# Reproducible Research: Peer Assessment 1

Luis Fernando Agottani

09/08/2019

## Loading and preprocessing the data

Verify if the filename has been downloaded/unzip yet. If not, download/unzip the file with the fileURL link.

```
filename <- "repdata_data_activity.zip"

if (!file.exists(filename)){
  fileURL <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2Factivity.zip"
  download.file(fileURL, filename)
}
if (!file.exists("activity.csv")) {
  unzip(filename)
}
```

Reading data.csv.

```
Act<- read.csv("activity.csv",
               head = TRUE, sep= ",")
```

Changing the Factor type of the column date to Date with as.Date.

```
Act$date<- as.Date(Act$date)
```

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

Library Dplyr to work with “%>%”

```
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
```

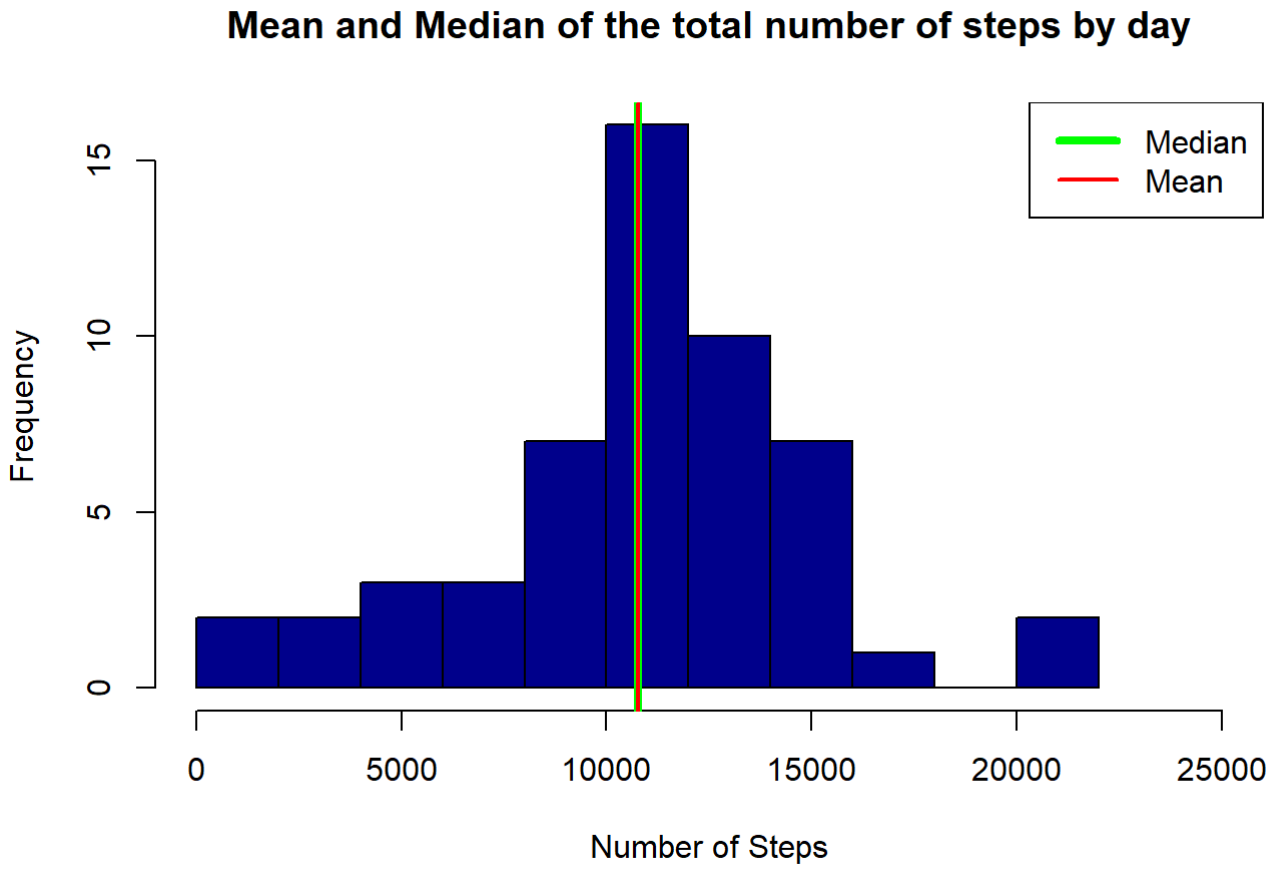
The total number of steps is the sum of all steps measured by day. The mean is the sum divided by the number of intervals..

```
Act_Day<- Act %>%
  group_by(date) %>%
  summarise(steps = sum(steps))
```

