plot2

Harold Pollack

9/15/2018

This is the markdown file that creates plot2. I didn’t print out all the intermediate products that are the same as those for plot one.

I first read the UCI data. Then save the dates 2/1/2007 and 2/2/2007. Note that the date variable is dd/mm/y. So some care must be taken. ALso cache=True given size of the dataset.

mydata <- read.csv("/Users/haroldpollack/Documents/coursera\_EDA/pS1/household\_power\_consumption.txt",header=TRUE,sep=";")  
mydata$date\_formatted <-as.Date(mydata$Date,format="%d/%m/%Y")  
# str(mydata)  
# summary(mydata)  
# table(mydata$Date)  
#  
# now extract the two dates called for in the assignment  
#  
data\_for\_graphing <-subset(mydata,(Date== "1/2/2007" | Date == "2/2/2007"))  
# str(data\_for\_graphing)  
# summary(data\_for\_graphing)  
# table(data\_for\_graphing$Date)  
table(data\_for\_graphing$date\_formatted)

##   
## 2007-02-01 2007-02-02   
## 1440 1440

#  
# now figure out weekday  
#  
data\_for\_graphing$weekday <- weekdays(data\_for\_graphing$date\_formatted)  
table(data\_for\_graphing$weekday)

##   
## Friday Thursday   
## 1440 1440

Now manipulate date/time. I used code found here <http://biostat.mc.vanderbilt.edu/wiki/pub/Main/ColeBeck/datestimes.pdf>

data\_for\_graphing$dts1 <- paste(data\_for\_graphing$date\_formatted, data\_for\_graphing$Time)  
time\_scr <-strptime(data\_for\_graphing$dts1,format = "%Y-%m-%d %H:%M:%S")  
#  
# time\_scr uses strptime. I just included this to learn.  
#  
str(time\_scr)

## POSIXlt[1:2880], format: "2007-02-01 00:00:00" "2007-02-01 00:01:00" ...

data\_for\_graphing$dts2 <- as.POSIXct(data\_for\_graphing$dts1, format = "%Y-%m-%d %H:%M:%S")  
str(data\_for\_graphing$dts1)

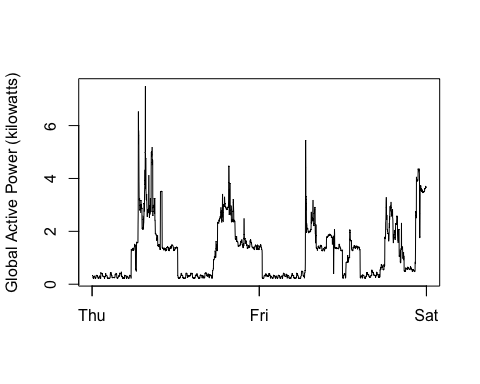
## chr [1:2880] "2007-02-01 00:00:00" "2007-02-01 00:01:00" ...

str(data\_for\_graphing$dts2)

## POSIXct[1:2880], format: "2007-02-01 00:00:00" "2007-02-01 00:01:00" ...

make the graph and save as plot2.png

#  
# note that data\_for\_graphing$Global\_active\_power is a factor. It has to be changed to numeric for graphing  
#  
# Global\_active\_power is a factor. So one must be careful converting it to numeric  
#  
data\_for\_graphing$Global\_active\_power\_numeric <- as.numeric(as.character(data\_for\_graphing$Global\_active\_power))   
# hist(Global\_active\_power\_numeric,col="red",xlab="Global Active Power (kilowatts)",main="Global Active Power",ylab="Frequency")  
  
#  
# time series by day  
#  
plot(data\_for\_graphing$dts2,data\_for\_graphing$Global\_active\_power\_numeric,type="l", xlab=" ", ylab = "Global Active Power (kilowatts)")



#  
# now save the png file  
#  
dev.copy(png,'/Users/haroldpollack/Documents/coursera\_datascience3/Plot2/plot2.png')

## quartz\_off\_screen   
## 3

dev.off()

## quartz\_off\_screen   
## 2