Electric Power consumption

Introduction

This assignment uses data from the UC Irvine Machine Learning Repository, a popular repository for machine learning datasets. In particular, we will be using the “Individual household electric power consumption Data Set” which I have made available on the course web site:

Dataset: Electric power consumption [20Mb]

Description: Measurements of electric power consumption in one household with a one-minute sampling rate over a period of almost 4 years. Different electrical quantities and some sub-metering values are available. The following descriptions of the 9 variables in the dataset are taken from the UCI web site:

Date: Date in format dd/mm/yyyy

Time: time in format hh:mm:ss

Global_active_power: household global minute-averaged active power (in kilowatt)

Global_reactive_power: household global minute-averaged reactive power (in kilowatt)

Voltage: minute-averaged voltage (in volt)

Global_intensity: household global minute-averaged current intensity (in ampere)

Sub_metering_1: energy sub-metering No. 1 (in watt-hour of active energy). It corresponds to the kitchen, containing mainly a dishwasher, an oven and a microwave (hot plates are not electric but gas powered).

Sub_metering_2: energy sub-metering No. 2 (in watt-hour of active energy). It corresponds to the laundry room, containing a washing-machine, a tumble-drier, a refrigerator and a light.

Sub_metering_3: energy sub-metering No. 3 (in watt-hour of active energy). It corresponds to an electric water-heater and an air-conditioner.

Loading the data

```
data <- read.table("household_power_consumption.txt", header= TRUE, sep=";", stringsAsFac
tors=FALSE, dec=".")

summary(data)
```

```
##      Date               Time           Global_active_power
##  Length:2075259     Length:2075259     Length:2075259
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##  Global_reactive_power   Voltage          Global_intensity
##  Length:2075259        Length:2075259     Length:2075259
##  Class :character      Class :character   Class :character
```

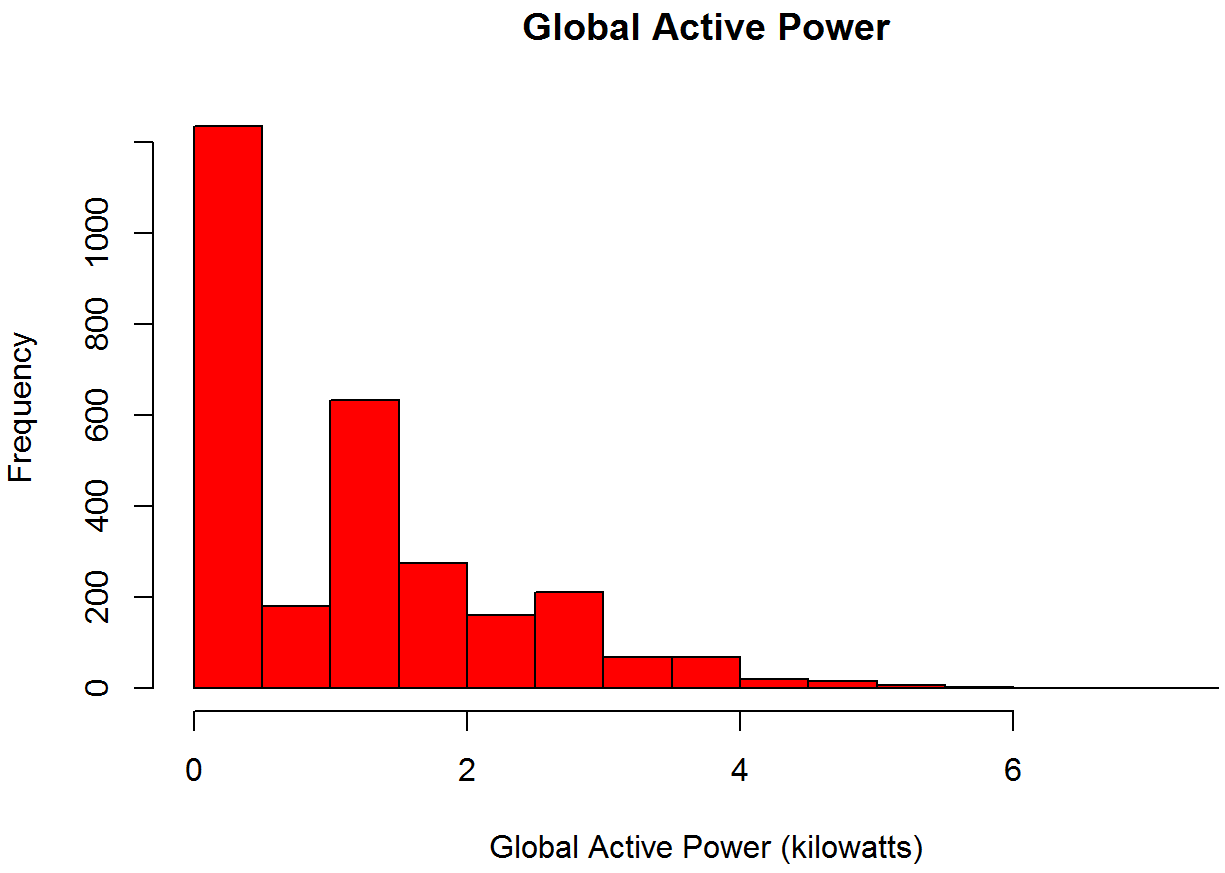
```
##  Mode   :character      Mode   :character   Mode   :character
##
##
##
##
##  Sub_metering_1     Sub_metering_2     Sub_metering_3
##  Length:2075259     Length:2075259     Min.   : 0.000
##  Class :character   Class :character   1st Qu.: 0.000
##  Mode  :character   Mode  :character   Median : 1.000
##                                        Mean   : 6.458
##                                        3rd Qu.:17.000
##                                        Max.   :31.000
##                                        NA's   :25979
```

