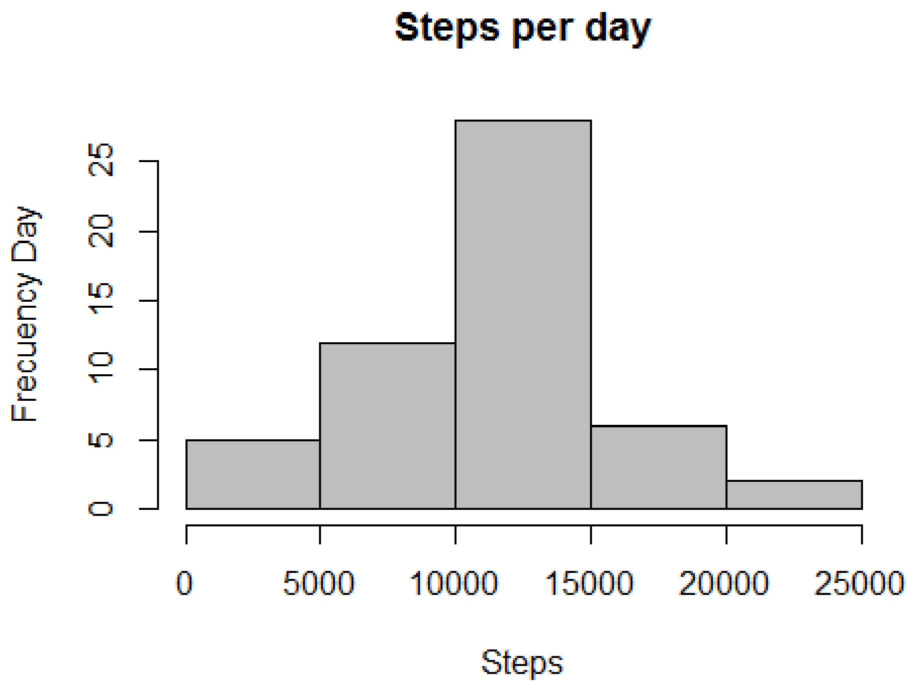
Project Markdown

Loading and preprocessing the data

```
if(!file.exists("getdata-projectfiles-UCI HAR Dataset.zip")) {
  Data <- tempfile()

download.file("http://d396qusza40orc.cloudfront.net/repdata%2Fdata%2Factivity.zip",Data)
    unzip(Data)
    unlink(Data)}
activity <- read.csv("activity.csv")
```

What is mean total number of steps taken per day

```
stepsday<-aggregate(steps~date,activity, sum)

hist(stepsday$steps, main="Steps per day", xlab="Steps", ylab =
"Frecuency Day", col="grey")
```



