

Data Visualization using MI band 4

Introduction

In this project, I am analyzing the information stored by my MI band 4. MI band captures various activities such as heart rate, steps count, calories burned, sleep, and sports activities like walking, running etc. The data is calculated from 1 November 2019 to 4 March 2021 i.e., total of 477 days of data. Data has distance in km and calories in kcal While counting the steps, and sports activities. Heart rate is measured in BPM(beats per minute). Sleep time is measured in minutes. Sleep time is divided into 3 types which are nothing but deep sleep time, shallow sleep time and wake time. Each activity has a different dataset file. Variables considered from datasets are given below-

- date
- steps
- distance: calculated in km
- calories: in kcal
- heartRate: captured every second or two in BPM
- deepSleepTime
- shallowSleepTime
- wakeTime

Preprocessing

1. Remove unnecessary columns from the data
2. Change the column names
3. Convert date from char to date object
4. Convert epoch to date object [1]
5. Convert categorical variable into factors
6. Merge activity and sleep data frames
7. Extract days from the date object [2]
8. Create a new column for weekdays and weekends
9. Calculate monthly steps and monthly sleep [3] and reshaping the data
10. Convert sports-type from number to character and then convert it into factors

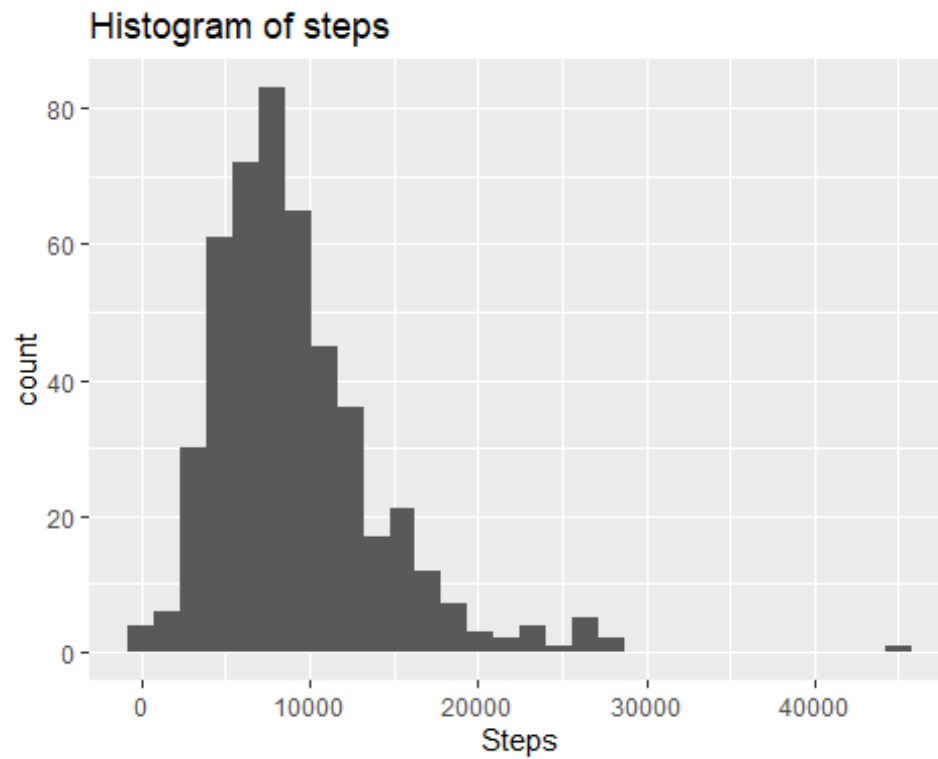
Analysis

Activity Analysis

Summary of Activity and Sleep data is

```
##    date                steps          distance    runDistance
## Min. :2019-11-12 Min. : 0          Min. : 0      Min. : 0.0
## 1st Qu.:2020-03-10 1st Qu.: 5900        1st Qu.: 3869 1st Qu.: 346.0
## Median :2020-07-09 Median : 8243        Median : 5567 Median : 455.0
## Mean   :2020-07-08 Mean   : 9159        Mean   : 6210 Mean   : 579.9
## 3rd Qu.:2020-11-05 3rd Qu.:11330        3rd Qu.: 7759 3rd Qu.: 605.0
## Max.   :2021-03-04 Max.   :44996        Max.   :30807 Max.   :5132.0
##
##    calories    deepSleepTime  shallowSleepTime  wakeTime
## Min. : 0.0      Min. : 0.00    Min. : 0.0        Min. : 0.00
## 1st Qu.: 137.0    1st Qu.: 10.00 1st Qu.:250.0      1st Qu.: 0.00
## Median : 191.0    Median : 28.00 Median :317.0      Median : 0.00
## Mean   : 216.8    Mean   : 32.14 Mean   :300.7      Mean   : 12.04
## 3rd Qu.: 269.0    3rd Qu.: 45.00 3rd Qu.:385.0      3rd Qu.: 18.00
## Max.   :1004.0    Max.   :148.00 Max.   :596.0      Max.   :120.00
##
##    start                stop          totalSleepTime    day
## Min. :1.573e+09 Min. :1.573e+09    Min. : 0.0      Friday :67
## 1st Qu.:1.584e+09 1st Qu.:1.584e+09    1st Qu.:290.0    Monday :68
## Median :1.594e+09 Median :1.594e+09    Median :372.0    Saturday:67
## Mean   :1.594e+09 Mean   :1.594e+09    Mean   :344.9    Sunday :68
```

```
## 3rd Qu.:1.605e+09 3rd Qu.:1.605e+09 3rd Qu.:441.0 Thursday :69
## Max. :1.615e+09 Max. :1.615e+09 Max. :642.0 Tuesday :69
##                                     Wednesday:69
## dayType
## Weekday:342
## Weekend:135
##
##
##
##
##
```

