Reproducible Research: Peer Assessment 1

## Loading and preprocessing the data

We assume that the data file is loaded and stored in the current directory.

unzip("activity.zip","activity.csv")  
activity<-read.csv("activity.csv")

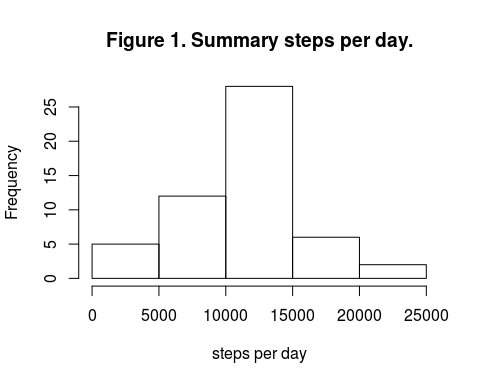
Now we convert activity$date in Date class. Because activity$date is factor we convert it indirectly.

activity$date<-as.Date(  
 as.character(activity$date),  
 format="%Y-%m-%d"  
)

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

Now we summerazing total number steps per day and plot histogram. The aggregate function splited dataframe by date and summarized steps into each split. Results are saved in agg dataframe.

agg<-aggregate(steps ~ date, data=activity, sum)  
hist(agg$steps, xlab = "steps per day", main="Figure 1. Summary steps per day.")



Mean steps per day is:

mean(agg$steps)

## [1] 10766.19

Median steps per day is:

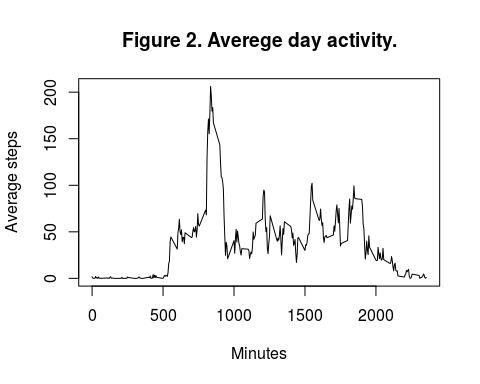
median(agg$steps)

## [1] 10765

## What is the average daily activity pattern?

We calculate averege number of step per every 5-minutes interval for all days.

typical\_day<-aggregate(steps ~ interval, data=activity, mean)  
plot(typical\_day$interval,typical\_day$steps, type="l", main ="Figure 2. Averege day activity.", ylab ="Average steps", xlab = "Minutes" )



This 5-minute interval contains the maximum number of step:

typical\_day$interval[which.max(typical\_day$steps)]

## [1] 835

## Imputing missing values

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