|  |
| --- |
| getwd()  [1] "C:/Users/user/Desktop"  > setwd("C:/Users/user/Desktop/New folder (3)")  > # Set Global Echo = On  >  > # Load data  > if (!file.exists("activity.csv") )  + {  + dlurl <- 'http://d396qusza40orc.cloudfront.net/repdata%2Fdata%2Factivity.zip'  + download.file(dlurl,destfile='repdata%2Fdata%2Factivity.zip',mode='wb')  + unzip('repdata%2Fdata%2Factivity.zip')  + }  >  > # Read data  > data <- read.csv("activity.csv")  > steps\_by\_day <- aggregate(steps ~ date, data, sum)  > hist(steps\_by\_day$steps, main = paste("Total Steps Each Day"), col="green",xlab="Number of Steps") |
|  |
| |  | | --- | | > | |

rmean <- mean(steps\_by\_day$steps)

> rmean

[1] 10766.19

> rmedian <- median(steps\_by\_day$steps)

> rmedian

[1] 10765

> steps\_by\_interval <- aggregate(steps ~ interval, data, mean)

> plot(steps\_by\_interval$interval,steps\_by\_interval$steps, type="l", xlab="Interval", ylab="Number of Steps",main="Average Number of Steps per Day by Interval")

max\_interval <- steps\_by\_interval[which.max(steps\_by\_interval$steps),1]

> max\_interval

[1] 835

> NATotal <- sum(!complete.cases(data))

> NATotal

[1] 2304

> StepsAverage <- aggregate(steps ~ interval, data = data, FUN = mean)

> fillNA <- numeric()

> for (i in 1:nrow(data)) {

+ obs <- data[i, ]

+ if (is.na(obs$steps)) {

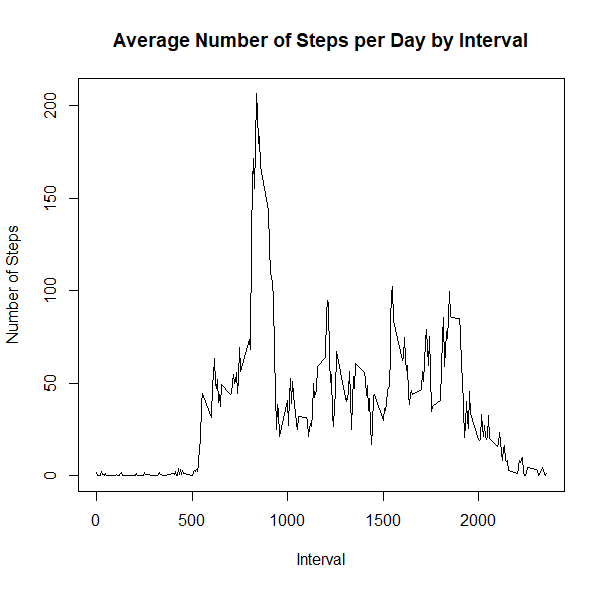
+ steps <- subset(StepsAverage, interval == obs$interval)$steps

+ } else {

+ steps <- obs$steps

+ }

+ fillNA <- c(fillNA, steps)



max\_interval <- steps\_by\_interval[which.max(steps\_by\_interval$steps),1]

> max\_interval

[1] 835

> NATotal <- sum(!complete.cases(data))

> NATotal

[1] 2304

> StepsAverage <- aggregate(steps ~ interval, data = data, FUN = mean)

> fillNA <- numeric()

> for (i in 1:nrow(data)) {

+ obs <- data[i, ]

+ if (is.na(obs$steps)) {

+ steps <- subset(StepsAverage, interval == obs$interval)$steps

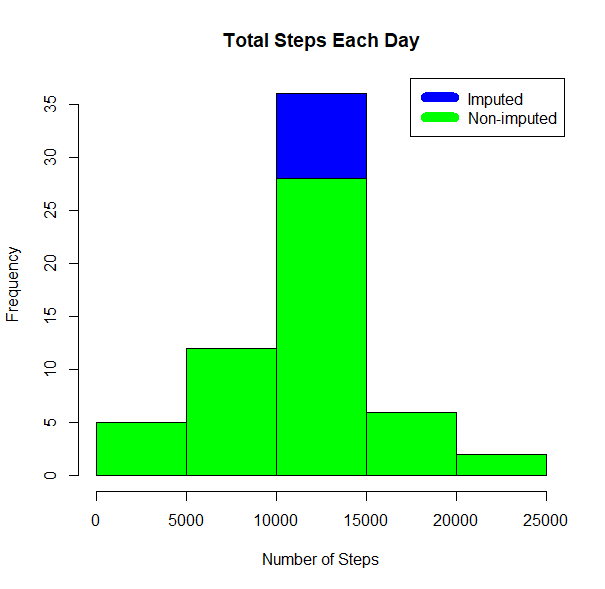
+ } else {

+ steps <- obs$steps

+ }

+ fillNA <- c(fillNA, steps)

|  |
| --- |
| StepsAverage <- aggregate(steps ~ interval, data = data, FUN = mean)  > fillNA <- numeric()  > for (i in 1:nrow(data)) {  + obs <- data[i, ]  + if (is.na(obs$steps)) {  + steps <- subset(StepsAverage, interval == obs$interval)$steps  + } else {  + steps <- obs$steps  + }  + fillNA <- c(fillNA, steps)  + }  >  > new\_activity <- data  > new\_activity$steps <- fillNA  > StepsTotalUnion <- aggregate(steps ~ date, data = new\_activity, sum, na.rm = TRUE)  > hist(StepsTotalUnion$steps, main = paste("Total Steps Each Day"), col="blue", xlab="Number of Steps")  > #Create Histogram to show difference.  > hist(steps\_by\_day$steps, main = paste("Total Steps Each Day"), col="green", xlab="Number of Steps", add=T)  > legend("topright", c("Imputed", "Non-imputed"), col=c("blue", "green"), lwd=10) |
|  |
| |  | | --- | | > | |



|  |
| --- |
| rmeantotal <- mean(StepsTotalUnion$steps)  > rmeantotal  [1] 10766.19  > rmediantotal <- median(StepsTotalUnion$steps)  > rmediantotal  [1] 10766.19  > rmediandiff <- rmediantotal - rmedian  > rmediandiff  [1] 1.188679  > rmeandiff <- rmeantotal - rmean  > rmeandiff  [1] 0 |
|  |
| |  | | --- | | > | |
| > weekdays <- c("Monday", "Tuesday", "Wednesday", "Thursday",  + "Friday")  > new\_activity$dow = as.factor(ifelse(is.element(weekdays(as.Date(new\_activity$date)),weekdays), "Weekday", "Weekend"))  > StepsTotalUnion <- aggregate(steps ~ interval + dow, new\_activity, mean)  > library(lattice)  > xyplot(StepsTotalUnion$steps ~ StepsTotalUnion$interval|StepsTotalUnion$dow, main="Average Steps per Day by Interval",xlab="Interval", ylab="Steps",layout=c(1,2), type="l") |
|  |
| |  | | --- | | > | |

