Reproducible Research: Peer Assesment 1

Matthew Pang

August 2, 2016

## Reproducible Research: Peer Assignment 1

# Loading and preprocessing the data

data <- read.csv("activity.csv")

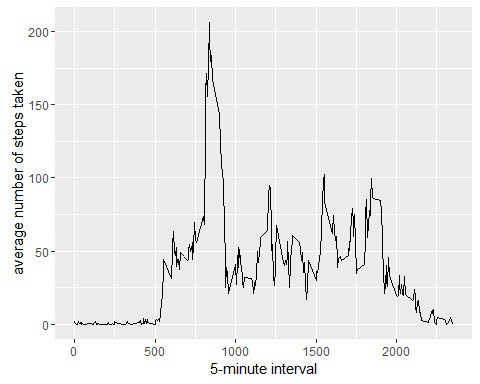
# What is mean total number of steps taken per day?



## [1] 9354.23

## [1] 10395

# What is the average daily activity pattern?

 Which 5-minute interval, on average across all the days in the dataset, contains the maximum number of steps?

averages[which.max(averages$steps),]

## interval steps  
## 104 835 206.1698

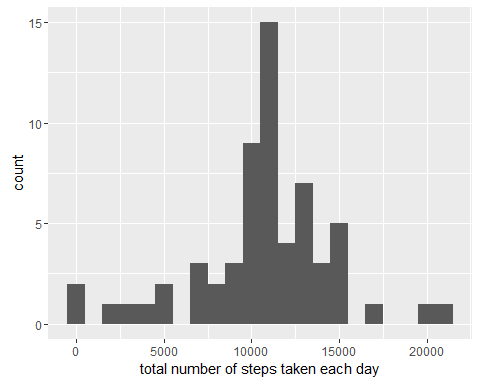
# Imputing missing values

Calculate and report the total number of missing values in the dataset (i.e. the total number of rows with NAs)

## missing  
## FALSE TRUE   
## 15264 2304

Missing Value filling in with mean value with 5-minute interval

fill.value <- function(steps, interval) {  
 filled <- NA  
 if (!is.na(steps))  
 filled <- c(steps)  
 else  
 filled <- (averages[averages$interval==interval, "steps"])  
 return(filled)  
}  
filled.data <- data  
filled.data$steps <- mapply(fill.value, filled.data$steps, filled.data$interval)

Histogram of total number of steps 

## [1] 10766.19

## [1] 10766.19

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

Fill in Weekday or weekend

weekday.or.weekend <- function(date) {  
 day <- weekdays(date)  
 if (day %in% c("Monday", "Tuesday", "Wednesday", "Thursday", "Friday"))  
 return("weekday")  
 else if (day %in% c("Saturday", "Sunday"))  
 return("weekend")  
 else  
 stop("invalid date")  
}  
filled.data$date <- as.Date(filled.data$date)  
filled.data$day <- sapply(filled.data$date, FUN=weekday.or.weekend)

Plot of average number of steps taken on weekdays and weekends 