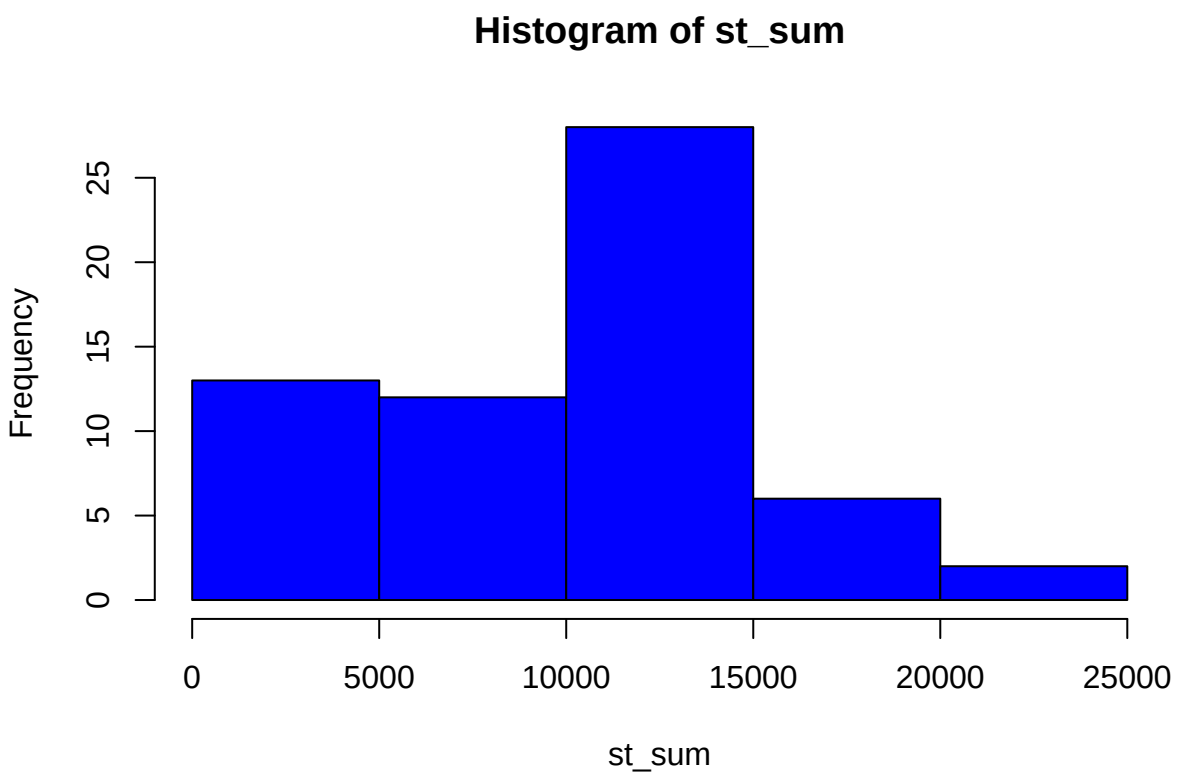
Reproducible Research: Peer Assessment 1

Loading and preprocessing the data

```
data<-read.csv('activity.csv',colClasses=c('numeric','Date','numeric'))
```

What is mean total number of steps taken per day?

```
st_sum<-tapply(data$steps,data$date,sum,na.rm=TRUE)
hist(st_sum,col='blue')
```




```
mean(st_sum)
```

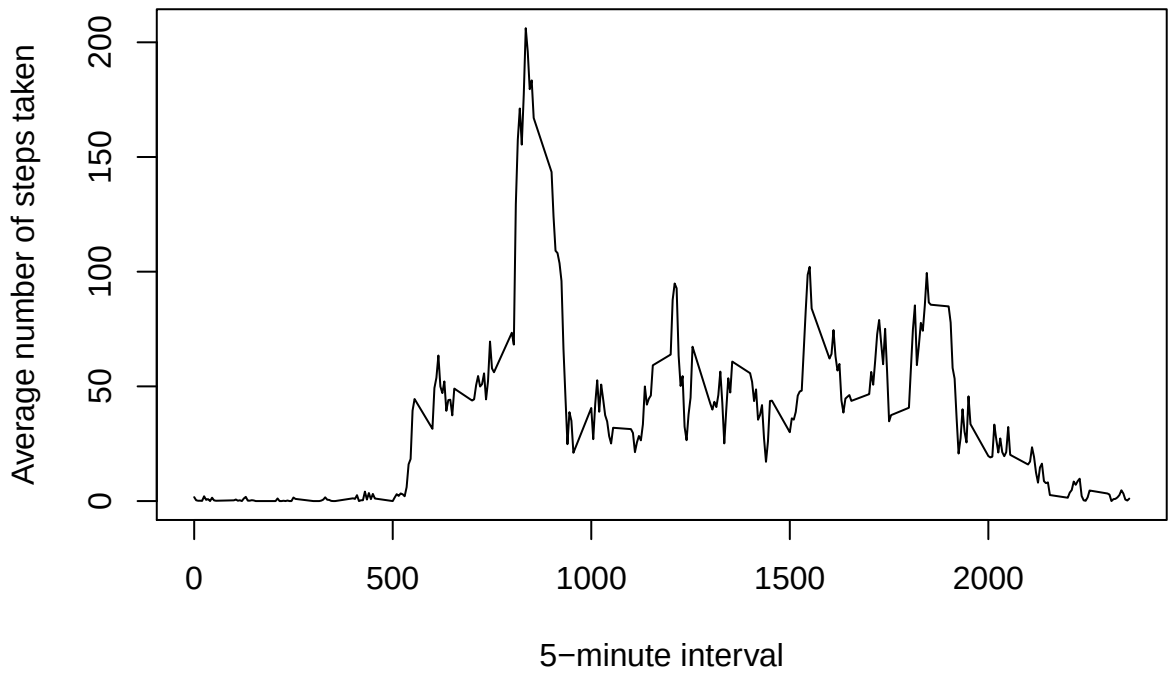
```
## [1] 9354.23
```

```
median(st_sum)
```

```
## [1] 10395
```

