

#Peer Assessment 1 : A.J. Arango

1. First I will read the downloaded data and load

```
#
data <- "E:/R/activity.csv"
info <- read.csv(data, header=TRUE , sep=",", stringsAsFactors = FALSE, colClasses = c("numeric","Date","numeric"))
print(str(info))
```

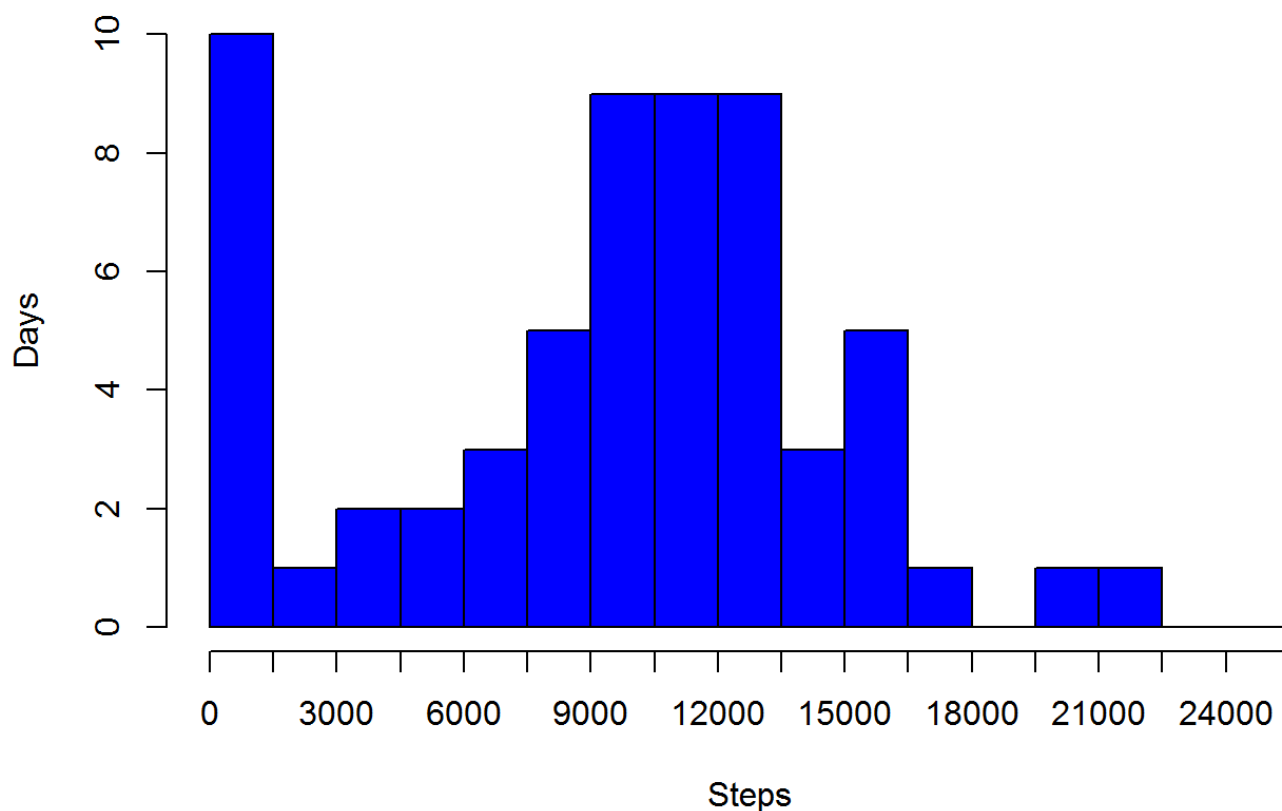
```
## 'data.frame':   17568 obs. of  3 variables:
## $ steps      : num  NA NA NA NA NA NA NA NA NA NA NA ...
## $ date       : Date, format: "2012-10-01" "2012-10-01" ...
## $ interval: num   0 5 10 15 20 25 30 35 40 45 ...
## NULL
```

2. What is the mean of the number of steps taken per day?

```
StepsInDay <- aggregate(info$steps, list(date=info$date),sum,na.rm=TRUE)

xAxisBreakdown = seq(from=0,to=26000,by=1500)
hist(StepsInDay$x,
     breaks = xAxisBreakdown,
     main="Frequency of Total Steps per Day",
     col="blue",
     xlab="Steps",
     ylab="Days",
     xaxt="n")
axis(side=1,at=xAxisBreakdown,labels=xAxisBreakdown)
```

Frequency of Total Steps per Day



##Mean steps

```
meansteps <- mean(StepsInDay$x, na.rm =TRUE)
print(paste("The Mean number of imputed steps per day is",meansteps))
```

```
## [1] "The Mean number of imputed steps per day is 9354.22950819672"
```

Median Steps

```
mediansteps <- median(StepsInDay$x, na.rm = TRUE)
print(paste("The median number of imputed steps per day is",mediansteps))
```

```
## [1] "The median number of imputed steps per day is 10395"
```

