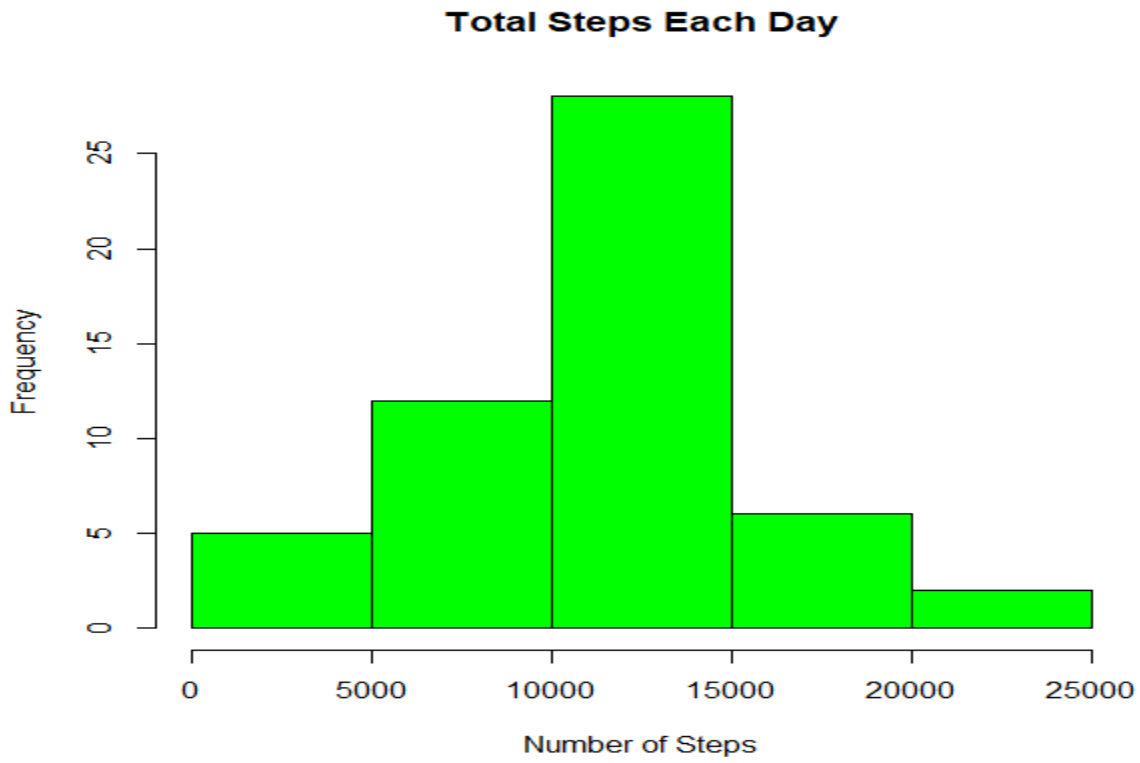
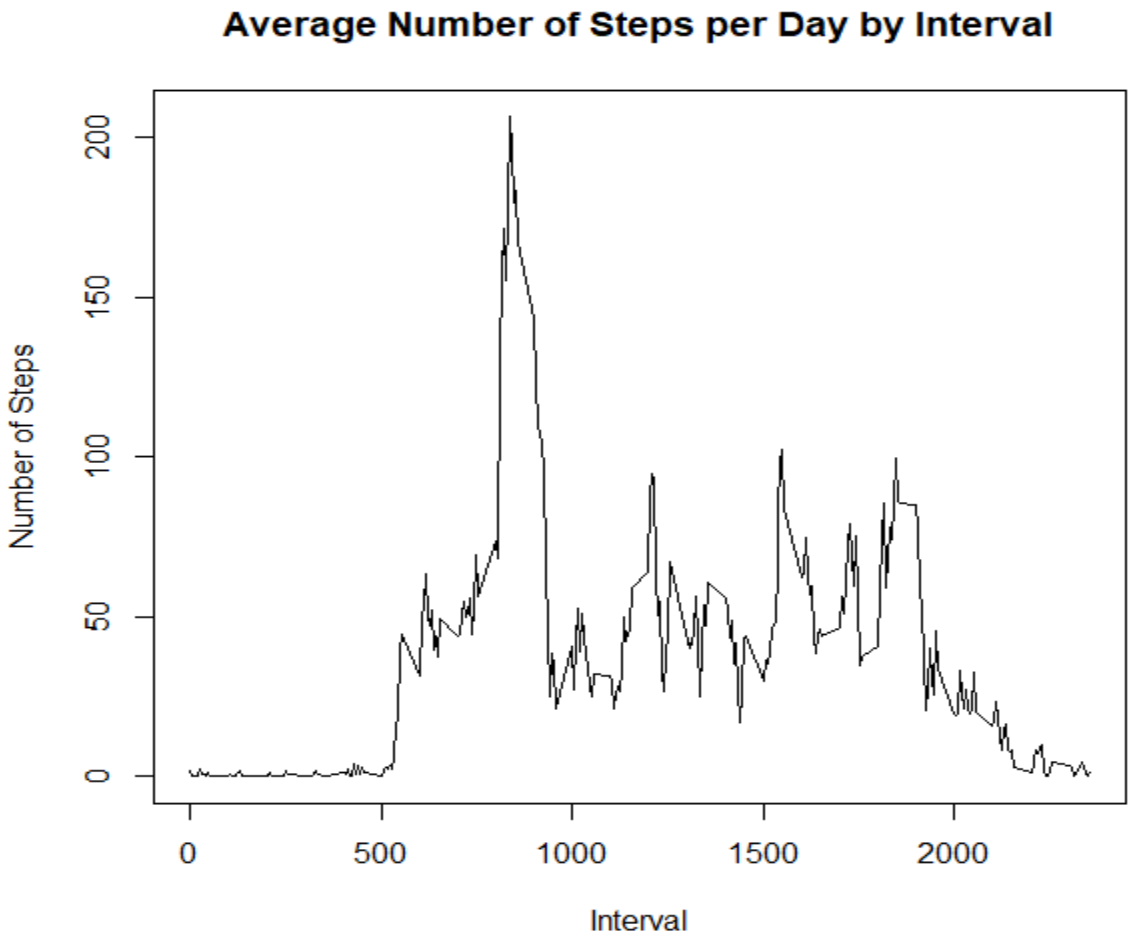
```
getwd()
[1] "C:/Users/user/Desktop"
> setwd("C:/Users/user/Desktop/New folder (3)")
> # Set Global Echo = On
>
> # Load data
> if (!file.exists("activity.csv") )
+ {
+     dlurl <- 'http://d396qusza40orc.cloudfront.net/repdata%2Fdata%2Factivity.zip'
+     download.file(dlurl,destfile='repdata%2Fdata%2Factivity.zip',mode='wb')
+     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="Numbe
r of Steps")

>
```



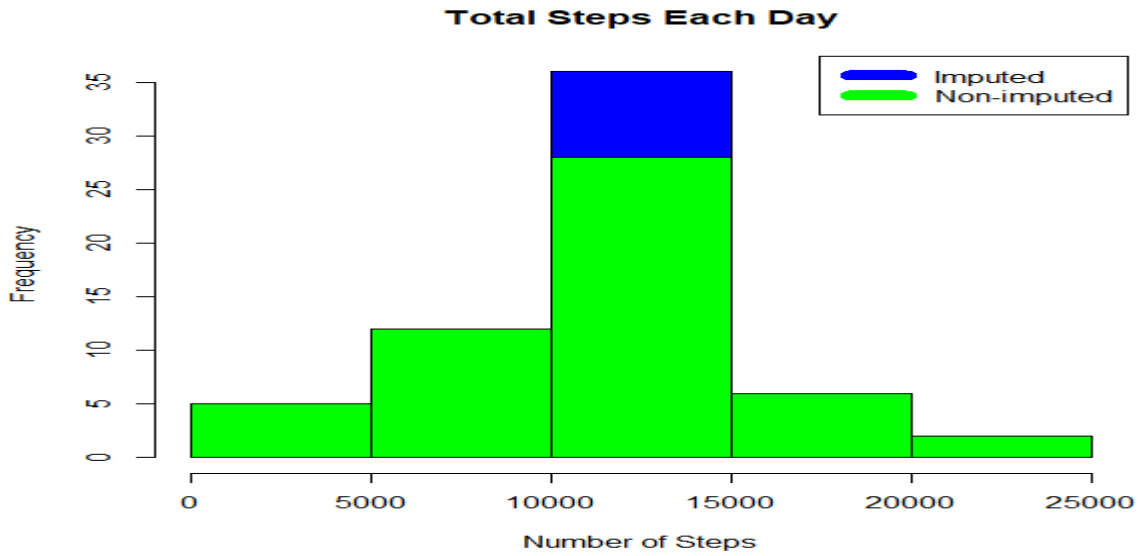

```
rmean <- mean(steps_by_day$steps)
> rmean
[1] 10766.19
> rmedian <- median(steps_by_day$steps)
> rmedian
[1] 10765
> steps_by_interval <- aggregate(steps ~ interval, data, mean)
> plot(steps_by_interval$interval,steps_by_interval$steps, type="l", xlab="In
terval", ylab="Number of Steps",main="Average Number of Steps per Day by Inte
rval")
```

```
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$steps, main = paste("Total Steps Each Day"), col="blue", xlab="Nu
mber of Steps")
 > #Create Histogram to show difference.
 > hist(steps_by_day$steps, main = paste("Total Steps Each Day"), col="green", xlab="Numb
er 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)),we
ekdays), "Weekday", "Weekend"))
> StepsTotalUnion <- aggregate(steps ~ interval + dow, new_activity, mean)
> library(lattice)
> xyplot(StepsTotalUnion$steps ~ StepsTotalUnion$interval|StepsTotalUnion$dow, main="Ave
rage Steps per Day by Interval",xlab="Interval", ylab="Steps",layout=c(1,2), type="l")

>
```