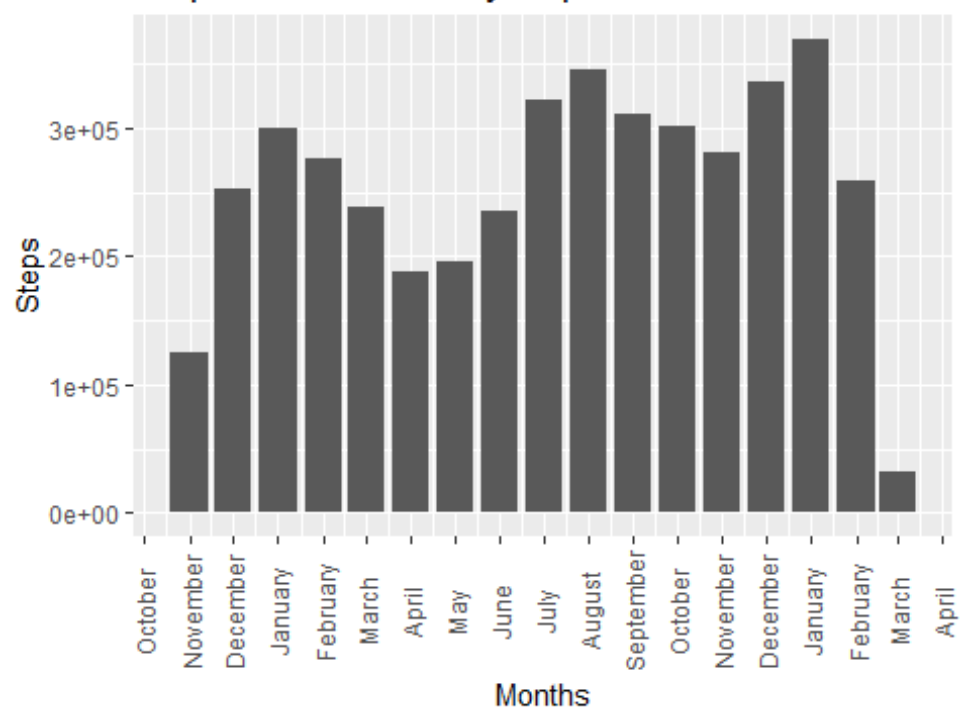


The average step count is 9159, the median is 8243 and the maximum steps count is 44996.