# Including Plots

Ploting a histogram with the sum, mean and median of steps by day.

```
hist(Act_Day$steps,
     breaks= 15,
     col="darkblue",
     border="black",
     xlab="Number of Steps",
     main="Mean and Median of the total number of steps by day",
     xlim=c(0,25000))

abline(v=median(Act_Day$steps,
                na.rm=TRUE),
       col="green",
       lwd=4)

abline(v=mean(Act_Day$steps,
              na.rm=TRUE),
       col="red",lwd=2)

legend(x="topright", #posicao da legenda
      c("Median","Mean"), #nomes da legenda
      col=c("green","red"), #cores
      lty=c(1,1), #estilo da linha
      lwd=c(4,2)) #grossura das linhas
```



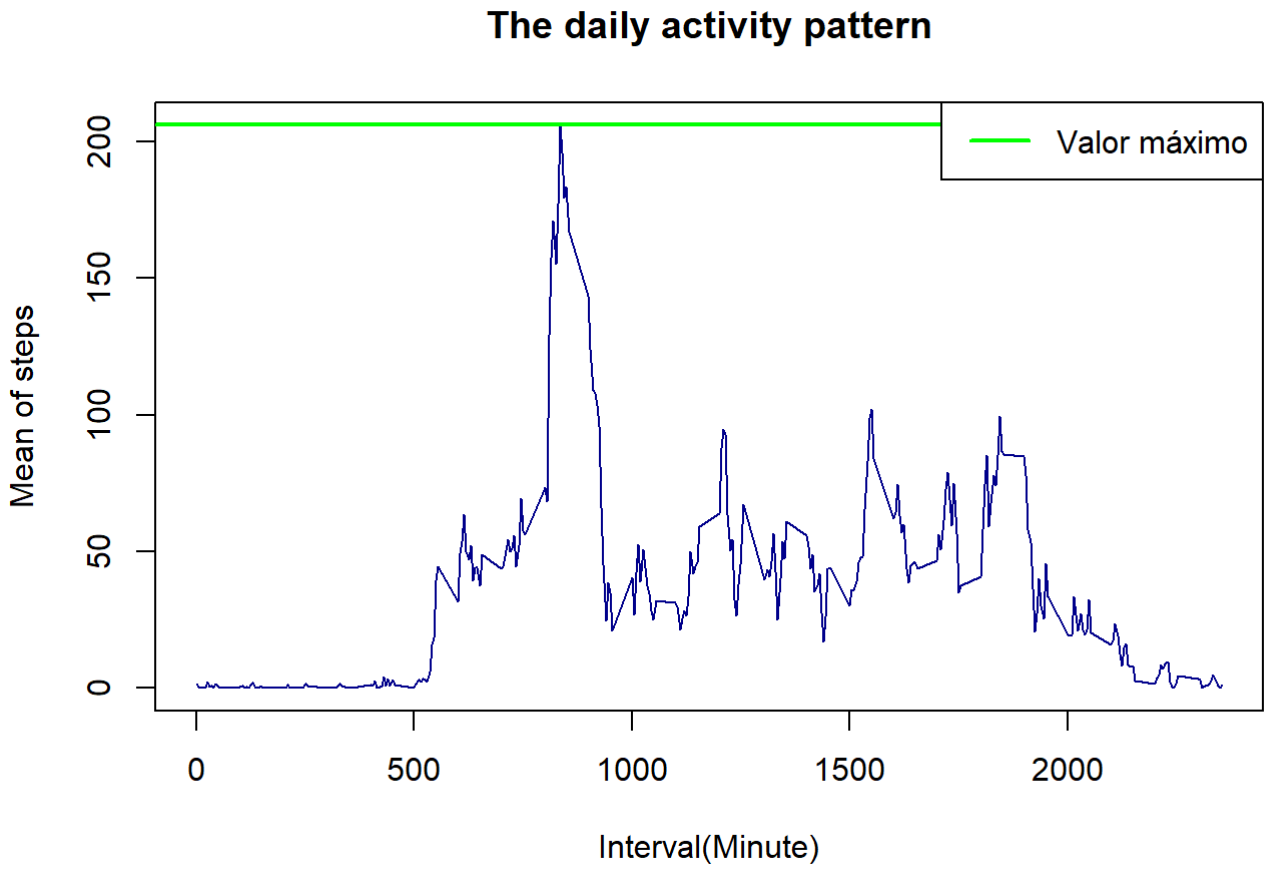

