

project1_activity

weimf1990

12/05/2021

Loading and preprocessing the data

```
#setting the directory
setwd("C:/Users/user/Desktop/Data Science/project/Course 5")

#loading the data
activity <- read.csv("./repdata_data_activity/activity.csv")

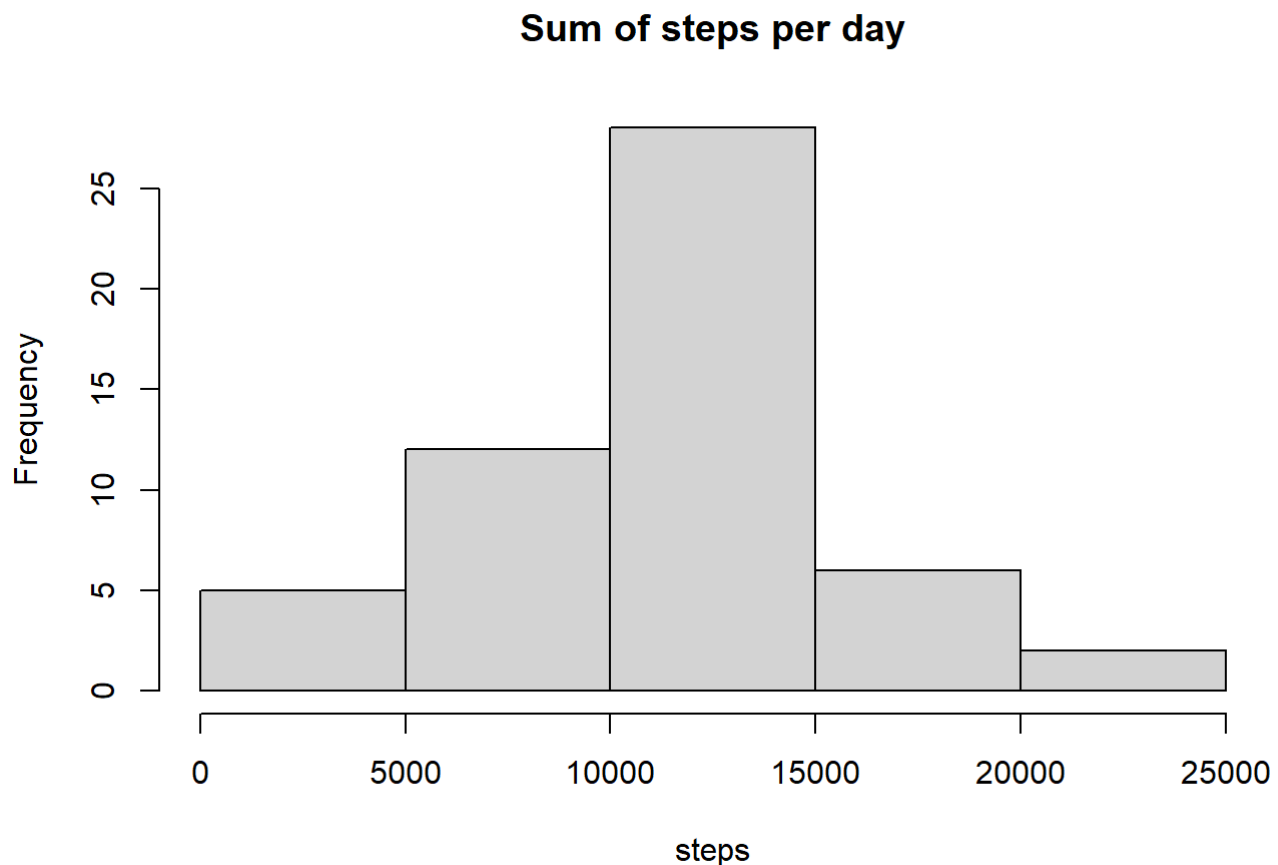
#processing the data
activity$day <- weekdays(as.Date(activity$date))
activity$DateTime<- as.POSIXct(activity$date, format="%Y-%m-%d")

##clean the data without na
data <- activity[!is.na(activity$steps),]
```

What is mean total number of steps taken per day?