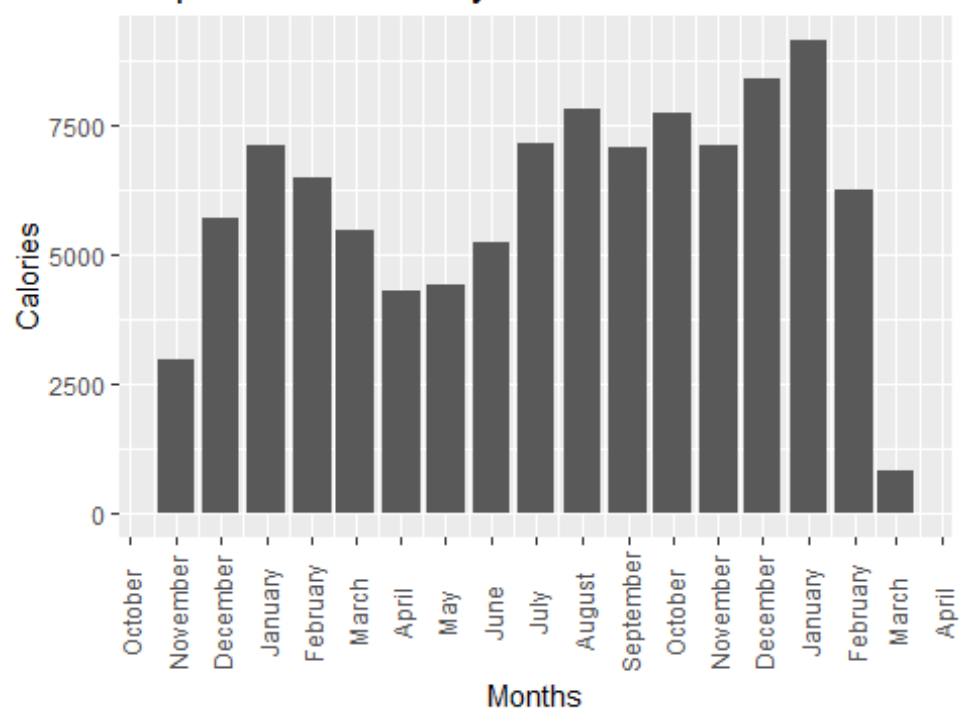


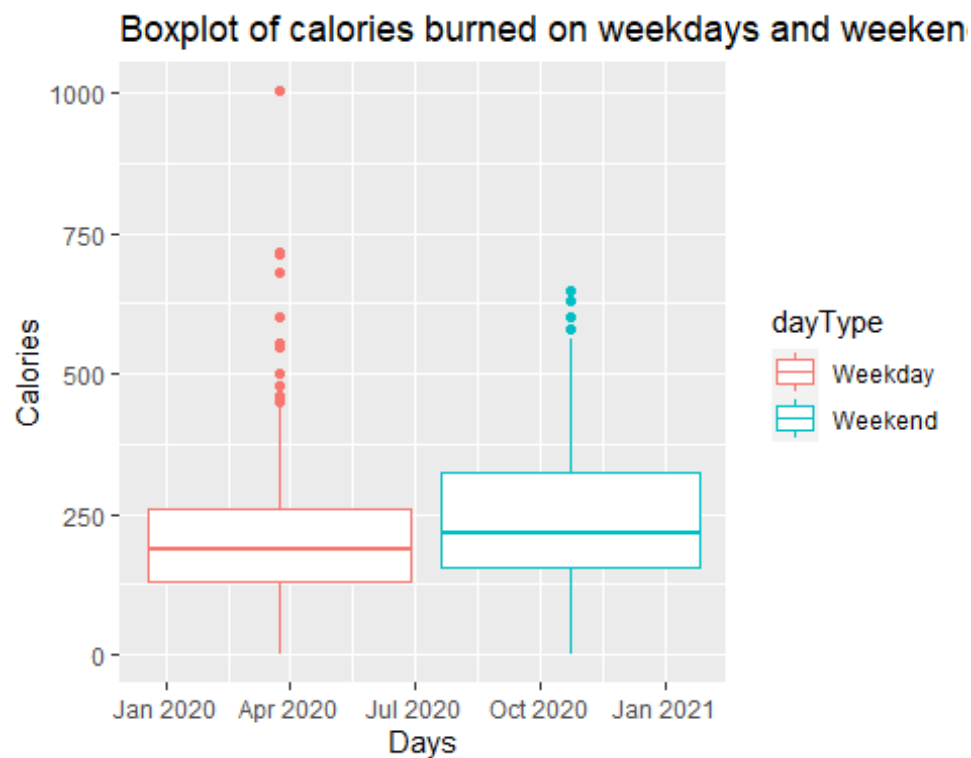
The above scatter plot shows the linear relationship between calories and steps. The more steps one walks, the more calories one will burn. The activity dataset has steps and calories captured every day. So I have calculated steps and calories burned each month. Below two graphs are showing the bar plots of monthly steps and monthly calories.

Barplot of total monthly steps



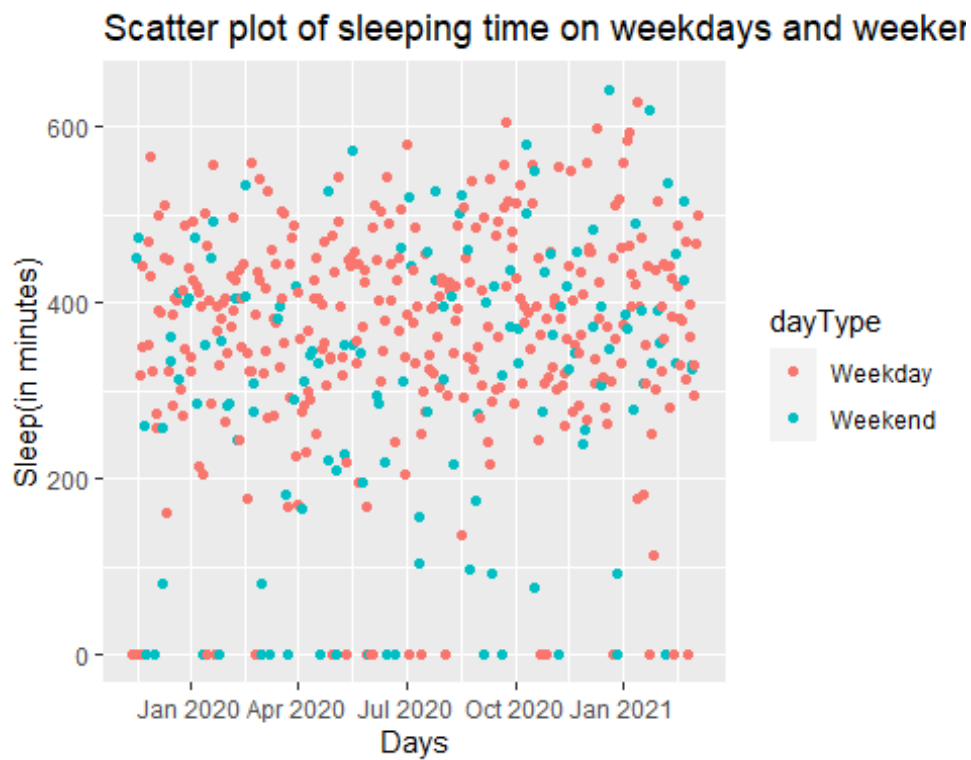
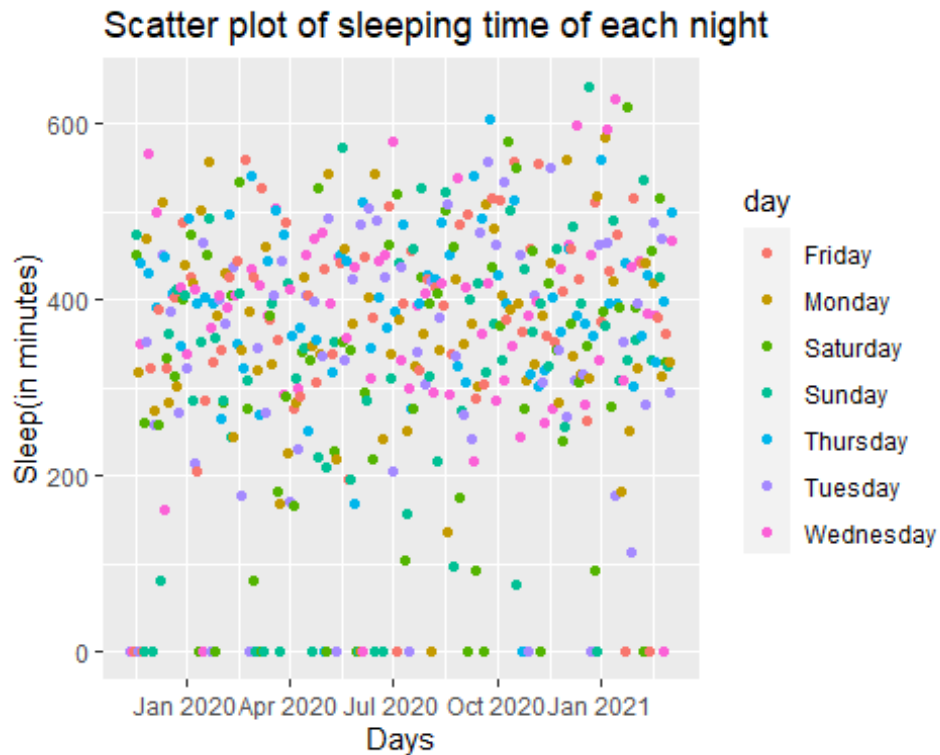
Barplot of total monthly calories