Subset the data from the dates 2007-02-01 and 2007-02-02.

```
subsetdata <- data[data$Date %in% c("1/2/2007","2/2/2007"),]


globalActivePower <- as.numeric(subsetdata$Global_active_power)
globalReactivePower <- as.numeric(subsetdata$Global_reactive_power)
voltage <- as.numeric(subsetdata$Voltage)
subMetering1 <- as.numeric(subsetdata$Sub_metering_1)
subMetering2 <- as.numeric(subsetdata$Sub_metering_2)
subMetering3 <- as.numeric(subsetdata$Sub_metering_3)
```

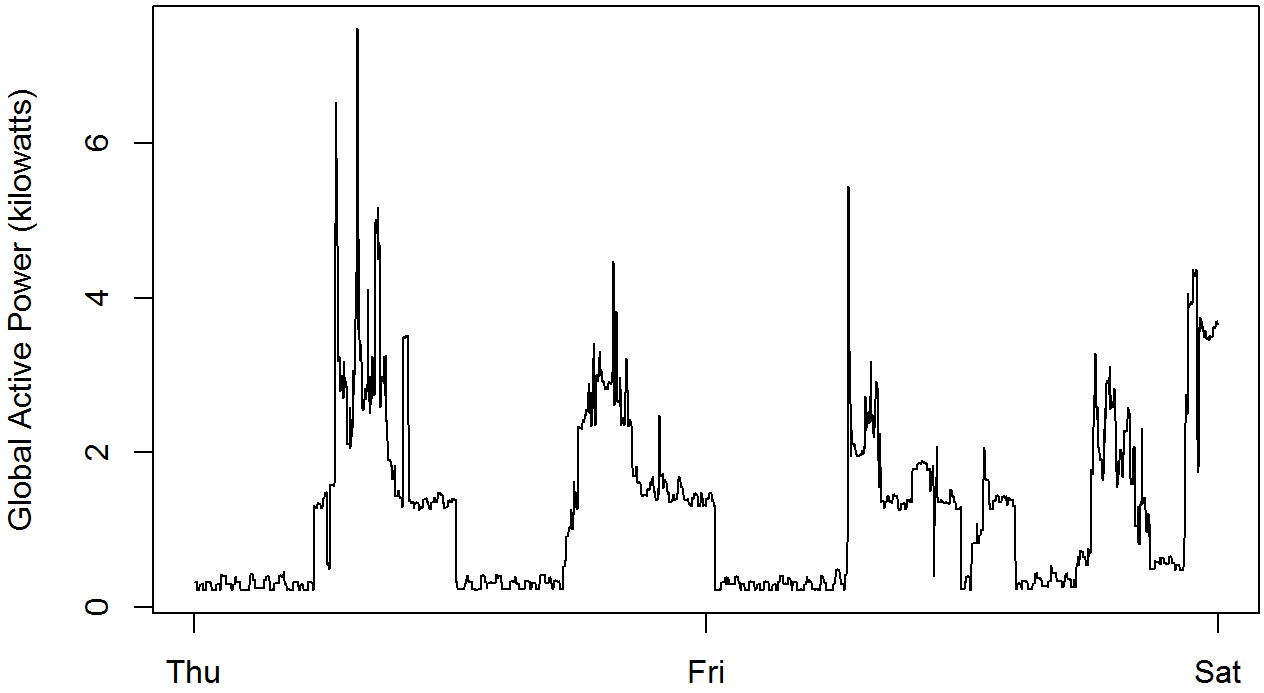
Created a histogram

```
hist(globalActivePower, col="red", main="Global Active Power", xlab="Global Active Power
(kilowatts)")
```



