Reproducible Research: Peer Assessment 1

## Loading and preprocessing the data

Set the working directory and import the data from the .csv file

# fileurl is here  
fileurl <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2Factivity.zip"  
  
# create a temporary directory  
temp = tempdir()  
  
# create the placeholder ready to receive  
acttemp = tempfile(tmpdir=temp, fileext=".zip")  
  
# download into the placeholder  
download.file(fileurl, acttemp)  
  
# get the name of the first file in the zip archive  
activity = unzip(acttemp, list=TRUE)$Name[1]  
  
# unzip the file to the temporary directory  
unzip(acttemp, files=activity, exdir=temp, overwrite=TRUE)  
  
# fpath is the full path to the extracted file  
fpath = file.path(temp, activity)  
  
# load the dataframe  
activity <- read.csv("activity.csv", as.is=TRUE)

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

# For this part of the assignment, you can ignore the missing values in the dataset.  
activitynona <- na.omit(activity)  
  
# Make a histogram of the total number of steps taken each day  
  
dailystepstot <- aggregate(steps ~ date, activitynona, sum)  
  
hist(dailystepstot$steps, col=1, main="Total number of steps per day",   
 xlab="Total number of per day")

# Calculate and report the mean and median total number of steps taken per day  
mean(dailystepstot$steps)

median(dailystepstot$steps)

## What is the average daily activity pattern?

## Imputing missing values

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