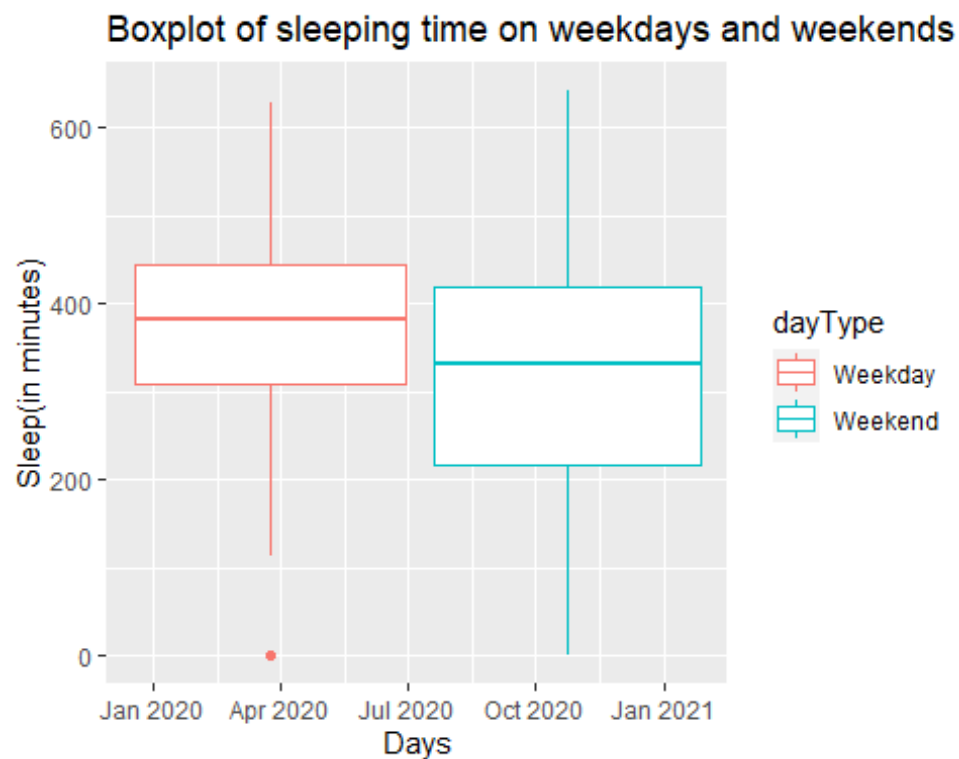
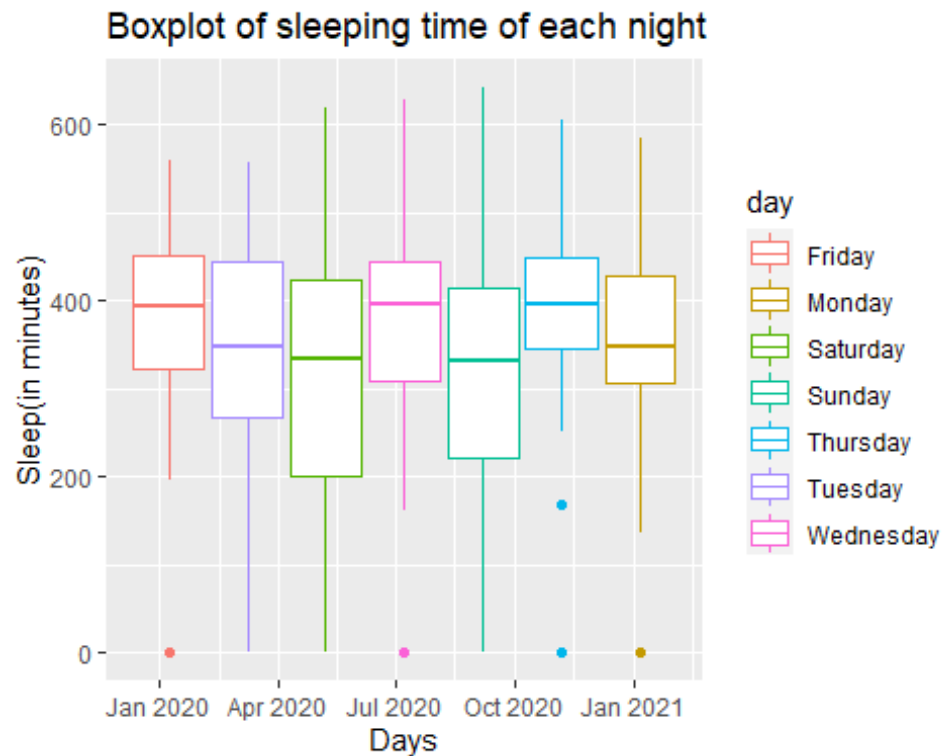
As shown in the above boxplot, calories burned on weekends are more than the calories burned on weekdays. There are outliers on both weekdays and weekends. Even though average calorie burn is less on weekdays, there are some weekdays where steps and calories count is more than usual. The highest point with calories burned more than 1000 kcal is present on weekdays.