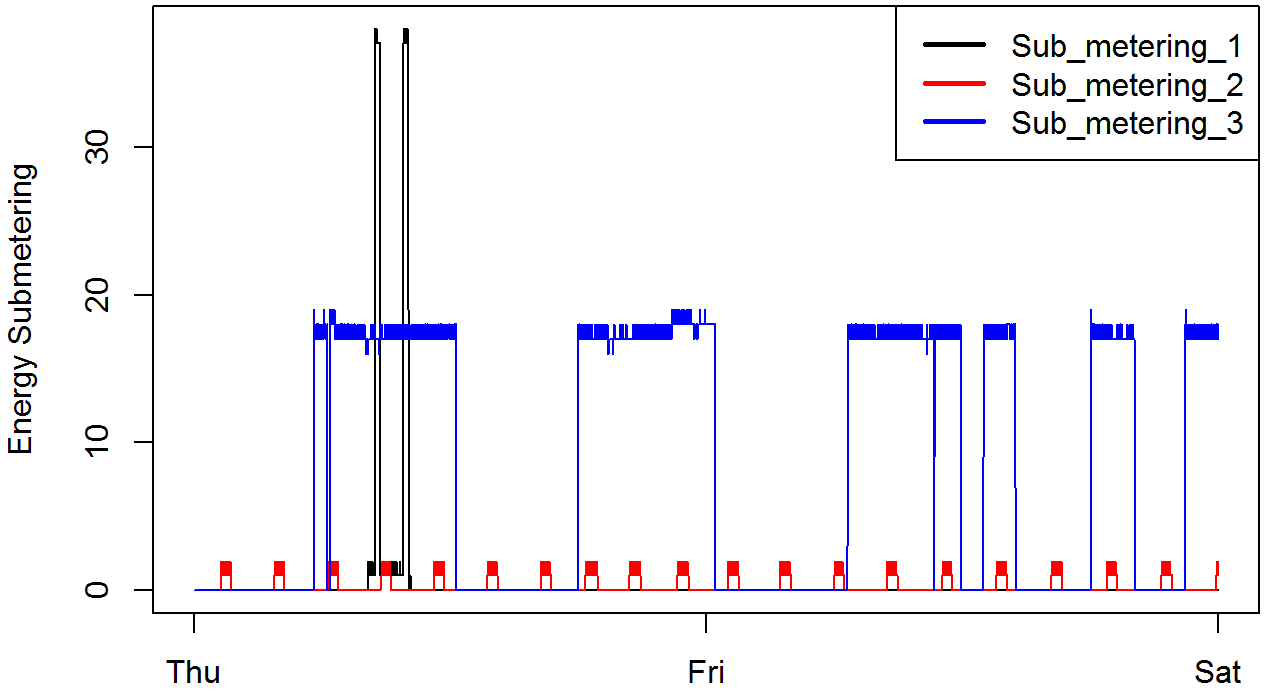
Create a Time series

```
datetime <- strptime(paste(subsetdata$Date, subsetdata$Time, sep=" "), "%d/%m/%Y %H:%M:%S
")

plot(datetime, globalActivePower, type="l", xlab="", ylab="Global Active Power (kilowatts
)")
```