Histogram of the total number of steps taken each day

Histogram of the total number of steps taken each day



Mean and median number of steps taken each day

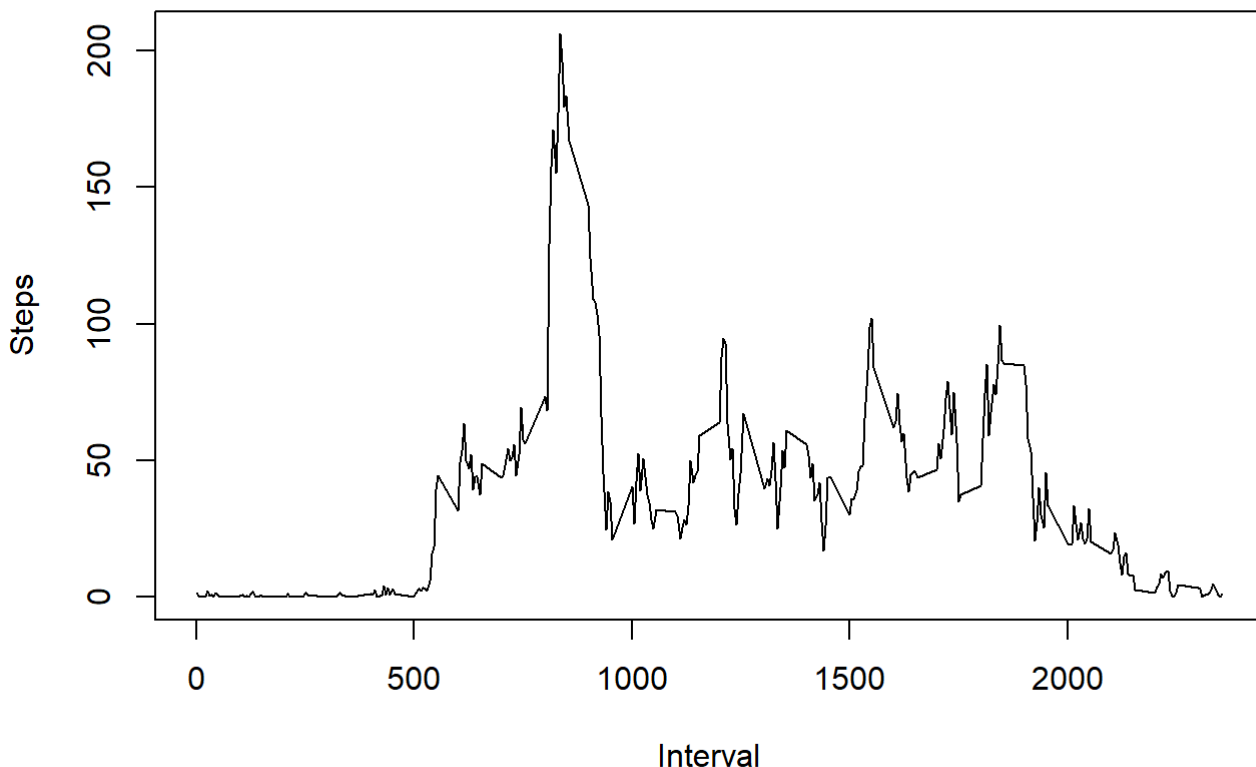
```
## [1] 10766.19
```

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

The mean and median number of steps taken each day is 10,766 and 10,765, respectively.

What is the average daily activity pattern?

Time series average number of steps taken



The 5-minute interval that, on average, contains the maximum number of steps

```
## [1] 206.1698
```

```
## [1] "835"
```

The maximum number of steps for a 5-minute interval is 206 steps.

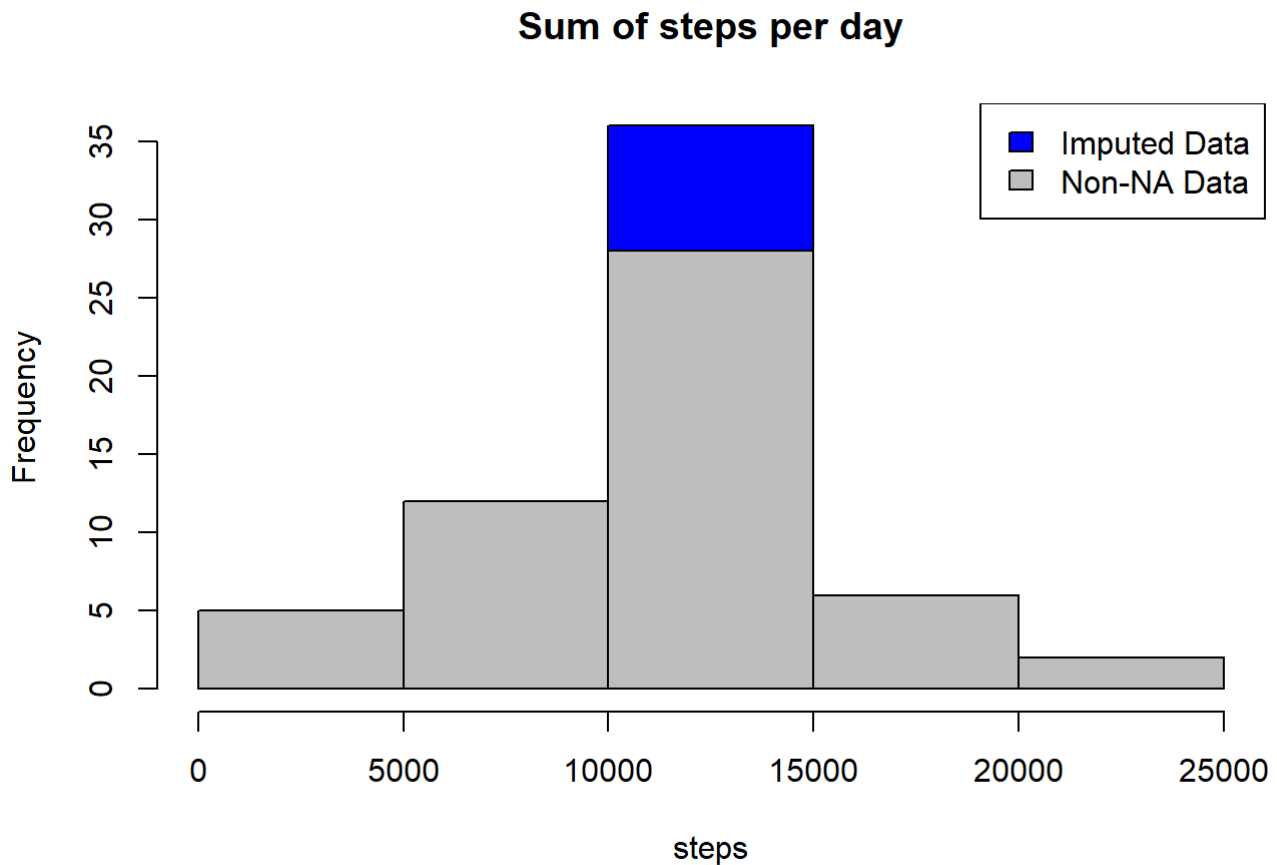
The 5-minute interval which has the maximum number of steps is the 835 interval.