Sleep Analysis

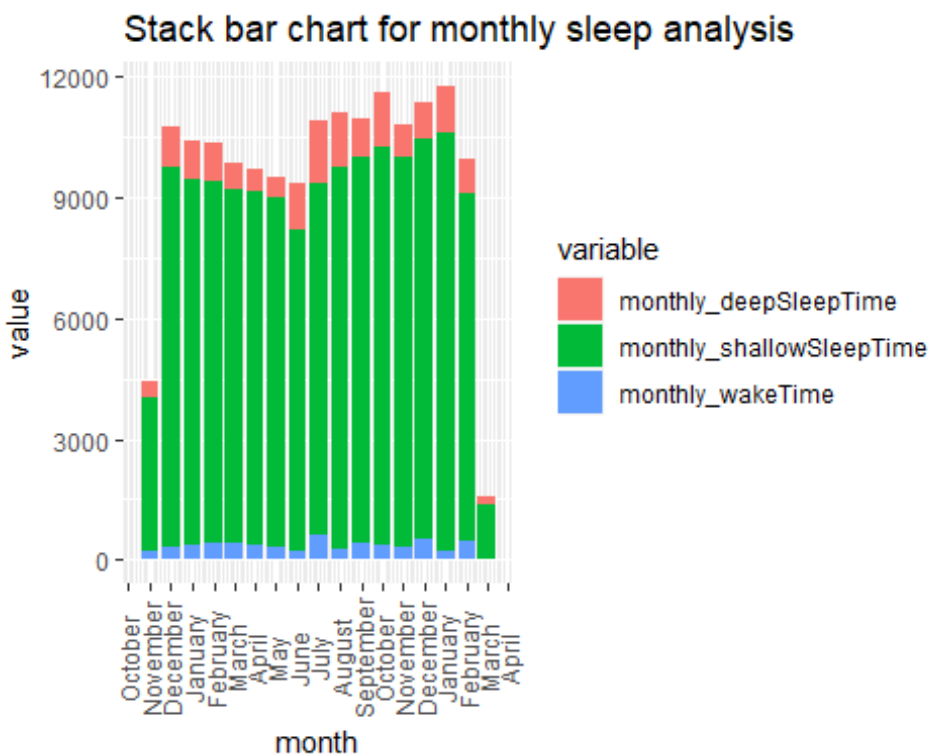
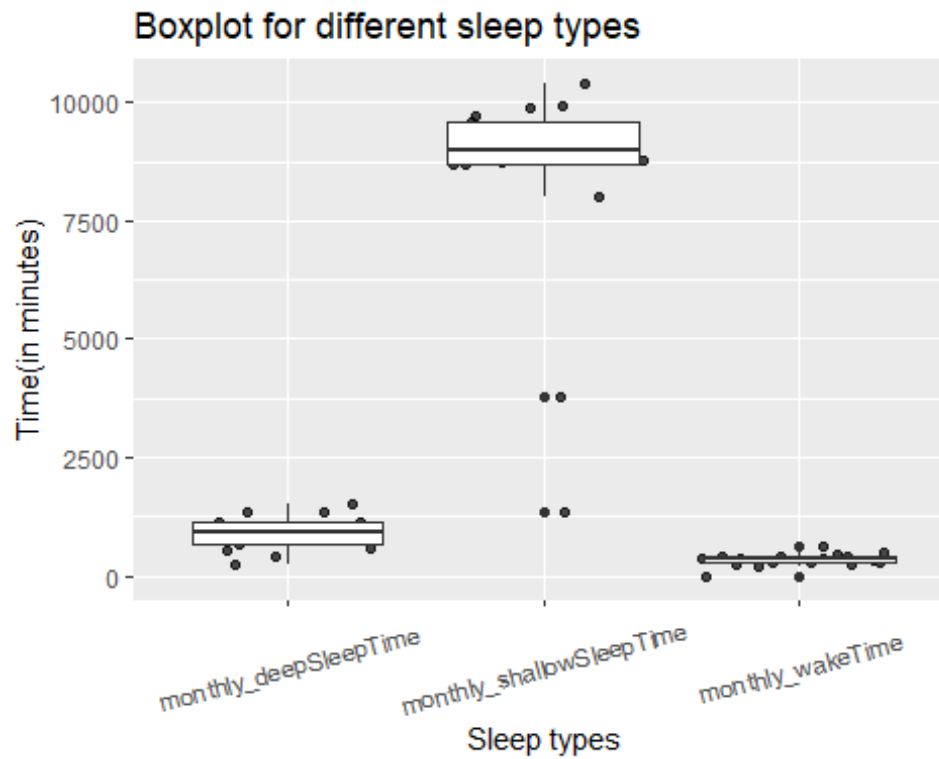


