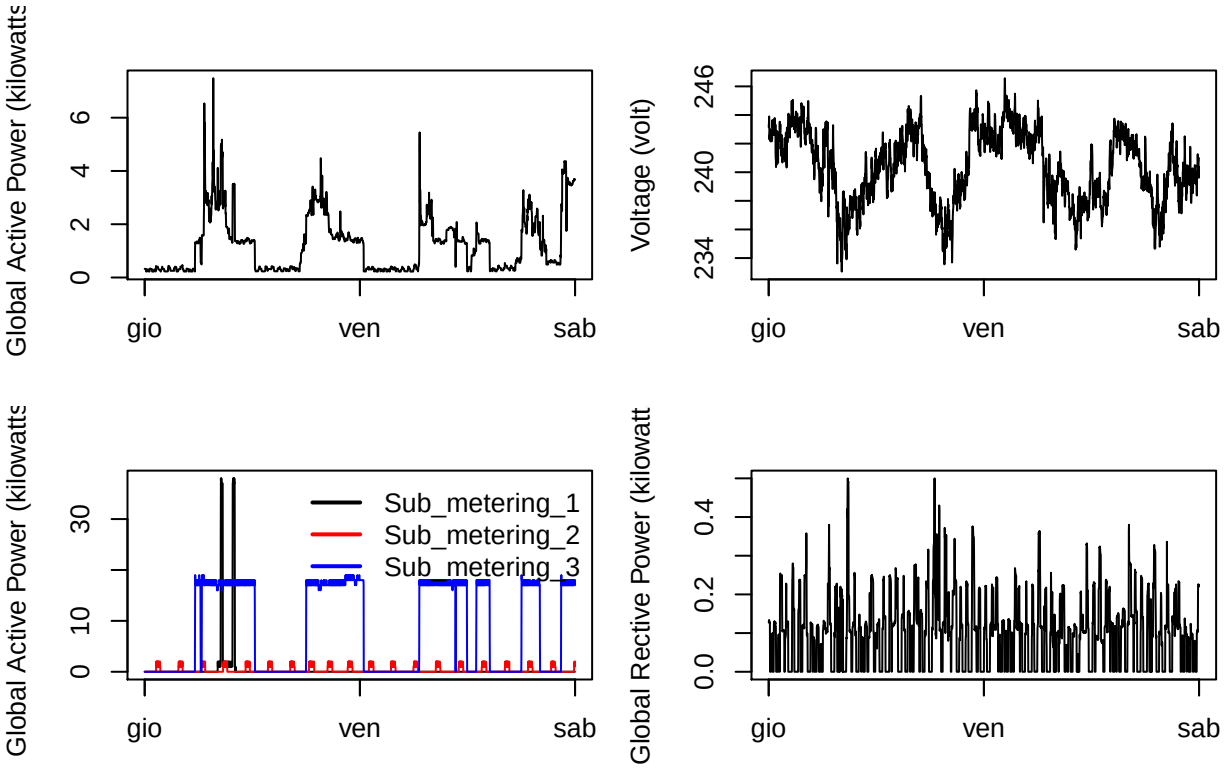

```
plot(t$Active_power~t$dateTime, type="l", ylab="Global Active Power (kilowatts)", xlab="")
```




```
with(t, {
  plot(type1~dateTime, type="l",
       ylab="Global Active Power (kilowatts)", xlab="")
  lines(type2~dateTime,col='Red')
  lines(type3~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"))
```




```
par(mfrow=c(2,2), mar=c(4,4,2,1), oma=c(0,0,2,0))
  with(t, {
    plot(Active_power~dateTime, type="l",
         ylab="Global Active Power (kilowatts)", xlab="")
    plot(Voltage~dateTime, type="l",
         ylab="Voltage (volt)", xlab="")
    plot(type1~dateTime, type="l",
         ylab="Global Active Power (kilowatts)", xlab="")
    lines(type2~dateTime,col='Red')
    lines(type3~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(Reactive_power~dateTime, type="l",
         ylab="Global Rective Power (kilowatts)",xlab="")
  })
```


Global Active Power (kilowatts
0
2
4
6
gio
ven
sab
Voltage (volt)
234
240
246
gio
ven
sab
Global Active Power (kilowatts
0
10
30
Sub_metering_1
Sub_metering_2
Sub_metering_3
gio
ven
sab
Global Rective Power (kilowatt
0.0
0.2
0.4
gio
ven
sab