Imputing missing values

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

There are 2,304 missing values in the data. I choose to impute these missing data using the mean number of steps by the particular interval across dates.

Histogram of the total number of steps taken each day after missing values are imputed



Report the mean and median total number of steps taken per day.

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

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

```
## [1] 0
```

```
## [1] 1.188679
```

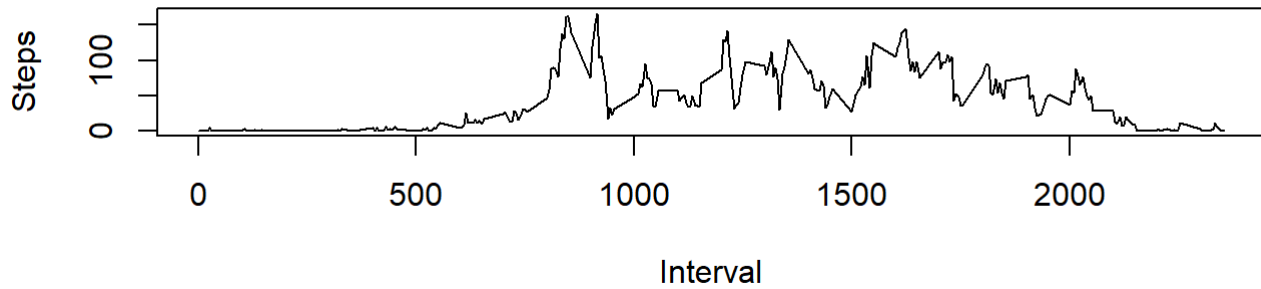
The mean remains the same as prior to imputation, and the median value increases by 1.18. Based on the figure plotted above, the difference lies in the steps ranging from 10,000 to 15,000.

Are there differences in activity patterns between weekdays and weekends?

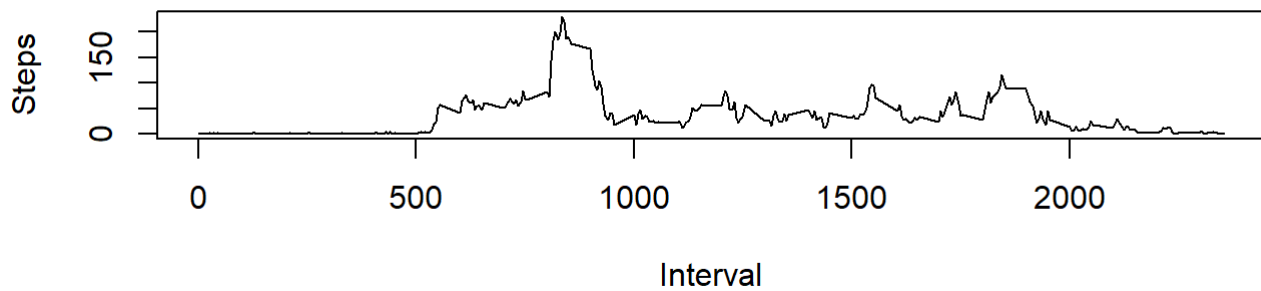
##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.000	1.241	32.340	42.366	74.654	166.639

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.000	2.247	25.803	35.611	50.854	230.378

average number of steps taken_weekend



average number of steps taken_weekday



The step activity trends are different between the weekend and weekday. Specifically, people seem to have more opportunity in activities on weekends.