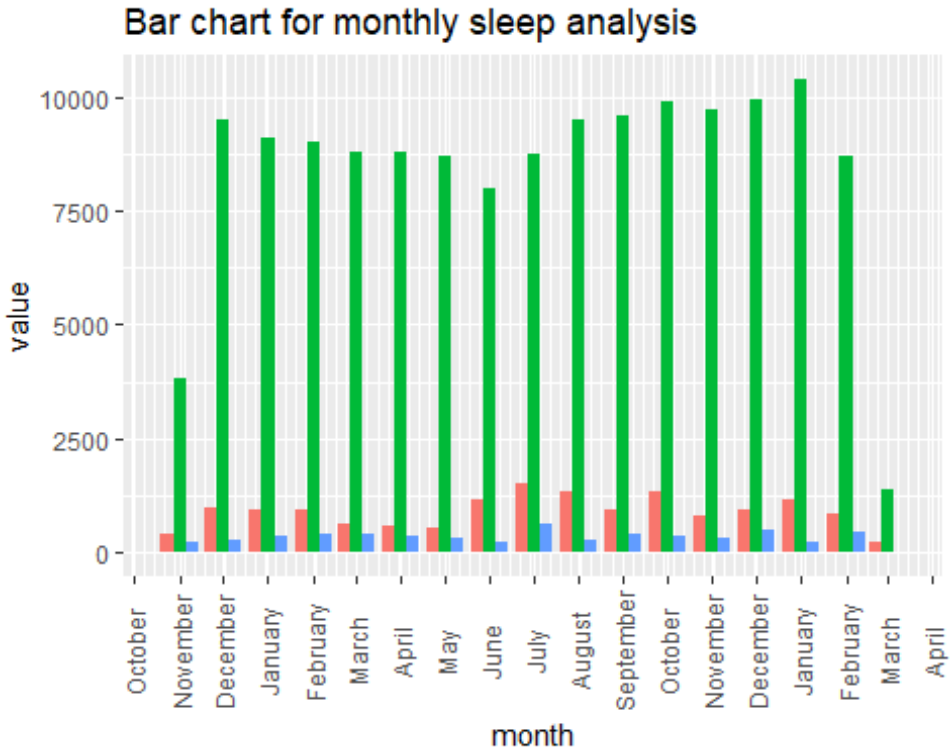
The above scatter plots shows the total sleep of each night. The first plot is divided with 7 colours of a week and the second plot bifurcated with weekdays and weekends. Some points are

at 0 which says 0 minutes of sleep. But it is impossible. Those were the nights I forgot to wear a band. So it couldn't capture the sleep for those nights.