3. Average daily activity pattern

```

TimeHours <- info$interval %% 100

TimeHours <- ifelse(TimeHours < 10, paste("0", TimeHours, sep=" "), TimeHours)

TimeMinutes <- info$interval %% 100
TimeMinutes <- ifelse(TimeMinutes < 10, paste("0",TimeMinutes, sep=" "), TimeMinutes)

InterTime <- paste(TimeHours, ":", TimeMinutes, sep="")
InterTime <- strptime(InterTime, format="%H:%M")

info <- cbind(info,InterTime)

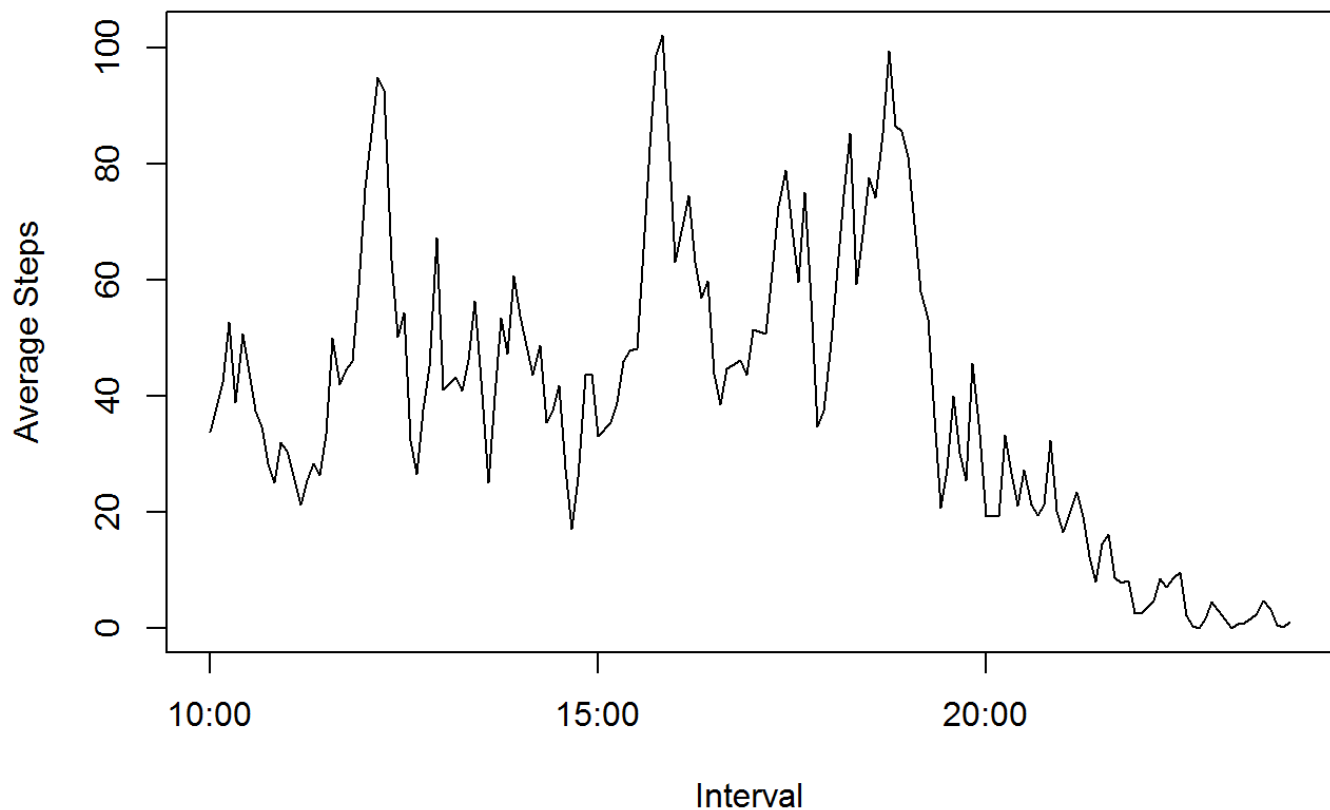
```

```

SPI <- aggregate(info$steps,list(InterTime=info$InterTime),mean,na.rm=TRUE)
plot(SPI$InterTime,SPI$x,
     type = "l",
     main = "Average Steps per Interval",
     xlab = "Interval",
     ylab = "Average Steps")

```

Average Steps per Interval



###4. Handle Missing Values ##Part a

```
CountNA <- sum(is.na(info$steps))  
print(CountNA)
```

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

Part b

Part c

```
StepsInDay <- aggregate(info$steps, list(date=info$date), sum, na.rm=TRUE)  
  
xAxisBreakdown = seq(from=0, to=26000, by=1500)  
hist(StepsInDay$x,  
      breaks = xAxisBreakdown,  
      main="Frequency of Total Steps per Day",  
      col="blue",  
      xlab="Steps",  
      ylab="Days",  
      xaxt="n")  
axis(side=1, at=xAxisBreakdown, labels=xAxisBreakdown)
```

Frequency of Total Steps per Day