# What is the average daily activity pattern?

The daily activity pattern is the mean of the number of steps of all days by interval.

```
Act_Interval<- Act %>%
  group_by(interval) %>%
  summarise(steps = mean(steps,
                         na.rm=TRUE))
```

Ploting the Mean of steps of all days by Interval with the line plot.

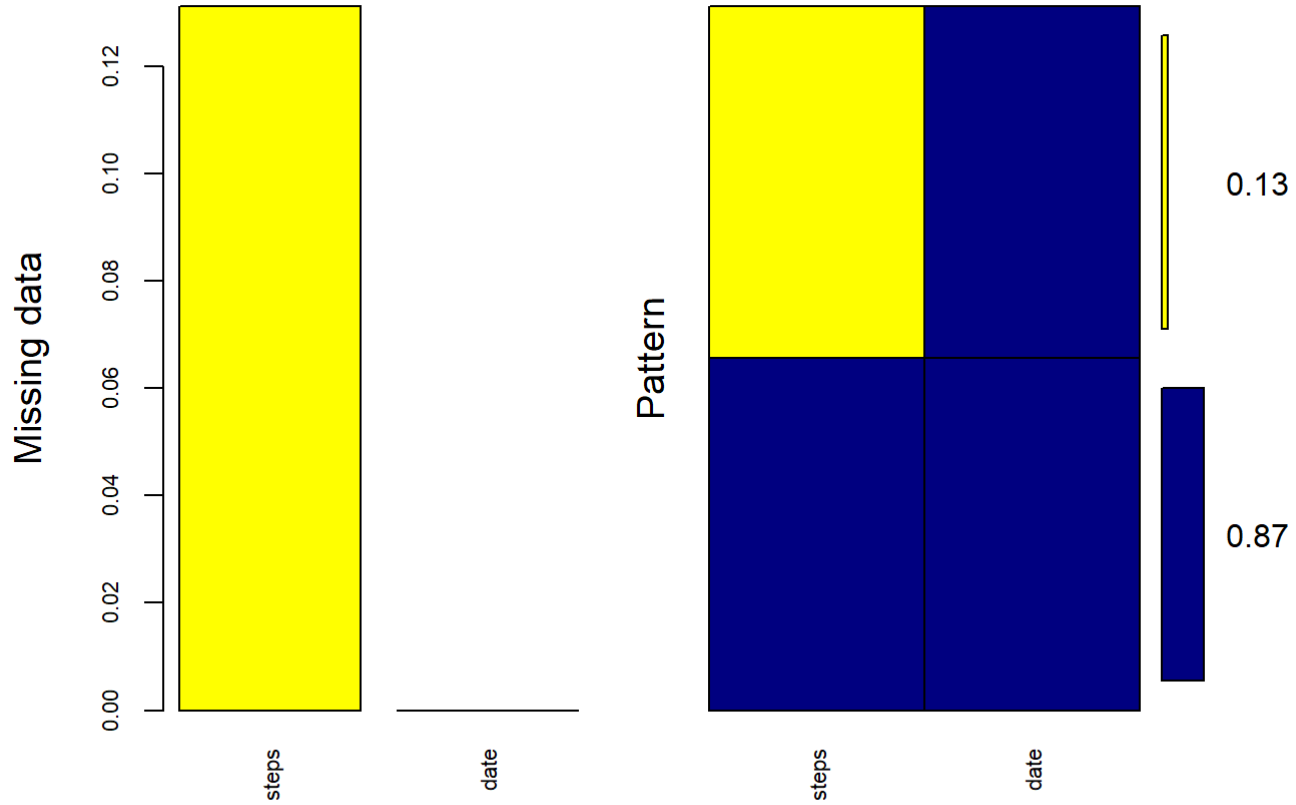
```
plot(Act_Interval$interval,
     Act_Interval$steps,
     type="l",
     col="darkblue",
     xlab="Interval(Minute)",
     ylab="Mean of steps",
     main="The daily activity pattern")

abline(h=max(Act_Interval$steps),
       col="green",
       lwd=2)

legend(x="topright", #posicao da legenda
      "Valor máximo", #nomes da legenda
      col="green",#cores
      lty=1,#estilo da linha
      lwd=2) #grossura das linhas
```