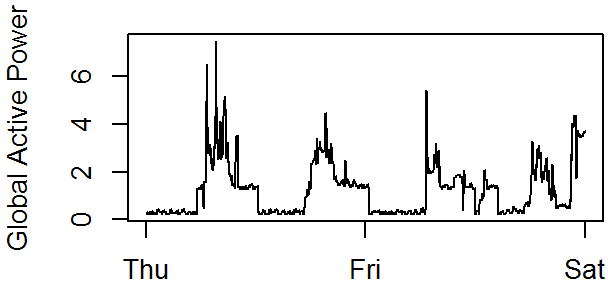
Create a plot for sub metering

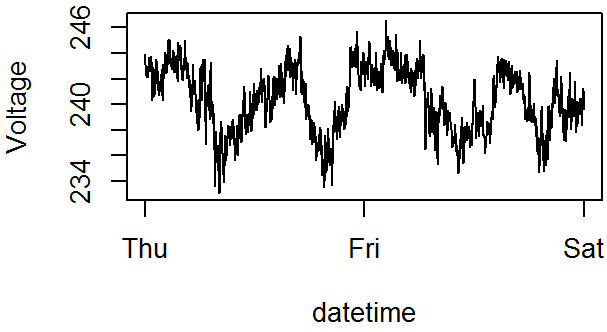
```
plot(datetime, subMetering1, type="l", ylab="Energy Submetering", xlab="")
lines(datetime, subMetering2, type="l", col="red")
lines(datetime, subMetering3, type="l", col="blue")
legend("topright", c("Sub_metering_1", "Sub_metering_2", "Sub_metering_3"), lty=1, lwd=2.5, col=c("black", "red", "blue"))
```

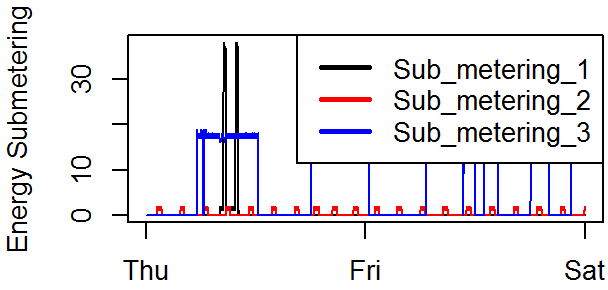


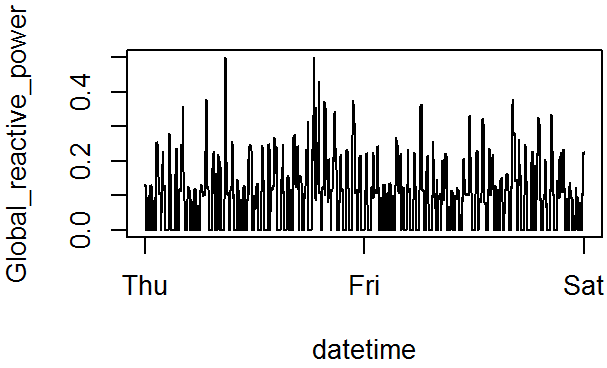
Create multiple plot

```
par(mfrow = c(2, 2))

# First plot

plot(datetime, globalActivePower, type="l", xlab="", ylab="Global Active Power", cex=0.2)

# Second plot

plot(datetime, voltage, type="l", xlab="datetime", ylab="Voltage")

# Third plot

plot(datetime, subMetering1, type="l", ylab="Energy Submetering", xlab="")

lines(datetime, subMetering2, type="l", col="red")

lines(datetime, subMetering3, type="l", col="blue")

legend("topright", c("Sub_metering_1", "Sub_metering_2", "Sub_metering_3"), lty=, lwd=2.5
, col=c("black", "red", "blue"), bty="o")

# Fourth plot

plot(datetime, globalReactivePower, type="l", xlab="datetime", ylab="Global_reactive_powe
r", cex=0.2)
```


Global Active Power
0
2
4
6
Thu
Fri
Sat


Voltage
234
240
246
Thu
Fri
Sat
datetime


Energy Submetering
0
10
30
Sub_metering_1
Sub_metering_2
Sub_metering_3
Thu
Fri
Sat


Global_reactive_power
0.0
0.2
0.4
Thu
Fri
Sat
datetime