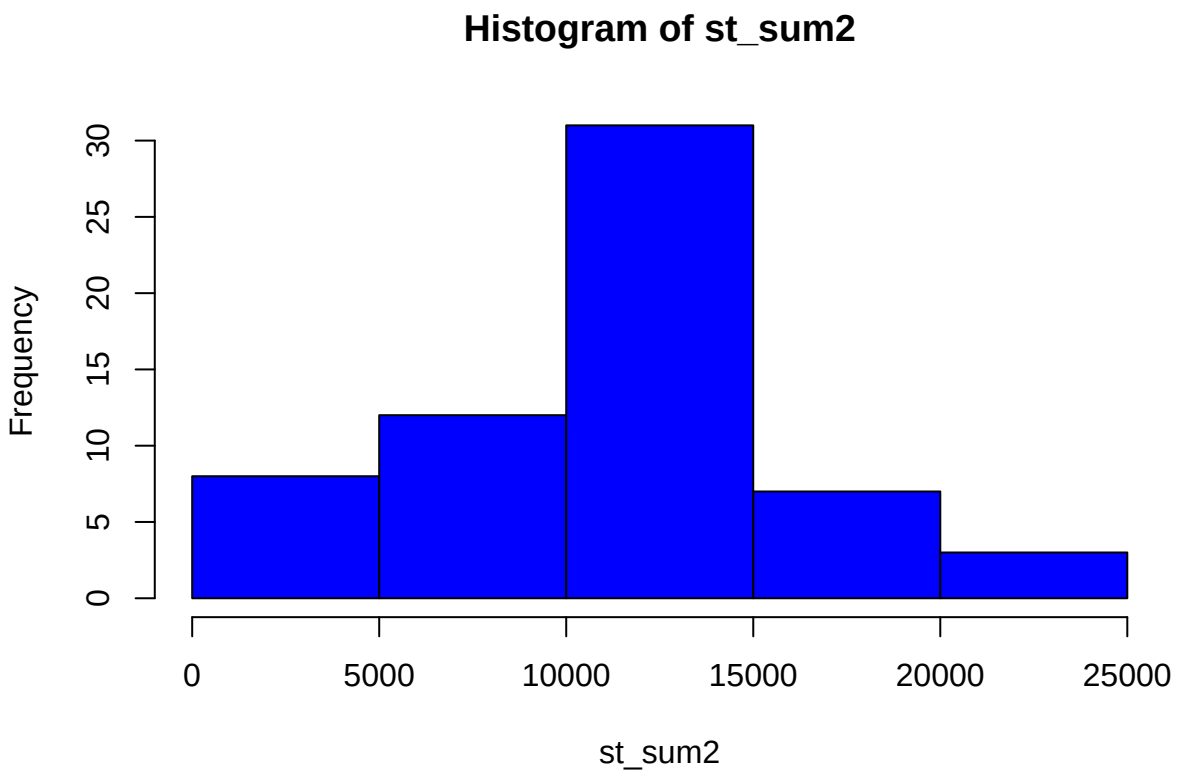
What is the average daily activity pattern?

```
st_mean<-tapply(data$steps,data$interval,mean,na.rm=TRUE)
x<-data$interval[1:288]
plot(x,st_mean,type='l',xlab='5-minute interval',ylab='Average number of steps taken')
```



Imputing missing values

```
miss<-sum(is.na(data))
index<-which(is.na(data))
df<-data.frame(interval=x,mean=st_mean)
u=merge(df,data[index,],by='interval')
data$steps[index]<-u$mean
st_sum2<-tapply(data$steps,data$date,sum,na.rm=TRUE)
hist(st_sum2,col='blue')
```




```
mean(st_sum2)
```

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

```
median(st_sum2)
```

```
## [1] 11015
```

Are there differences in activity patterns between weekdays and weekends?

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following object is masked from 'package:stats':
##
##     filter
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(timeDate)
newdata<-mutate(data,week=factor(isWeekday(date)))
levels(newdata$week)<-c('weekend','weekday')
sub1<-subset(newdata,week=='weekend')
sub2<-subset(newdata,week=='weekday')
st_mean2<-tapply(sub1$steps,sub1$interval,mean)
st_mean3<-tapply(sub2$steps,sub2$interval,mean)
par(mfrow=c(1,2))
plot(st_mean2~x,type='l',xlab='5-minute interval',ylab='Average number of steps taken',ylim=c(0,max(st_
plot(st_mean3~x,type='l',xlab='5-minute interval',ylab='Average number of steps taken',main='weekday')
```