Loading and preprocessing the data

```
r   library(dplyr)

## Warning: package 'dplyr' was built under R version 3.3.3

##    ## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':   ##    ##
filter, lag

## The following objects are masked from 'package:base':   ##    ##
intersect, setdiff, setequal, union

r   library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.3.3

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

## 'data.frame':  17568 obs. of  3 variables:   ##  $ steps    : int  NA NA NA
NA NA NA NA NA NA NA ...   ##  $ date     : Factor w/ 61 levels "2012-10-
01","2012-10-02",..: 1 1 1 1 1 1 1 1 1 1 ...   ##  $ interval: int  0 5 10 15
20 25 30 35 40 45 ...

r   summary(activity.data)

##      steps                date          interval        ##  Min.   :  0.00
2012-10-01:  288   Min.   :   0.0     ##  1st Qu.:  0.00   2012-10-02:  288
1st Qu.: 588.8     ##  Median :  0.00   2012-10-03:  288   Median :1177.5
##  Mean   : 37.38   2012-10-04:  288   Mean   :1177.5     ##  3rd Qu.: 12.00
2012-10-05:  288   3rd Qu.:1766.2     ##  Max.   :806.00   2012-10-06:  288
Max.   :2355.0     ##  NA's   :2304     (Other)   :15840

r   head(activity.data,3)

##   steps       date interval   ## 1    NA 2012-10-01        0   ## 2    NA
2012-10-01        5   ## 3    NA 2012-10-01       10

r   act.data.completecase <- na.omit(activity.data)
head(act.data.completecase,3)

##     steps       date interval   ## 289     0 2012-10-02        0   ## 290
0 2012-10-02        5   ## 291     0 2012-10-02       10
```

What is mean total number of steps taken per day?

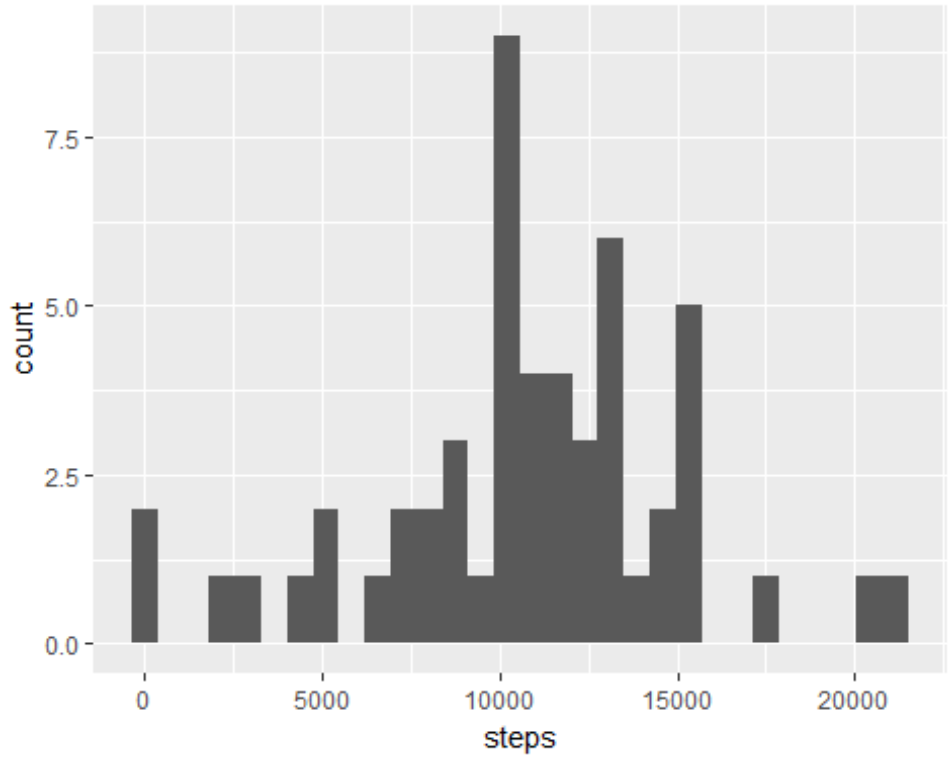
```
activity.day <- group_by(act.data.completecase, date)
activity.day <- summarize(activity.day, steps=sum(steps))
summary(activity.day)

##          date        steps
##  2012-10-02: 1   Min.   :   41
##  2012-10-03: 1   1st Qu.: 8841
```

```
##  2012-10-04: 1   Median :10765
##  2012-10-05: 1   Mean   :10766
##  2012-10-06: 1   3rd Qu.:13294
##  2012-10-07: 1   Max.   :21194
##  (Other)   :47
```

```
qplot(steps, data=activity.day)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
mean(activity.day$steps)
```