# Imputing missing values

To represent the missing data and verify the number of the missing data in each column, a nice way to do it is with the package “mice” and “VIM”.

```
library(mice)
library(VIM)

mice_plot <- aggr(Act_Day,
                  col=c('navyblue','yellow'),
                  numbers=TRUE,
                  sortVars=TRUE,
                  labels=names(Act_Day),
                  cex.axis=.7,
                  gap=3,
                  ylab=c("Missing data","Pattern"))
```




```
##
##  Variables sorted by number of missings:
##  Variable     Count
##     steps 0.1311475
##      date 0.0000000
```

The result show us that is a big amount, x>5%, of NA's in the data, we have to take care to impute the a value that help us with accuracy.

With library "mice" and the pmm method, we will impute values to missing data.

```
imputed_Data <- mice(Act,
                     m=5,
                     maxit = 50,
                     method = 'pmm',
                     seed = 500)
```

```
##
##  iter imp variable
##   1   1  steps
##   1   2  steps
##   1   3  steps
##   1   4  steps
##   1   5  steps
##   2   1  steps
##   2   2  steps
##   2   3  steps
##   2   4  steps
##   2   5  steps
##   3   1  steps
##   3   2  steps
##   3   3  steps
##   3   4  steps
##   3   5  steps
##   4   1  steps
##   4   2  steps
##   4   3  steps
##   4   4  steps
##   4   5  steps
##   5   1  steps
##   5   2  steps
##   5   3  steps
##   5   4  steps
##   5   5  steps
##   6   1  steps
##   6   2  steps
##   6   3  steps
##   6   4  steps
##   6   5  steps
##   7   1  steps
##   7   2  steps
##   7   3  steps
##   7   4  steps
##   7   5  steps
##   8   1  steps
##   8   2  steps
##   8   3  steps
##   8   4  steps
##   8   5  steps
##   9   1  steps
##   9   2  steps
##   9   3  steps
##   9   4  steps
##   9   5  steps
##   10   1  steps
##   10   2  steps
##   10   3  steps
##   10   4  steps
##   10   5  steps
##   11   1  steps
##   11   2  steps
##   11   3  steps
##   11   4  steps
##   11   5  steps
```

```
##   12   1  steps
##   12   2  steps
##   12   3  steps
##   12   4  steps
##   12   5  steps
##   13   1  steps
##   13   2  steps
##   13   3  steps
##   13   4  steps
##   13   5  steps
##   14   1  steps
##   14   2  steps
##   14   3  steps
##   14   4  steps
##   14   5  steps
##   15   1  steps
##   15   2  steps
##   15   3  steps
##   15   4  steps
##   15   5  steps
##   16   1  steps
##   16   2  steps
##   16   3  steps
##   16   4  steps
##   16   5  steps
##   17   1  steps
##   17   2  steps
##   17   3  steps
##   17   4  steps
##   17   5  steps
##   18   1  steps
##   18   2  steps
##   18   3  steps
##   18   4  steps
##   18   5  steps
##   19   1  steps
##   19   2  steps
##   19   3  steps
##   19   4  steps
##   19   5  steps
##   20   1  steps
##   20   2  steps
##   20   3  steps
##   20   4  steps
##   20   5  steps
##   21   1  steps
##   21   2  steps
##   21   3  steps
##   21   4  steps
##   21   5  steps
##   22   1  steps
##   22   2  steps
##   22   3  steps
##   22   4  steps
##   22   5  steps
##   23   1  steps
##   23   2  steps
##   23   3  steps
```

```
##   23   4  steps
##   23   5  steps
##   24   1  steps
##   24   2  steps
##   24   3  steps
##   24   4  steps
##   24   5  steps
##   25   1  steps
##   25   2  steps
##   25   3  steps
##   25   4  steps
##   25   5  steps
##   26   1  steps
##   26   2  steps
##   26   3  steps
##   26   4  steps
##   26   5  steps
##   27   1  steps
##   27   2  steps
##   27   3  steps
##   27   4  steps
##   27   5  steps
##   28   1  steps
##   28   2  steps
##   28   3  steps
##   28   4  steps
##   28   5  steps
##   29   1  steps
##   29   2  steps
##   29   3  steps
##   29   4  steps
##   29   5  steps
##   30   1  steps
##   30   2  steps
##   30   3  steps
##   30   4  steps
##   30   5  steps
##   31   1  steps
##   31   2  steps
##   31   3  steps
##   31   4  steps
##   31   5  steps
##   32   1  steps
##   32   2  steps
##   32   3  steps
##   32   4  steps
##   32   5  steps
##   33   1  steps
##   33   2  steps
##   33   3  steps
##   33   4  steps
##   33   5  steps
##   34   1  steps
##   34   2  steps
##   34   3  steps
##   34   4  steps
##   34   5  steps
##   35   1  steps
```