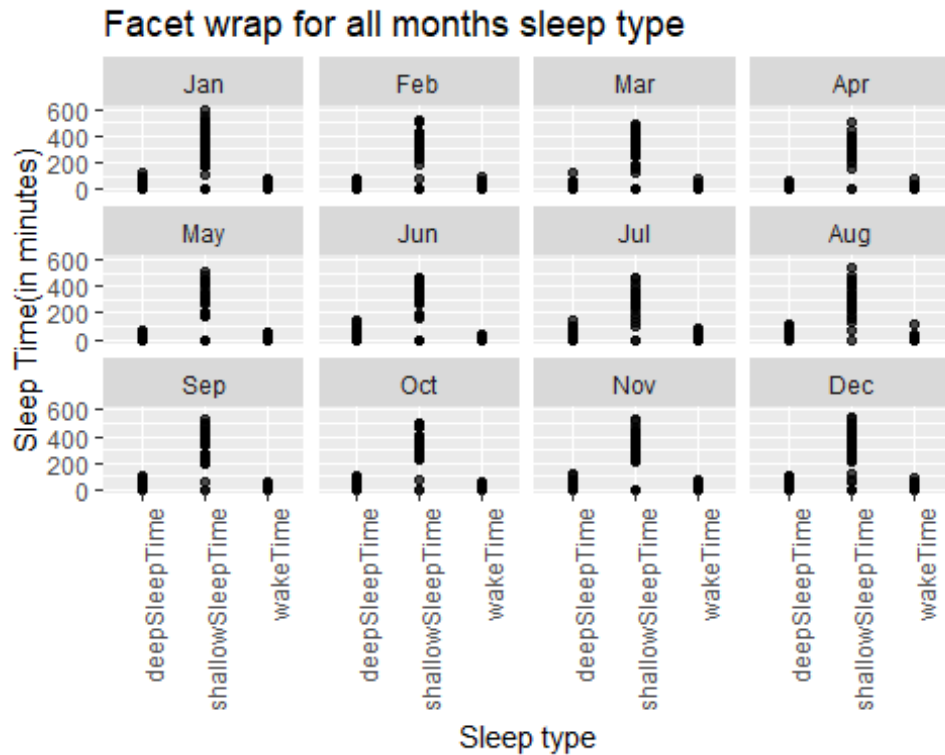
The above boxplots show the total sleep of each night. The first plot is divided with 7 colours of a week and the second plot is bifurcated with weekdays and weekends. The minimum sleep time is on Saturday. The average sleeping time of Wednesday, Thursday, and Friday is at the same level. Sleeping time on Friday is better than any other nights. The second boxplot tells us that the mean sleeping time on weekdays is more than on weekends and weekends has the lowest sleeping time.





[5]

The above stack bar chart shows the total time of sleep time in each month. Each bar consists of 3 different stacks. The lowest stack is wake time, the middle one is shallow sleep time, and the top part is deep sleep time. Maximum shallow sleep time and total sleep time is in January 2021. Sleep time in November 2019 and March 2021 is less because there are very few days in those months.



[6]

Sports Data Analysis

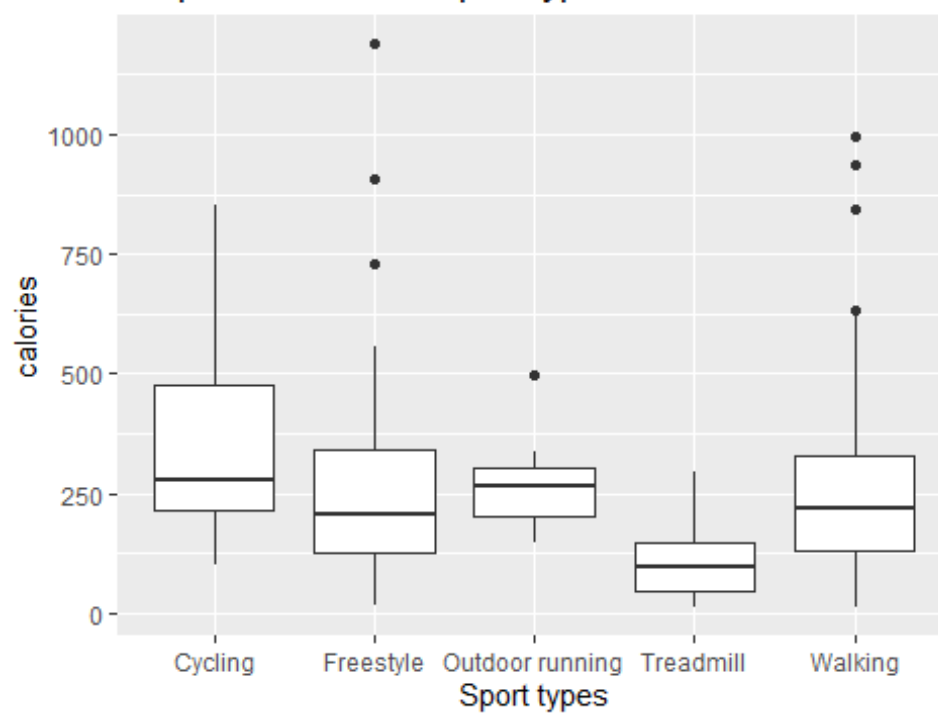
Walking, Outdoor Running, Treadmill, Cycling, and Freestyle are 5 different types of sports activities that can be tracked by MI band 4. And in each activity, calories are calculated. Distance has been recorded wherever needed. The boxplot below shows the calories burned in each sport. The maximum calorie burning sport is cycling. But there are some days in freestyle, which has extreme calorie burn.