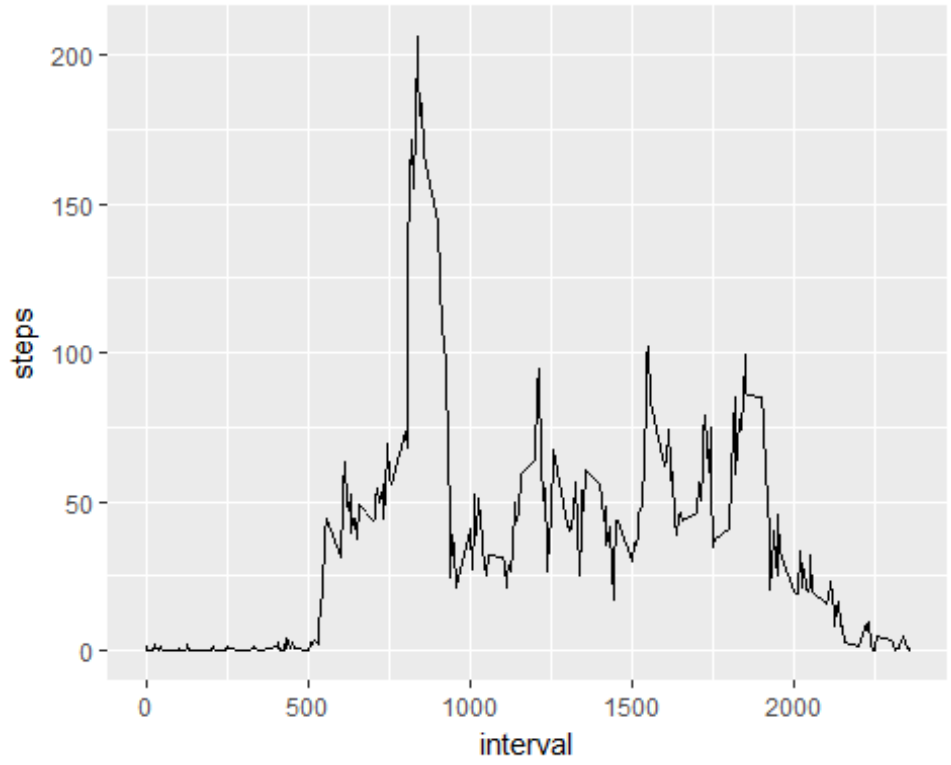
```
## [1] 10766.19
```

```
median(activity.day$steps)
```

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

What is the average daily activity pattern?

```
activity.interval <- group_by(act.data.completecase, interval)
activity.interval <- summarize(activity.interval, steps=mean(steps))
ggplot(activity.interval, aes(interval, steps)) + geom_line()
```




```
activity.interval[activity.interval$steps==max(activity.interval$steps),]
```

```
## # A tibble: 1 x 2
##   interval    steps
##      <int>    <dbl>
## 1      835 206.1698
```

Imputing missing values

```
nrow(activity.data)-nrow(act.data.completecase)
```

```
## [1] 2304
```

```
names(activity.interval)[2] <- "mean.steps"
activity.impute <- merge(activity.data, activity.interval)
activity.impute$steps[is.na(activity.impute$steps)] <-
activity.impute$mean.steps[is.na(activity.impute$steps)]
```

Are there differences in activity patterns between weekdays and weekends?

```
activity.impute$dayofweek <- weekdays(as.Date(activity.impute$date))
activity.impute$weekend <-
as.factor(activity.impute$dayofweek=="Saturday"|activity.impute$dayofweek=="S
unday")
levels(activity.impute$weekend) <- c("Weekday", "Weekend")

activity.weekday <- activity.impute[activity.impute$weekend=="Weekday",]
```

```r
activity.weekend <- activity.impute[activity.impute$weekend=="Weekend",]

act.int.weekday <- group_by(activity.weekday, interval)
act.int.weekday <- summarize(act.int.weekday, steps=mean(steps))
act.int.weekday$weekend <- "Weekday"
act.int.weekend <- group_by(activity.weekend, interval)
act.int.weekend <- summarize(act.int.weekend, steps=mean(steps))
act.int.weekend$weekend <- "Weekend"

activity.interval <- rbind(act.int.weekday, act.int.weekend)
activity.interval$weekend <- as.factor(activity.interval$weekend)
ggplot(activity.interval, aes(interval, steps)) + geom_line() +
facet_grid(weekend ~ .)
```