Mean steps by day

```
mean_steps <- mean(stepsday$steps)
```

Median steps by day

```
median_steps <- median(stepsday$steps)
```

The mean is

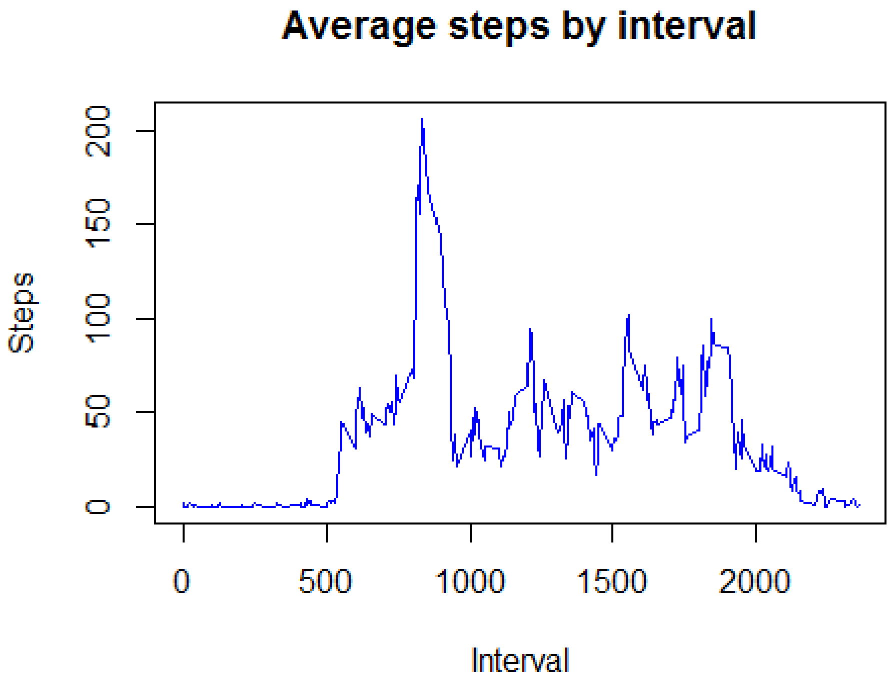
```
## [1] 10766.19
```

The median is

```
## [1] 10765
```

What is the average daily activity pattern?

```
steps_average<- aggregate(steps~interval, activity, mean)

plot(steps_average$interval,steps_average$steps, type ="l", col="blue",
xlab = "Interval", ylab="Steps", main ="Average steps by interval")
```



Interval that contains the maximum number of steps On average across all the days in the dataset

```
max_steps_average <- steps_average[which.max(steps_average$steps),1]
```

The 5-minute interval, on average across all the days in the data set, containing the maximum number of steps is

```
## [1] 835
```

Imputing missing values

Create a new dataset that is equal to the original dataset but with the missing data filled in.

number of missing values in the data

```
sum(!complete.cases(activity))

## [1] 2304
```

Devise a strategy for filling in all of the missing values in the dataset.

```
activity_imputed <- transform(activity,steps =
ifelse(is.na(activity$steps),
steps_average$steps[match(activity$interval,steps_average$interval)],acti
vity$steps))
```

the strategy is to assign the average to the intervals for the steps that are equal to NA for the first date 2012-10-01

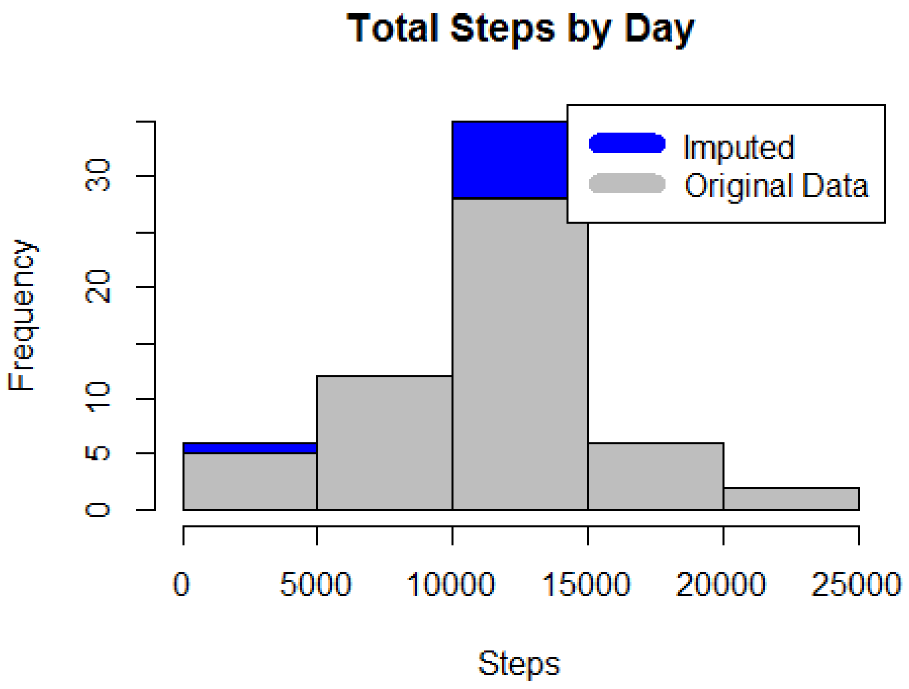
```
activity_imputed[as.character(activity_imputed$date) == "2012-10-01", 1]
<- 0
```

histogram of the total number of steps taken each day and Calculate and report the mean and total number of steps taken per day

```
stepsday2 <- aggregate(steps~date ,activity_imputed,sum)
```