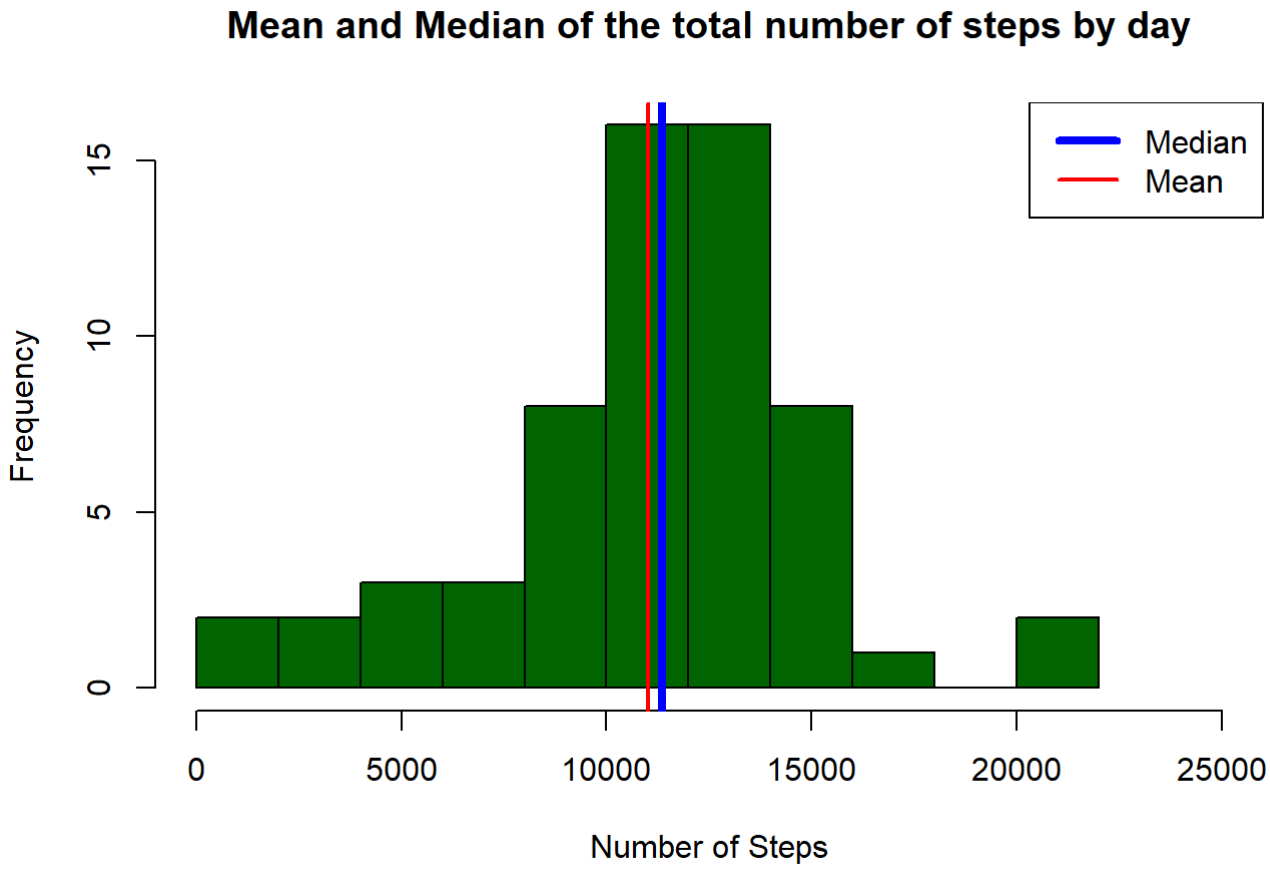
```
##   35   2  steps
##   35   3  steps
##   35   4  steps
##   35   5  steps
##   36   1  steps
##   36   2  steps
##   36   3  steps
##   36   4  steps
##   36   5  steps
##   37   1  steps
##   37   2  steps
##   37   3  steps
##   37   4  steps
##   37   5  steps
##   38   1  steps
##   38   2  steps
##   38   3  steps
##   38   4  steps
##   38   5  steps
##   39   1  steps
##   39   2  steps
##   39   3  steps
##   39   4  steps
##   39   5  steps
##   40   1  steps
##   40   2  steps
##   40   3  steps
##   40   4  steps
##   40   5  steps
##   41   1  steps
##   41   2  steps
##   41   3  steps
##   41   4  steps
##   41   5  steps
##   42   1  steps
##   42   2  steps
##   42   3  steps
##   42   4  steps
##   42   5  steps
##   43   1  steps
##   43   2  steps
##   43   3  steps
##   43   4  steps
##   43   5  steps
##   44   1  steps
##   44   2  steps
##   44   3  steps
##   44   4  steps
##   44   5  steps
##   45   1  steps
##   45   2  steps
##   45   3  steps
##   45   4  steps
##   45   5  steps
##   46   1  steps
##   46   2  steps
##   46   3  steps
##   46   4  steps
```

```
##   46   5  steps
##   47   1  steps
##   47   2  steps
##   47   3  steps
##   47   4  steps
##   47   5  steps
##   48   1  steps
##   48   2  steps
##   48   3  steps
##   48   4  steps
##   48   5  steps
##   49   1  steps
##   49   2  steps
##   49   3  steps
##   49   4  steps
##   49   5  steps
##   50   1  steps
##   50   2  steps
##   50   3  steps
##   50   4  steps
##   50   5  steps
```

```
completeData <- complete(imputed_Data,2)

completeDataSum<- completeData %>%
  group_by(date) %>%
  summarise(steps = sum(steps))
```

