# Project 1 Document

Claudio Freitas

## Section 1

This is the first code chunk to read the data file and transform data:

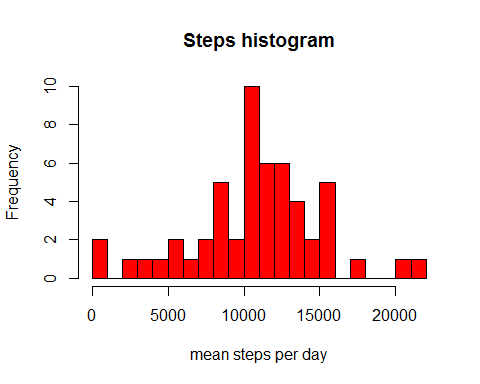
library(plyr)  
activity <- read.csv("./repdata-data-activity/activity.csv")  
activity$date <- as.Date(activity$date)  
activity$steps <- as.numeric(activity$steps)  
activity$interval <- as.numeric(activity$interval)  
apply(activity, 2, max, na.rm=T)

## steps date interval   
## "806" "2012-11-30" "2355"

## Section2

Subset data, calculate total steps per day and plot histogram

subdata <- ddply(activity, .(date), summarize, agregate = sum(steps))  
hist(subdata$agregate,col="red",breaks=20,xlab="mean steps per day", main="Steps histogram")



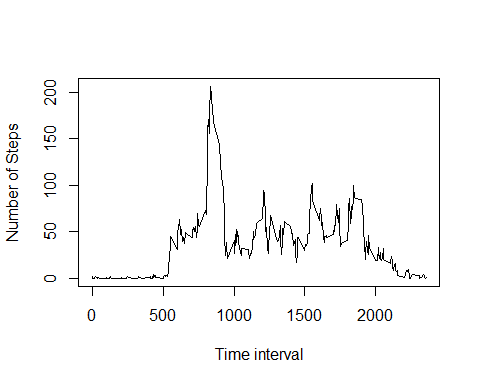
meansteps <- format(mean(subdata$agregate, na.rm=T), nsmall=2)  
mediansteps <- format(median(subdata$agregate,na.rm=T), nsamll=2)

The mean is 10766.19 and the median is 10765.

## Section 3

subset data to calculate average step for each interval and plot a linear chart

subdata1 <- ddply(activity, .(interval), summarize, agregateX = mean(interval))  
subdata2 <- with(activity, tapply(steps, interval, mean, na.rm=T))  
plot(subdata1$agregateX, subdata2, type="l", xlab = "Time interval", ylab = "Number of Steps")



Calculate the interval with max average steps

maxi <- max(subdata2)  
maxint <- which.max(subdata2)  
maxint<- subdata1$agregateX[maxint]

The maximum value for average number of steps in a particluar time interval is 206.1698113

The 5 minutes time interval for the maximum average number of steps is 835

## Section 4

Calculating the number of rows with NAs

totalna<- sum(is.na(activity$steps))  
print (totalna)

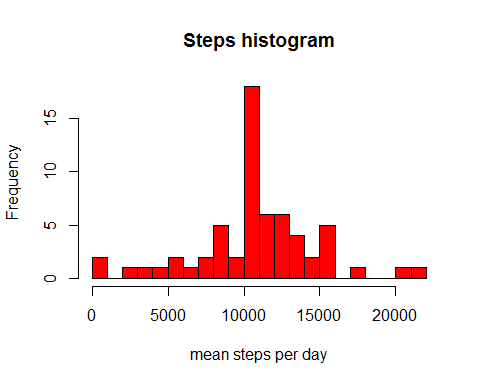
## [1] 2304

Create a new dataset without NAs

totalna<- is.na(activity$steps)  
subdata<- tapply(activity$steps, activity$interval, mean, na.rm=T)  
activity$steps[totalna] <- subdata[as.character(activity$interval[totalna])]

Plot a histogram with the sum of the steps for each day

result<- tapply(activity$steps, activity$date, sum, na.rm=T)  
hist(result, col="red",breaks=20, xlab="mean steps per day", main="Steps histogram")



mean1<- format(mean(result))  
median1<- format(median(result))

The mean is 10766.19

The median is 10766.19

## Section 5

Subseting weekday and weekend using the wekdays() function

activity$day<- weekdays(as.Date(activity$date))  
actwday<- subset(activity, !day %in% c("sabado", "domingo"))  
actwend<- subset(activity, day %in% c("sabado","domingo"))

creating a panel with the two charts (weekend and weekdays)

par(mfrow=c(1,2))  
Yweekdays<- tapply(actwday$steps, actwday$interval, mean, na.rm=T)  
Xweekdays<- tapply(actwday$interval, actwday$interval,unique, na.rm=T)  
Yweekends<- tapply(actwend$steps, actwend$interval, mean, na.rm=T)  
Xweekends<- tapply(actwend$interval, actwend$interval,unique, na.rm=T)  
plot(Xweekdays, Yweekdays, type="l",xlab = "Weekdays activity", ylab = "Total steps")  
plot(Xweekends, Yweekends, type="l",xlab = "Weekends activity", ylab = "Total steps")