Data vs imputed data through histrogram

```
stepsday_imputed <- aggregate(steps ~ date, activity_imputed, sum)
hist(stepsday_imputed$steps, main = "Total Steps by Day", col="blue",
xlab="Steps")

#Histogram that show difference.
hist(stepsday$steps, main = paste("Total Steps by Day"), col="grey",
xlab="Steps", add=T)
legend("topright", c("Imputed", "Original Data"), col=c("blue", "grey"),
lwd=10)
```



#### Values for imputed data

- Mean

```
mean_imput <- mean(stepsday_imputed$steps)
mean_imput

## [1] 10589.69
```

- Median

```
median_imput <- median(stepsday_imputed$steps)
median_imput

## [1] 10766.19
```

- Difference

The Diference in steps imputed vs steps original data is:

```
diff_steps <-( sum(stepsday_imputed$steps) - sum(stepsday$steps))
diff_steps

## [1] 75363.32
```

The Diference in mean imputed vs mean steps original data is:

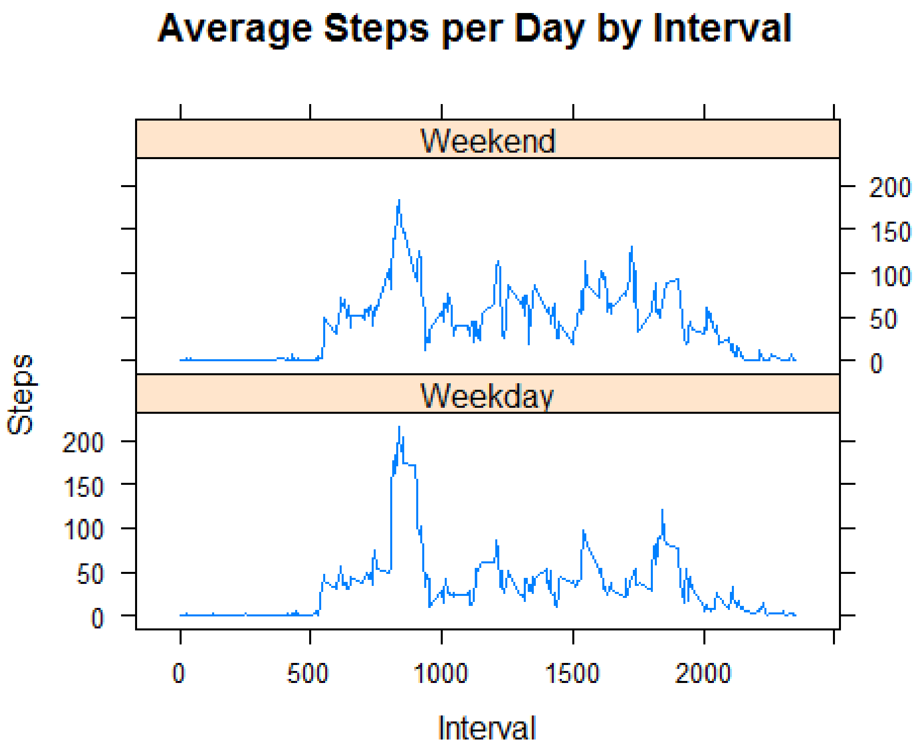
```
mean_diff_steps <-( sum(mean_imput) - sum(mean_steps))
mean_diff_steps

## [1] -176.4949
```

The Diference in median imputed vs median steps original data is:

```
median_diff_steps <-( sum(median_imput) - sum(median_steps))
median_diff_steps

## [1] 1.188679
```

Differences in activity patterns between weekdays and weekends

```
#Define the name days (It's in spanish because I'm in Colombia)
weekdays <- c("lunes", "martes", "miercoles", "jueves","viernes")

activity_imputed$typeday=as.factor(ifelse(is.element(weekdays(as.Date(act
ivity_imputed$date)),weekdays), "Weekday", "Weekend"))



steps_week_interval<- aggregate(steps~interval + typeday,
activity_imputed, mean)

library(lattice)

xyplot(steps_week_interval$steps ~
steps_week_interval$interval|steps_week_interval$typeday, main="Average
Steps per Day by Interval",xlab="Interval", ylab="Steps",layout=c(1,2),
type="l")
```



```
head(steps_week_interval$type)
```

```
## [1] Weekday Weekday Weekday Weekday Weekday Weekday
## Levels: Weekday Weekend
```