# Peer Assessment 1

HE Ying Daniel

## Loading and preprocessing the data

setwd("C:\\Temp\\test")  
data0<-read.csv("activity.csv",colClasses = c("numeric", "character","numeric"))  
good<-complete.cases(data0)  
data1<-data0[good,]

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

library(ggplot2)  
aggsteps<-aggregate(steps~date,data1,sum)  
mean(aggsteps$steps)

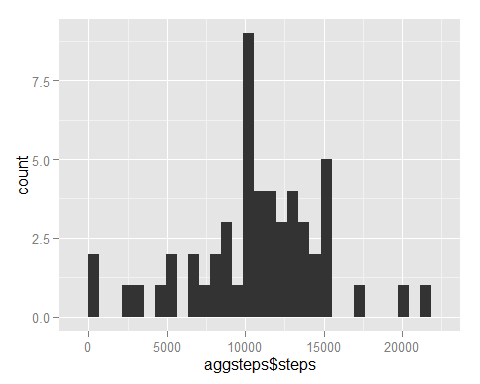
## [1] 10766

median(aggsteps$steps)

## [1] 10765

qplot(aggsteps$steps,data=aggsteps)

## stat\_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.

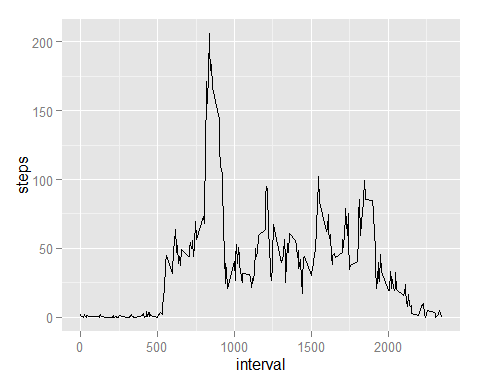


## What is the average daily activity pattern?

meaninterval<-aggregate(steps~interval,data1,mean)  
meaninterval[which.max(meaninterval$steps),]

## interval steps  
## 104 835 206.2

g<-ggplot(meaninterval,aes(x=interval,y=steps))  
g+geom\_line()



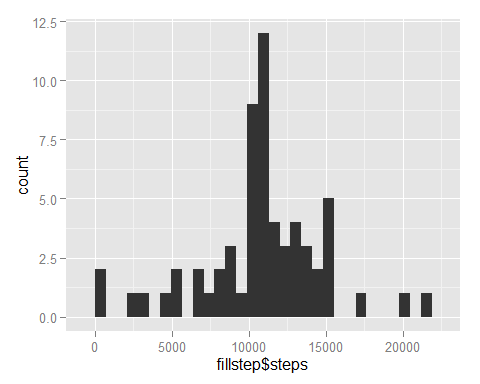
## Imputing missing values

mdata<-data0[!good,]  
nrow(mdata) ## row number for "NA"

## [1] 2304

meansteps<-aggregate(steps~interval,data=data0,FUN=mean)  
  
for(i in 1:nrow(data0)){  
 if(is.na(data0$steps[i])){  
 inver<-data0$interval[i]  
 rowid<-which(meansteps$interval==inver)  
 step<-meansteps$steps[rowid]  
 data0$steps[i]<-step  
 }   
 }   
   
fillstep<-aggregate(steps~date,data=data0,sum,na.rm=TRUE)  
qplot(fillstep$steps,data=fillstep)

## stat\_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.



mean(fillstep$steps)

## [1] 10766

median(fillstep$steps)

## [1] 10766

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

Sys.setlocale("LC\_TIME","Chinese")

## [1] "Chinese (Simplified)\_China.936"

weekdays(Sys.Date()+0:6)

## [1] "星期日" "星期一" "星期二" "星期三" "星期四" "星期五" "星期六"

Sys.setlocale("LC\_TIME","English")

## [1] "English\_United States.1252"

weekdays(Sys.Date()+0:6)

## [1] "Sunday" "Monday" "Tuesday" "Wednesday" "Thursday" "Friday"   
## [7] "Saturday"

data0$date<-as.Date(data0$date,"%Y-%m-%d")  
data0$day<-weekdays(data0$date)  
data0$type<-c("weekday")  
for(i in 1:nrow(data0)){  
 if(data0$day[i]=="Saturday"||data0$day[i]=="Sunday"){  
 data0$type[i]<-"weekend"  
 } else{  
 data0$type[i]<-"weekday"  
 }  
}  
  
data0$type<-as.factor(data0$type)  
tabledata<-aggregate(steps~interval+type,data0,mean)  
qplot(interval,steps,data=tabledata,geom=c("line"),xlab="Interval",ylab="Number of Steps",main="")+facet\_wrap(~type,ncol=1)

