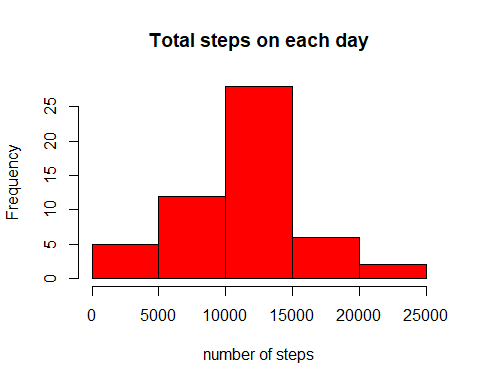
Reproducible Research: Peer Assessment 1

library(ggplot2)  
library(lattice)  
library(Hmisc)  
library(dplyr)

## Loading and preprocessing the data

file = "activity.csv"  
if (!file.exists(file)){  
 unzip("RepData\_PeerAssessment1/activity.zip")  
}  
data <- read.csv("activity.csv",header=TRUE,na.strings = "NA")  
  
totalstepsperdate <- aggregate(steps~date,data=data,sum)  
hist(totalstepsperdate$steps,breaks = 5,main="Total steps on each day",xlab="number of steps",col="red")



## What is mean total number of steps taken per day?

meansteps = mean(totalstepsperdate$steps)  
mediansteps = median(totalstepsperdate$steps)  
print (meansteps)

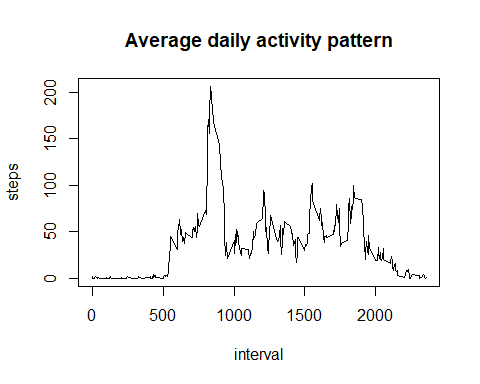
## [1] 10766.19

print (mediansteps)

## [1] 10765

## What is the average daily activity pattern?

totstepsbyinterval <- aggregate(steps~interval,data=data,mean)  
with(totstepsbyinterval,plot(interval,steps,type='l',main="Average daily activity pattern"))

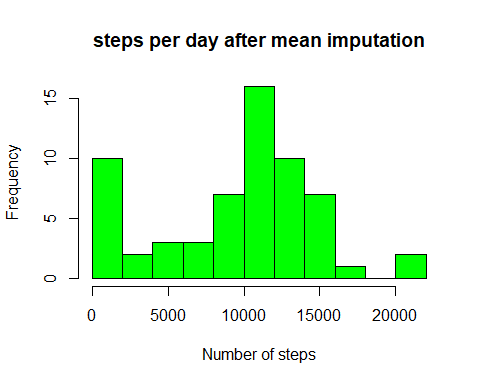


totstepsbyinterval[which.max(totstepsbyinterval$steps),]

## interval steps  
## 104 835 206.1698

## Imputing missing values

meandata <- data  
meandata <- mutate(meandata, steps=impute(steps))  
  
stepsperdata <- aggregate(steps~date,data=meandata,sum)  
hist(stepsperdata$steps,main="steps per day after mean imputation",xlab="Number of steps",col="green",breaks=15)



## Mean and median after imputation

meansteps = mean(stepsperdata$steps)  
mediansteps = median(stepsperdata$steps)

## Are there differences in activity patterns between weekdays and weekends?

daycreater <- function(date) {  
 day <- weekdays(date)  
 if (day %in% c("Monday","Tuesday","Wednesday","Thursday","Friday")) {  
 return ("weekday")  
 } else if (day %in% c("Saturday","Sunday")){  
 return ("weekend")  
 } else {  
 print (day)  
 return ("Invalid date")  
   
 }  
}  
  
meandata$date <- as.Date(meandata$date)  
meandata$day <- sapply(meandata$date,FUN = daycreater)  
  
stepsperinterval <- aggregate(steps~interval + day,meandata,mean)  
  
averages <- aggregate(steps ~ interval + day, data=meandata, mean)  
ggplot(averages, aes(interval, steps)) + geom\_line() + facet\_grid(day ~ .) +  
 xlab("5-minute interval") + ylab("Number of steps")

