title: “meanwhile”

#Get file and load libraries

path<- getwd()  
file.path(path, "repdata\_data\_activity")  
data<- unzip(zipfile = "repdata\_data\_activity.zip")  
data1<- read.csv(data)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.0.5

# Organize data

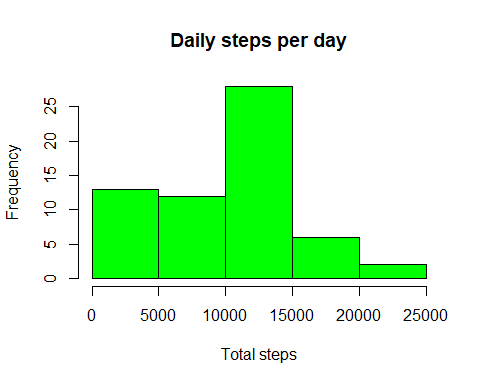
data1<- read.table(data, skip = 1, sep=",")  
names(data1)<-c("steps", "date", "interval")  
as.Date(data1$date)

# group steps by date and do a function with the mean steps

stepsperday<- data1 %>%  
 group\_by(date)%>%  
 summarise(sumsteps=sum(steps, na.rm = T))

# Make histogram and get the mean and the median

hist(stepsperday$sumsteps, freq = T, col="green", main = "Daily steps per day", xlab = "Total steps")



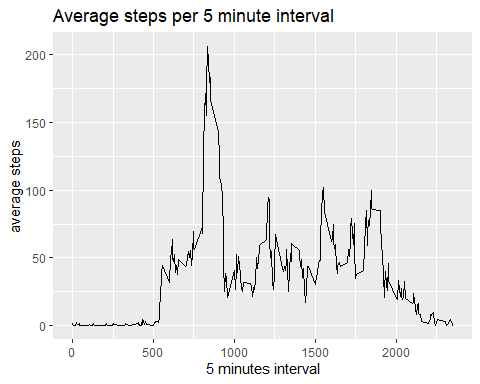
mean<- round(mean(stepsperday$sumsteps))  
median<- round(median(stepsperday$sumsteps))

#Group average per interval

databyinterval <- data1%>%   
 select(interval, steps) %>%   
 na.omit() %>%   
 group\_by(interval) %>% summarize(sumsteps= mean(steps))

##Time series plot

ggplot(databyinterval, aes(x=interval, y=sumsteps))+ geom\_line()+ ylab("average steps")+ xlab("5 minutes interval") + ggtitle("Average steps per 5 minute interval")



databyinterval[which(databyinterval$sumsteps== max(databyinterval$sumsteps)),]

## # A tibble: 1 x 2  
## interval sumsteps  
## <int> <dbl>  
## 1 835 206.

#replace missing values and check the data

missing\_values<- sum(is.na(data1))  
replacena <- function(x)replace(x, is.na(x), mean(x, na.rm=T))  
meandata<- data1%>% group\_by(interval) %>% mutate(steps= replacena(steps))  
is.na(meandata)

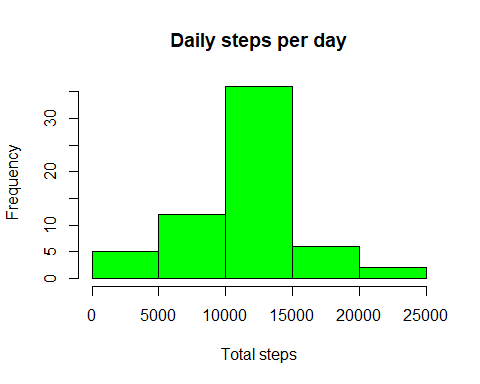
#Data in steps per day

stepsPerDay1 <- meandata %>%  
 group\_by(date) %>%  
 summarize(sumsteps = sum(steps, na.rm = TRUE))

# Histogram of total steps per day whithout NA

## In the lower spectrum of steps there is a great change, the mean changes for about a thousand

hist(stepsPerDay1$sumsteps, freq = T, col="green", main = "Daily steps per day", xlab = "Total steps")



mean1<- round(mean(stepsPerDay1$sumsteps))  
median1<- round(median(stepsPerDay1$sumsteps))  
mean1

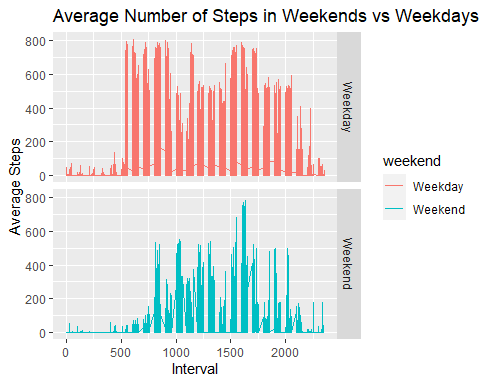
## [1] 10766

median1

## [1] 10766

# Differences in activity patterns between weekdays and weekends

meandata$date <- as.Date(meandata$date)  
meandata$weekday <- weekdays(meandata$date)  
meandata$weekend <- ifelse(meandata$weekday=="sabado" | meandata$weekday=="domingo", "Weekend", "Weekday" )  
  
  
ggplot(meandata, aes(x=interval, y=steps, color=weekend)) + geom\_line()+  
 facet\_grid(weekend ~.) + xlab("Interval") + ylab("Average Steps") +  
 ggtitle("Average Number of Steps in Weekends vs Weekdays")

 library(knitr) knit2html(“codeactivity.R”)