Ploting the complete data.

```
hist(completeDataSum$steps,
     breaks= 15,
     col="darkgreen",
     border="black",
     xlab="Number of Steps",
     main="Mean and Median of the total number of steps by day",
     xlim=c(0,25000))

abline(v=median(completeDataSum$steps,na.rm=TRUE),
       col="blue",
       lwd=4)

abline(v=mean(completeDataSum$steps,
              na.rm=TRUE),
       col="red",
       lwd=2)

legend(x="topright", #posicao da legenda
c("Median","Mean"), #nomes da legenda
col=c("blue","red"), #cores
lty=c(1,1), #estilo da linha
lwd=c(4,2)) #grossura das linhas
```



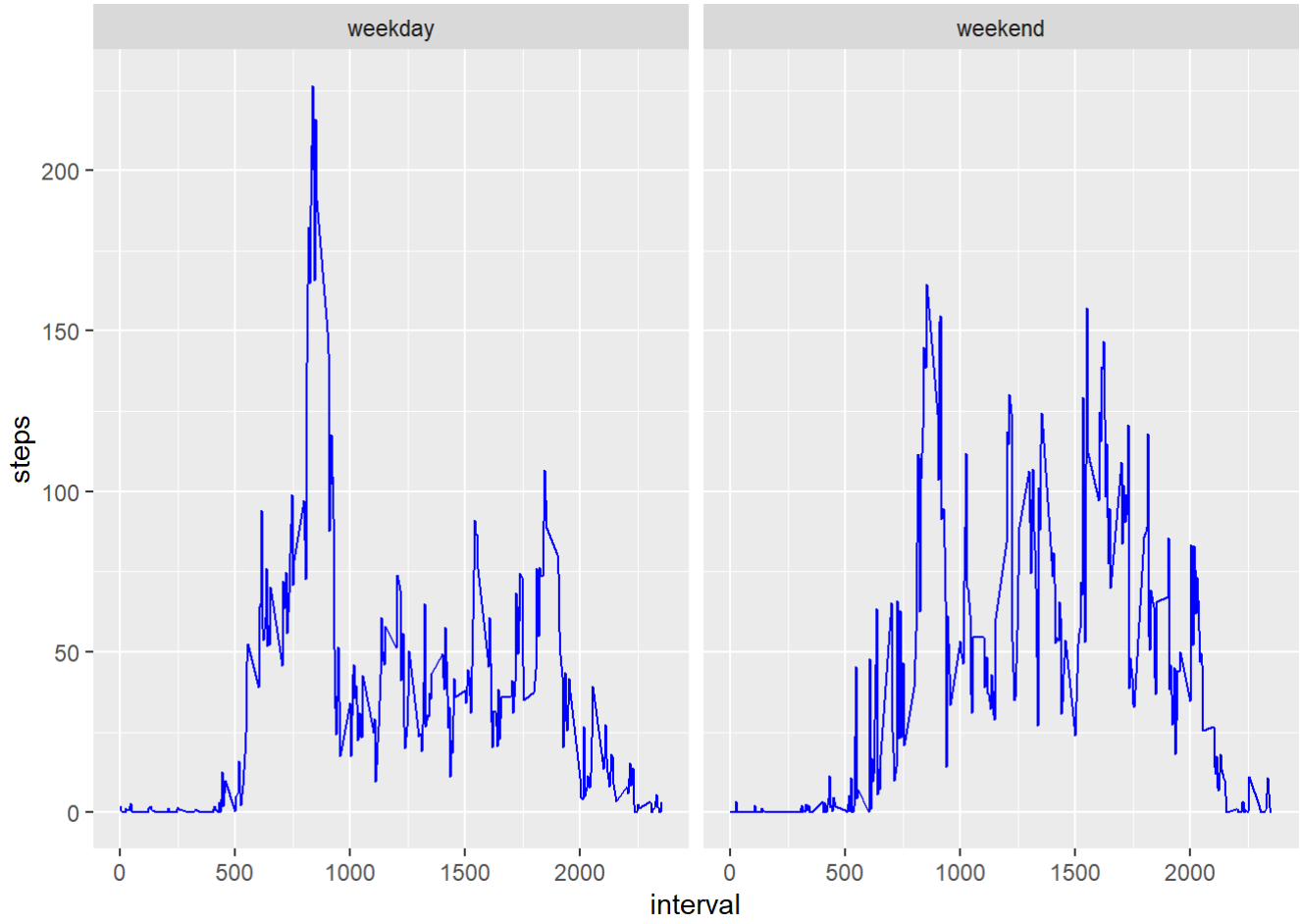

As we can see, imputing values with de pmm method showed us a similar result comparing with the first part of the assignment. So, we can assume that imputing missing values did not impacted signifcally.

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

```
Act_days<- mutate(completeData,
                  days=weekdays(completeData$date))

Act_week<- mutate(Act_days,
                  weekpart=(ifelse(Act_days$days=="sábado"|Act_days$days=="domingo","weekend"
,"weekday")))

Act_weekdays<- Act_week %>%
  group_by(interval,weekpart) %>%
  summarise(steps = mean(steps))
```

Ploting the comparassion between the parts of the week, “weekdays” and “weekend”, to see if have some pattern there.

```
library(ggplot2)

ggplot(aes(x=interval,y=steps),
       data=Act_weekdays)+
       geom_line(col="blue")+
       facet_wrap(~Act_weekdays$weekpart)
```


weekday
weekend
steps
200
150
100
50
0
0
500
1000
1500
2000
interval