Summary of sport data is

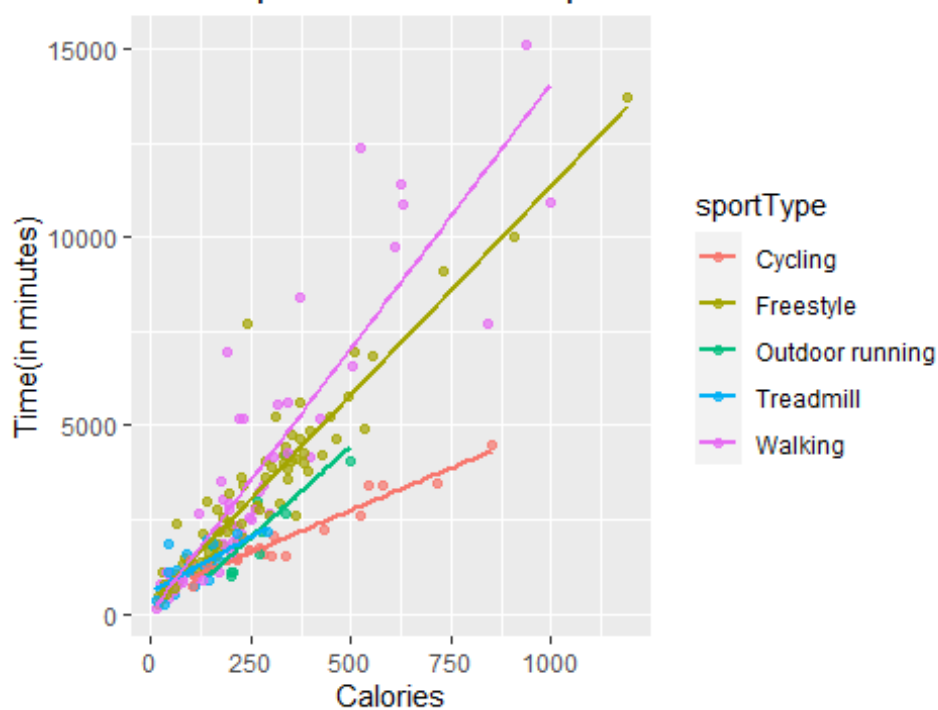
##	type	startTime	sportTime	distance	maxPace
## 1:11	Min.	:1.574e+09	Min. : 135	Min. : 0.000	Min. :0.07618
## 6:51	1st Qu.:	1.584e+09	1st Qu.: 1207	1st Qu.: 0.000	1st Qu.:0.43175
## 8:21	Median	:1.596e+09	Median : 2145	Median : 1.948	Median :0.53417
## 9:19	Mean	:1.596e+09	Mean : 2915	Mean : 2.672	Mean :1.01927
## 16:82	3rd Qu.:	1.607e+09	3rd Qu.: 3868	3rd Qu.: 4.652	3rd Qu.:1.80000
##	Max.	:1.615e+09	Max. :15099	Max. :10.000	Max. :1.80000
##	minPace	avgPace	calories	sportType	
## Min. :	0	Min. :0.0000	Min. : 12.0	Cycling	:19
## 1st Qu.:	0	1st Qu.:0.0000	1st Qu.: 118.2	Freestyle	:82
## Median :	511	Median :0.2941	Median : 206.5	Outdoor running:	11
## Mean :	1802	Mean :0.5981	Mean : 250.5	Treadmill	:21
## 3rd Qu.:	2481	3rd Qu.:0.9263	3rd Qu.: 334.2	Walking	:51
## Max. :	15046	Max. :3.3747	Max. :1190.0		

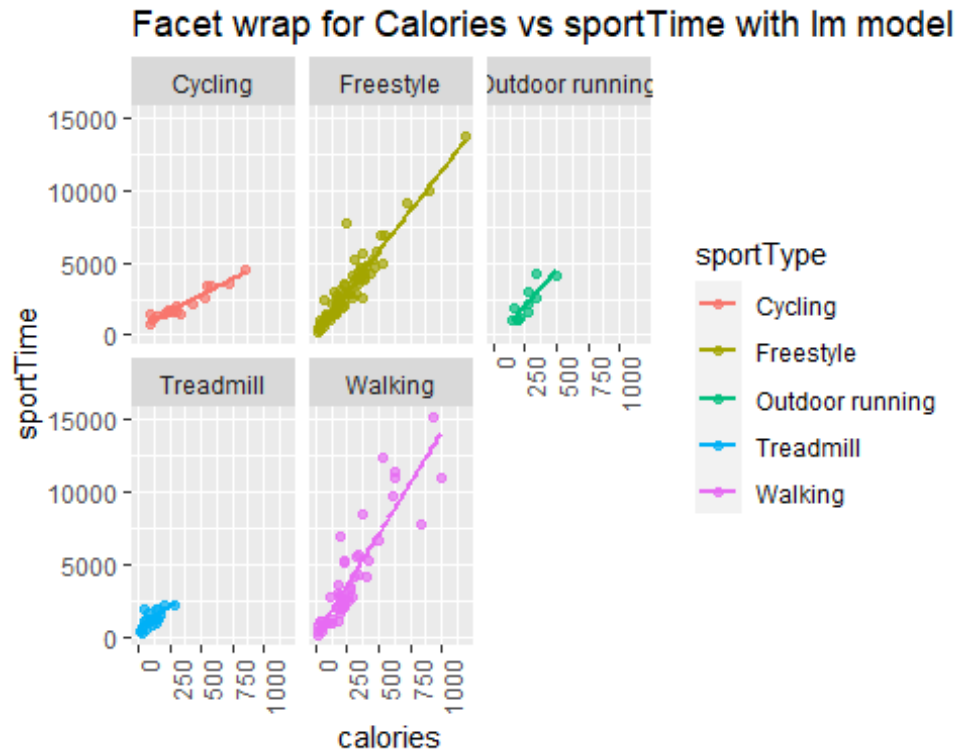
```
##      date
## Min.   :2019-11-14
## 1st Qu.:2020-03-11
## Median :2020-07-28
## Mean   :2020-07-23
## 3rd Qu.:2020-11-30
## Max.   :2021-03-01
```

Boxplot for different sport types



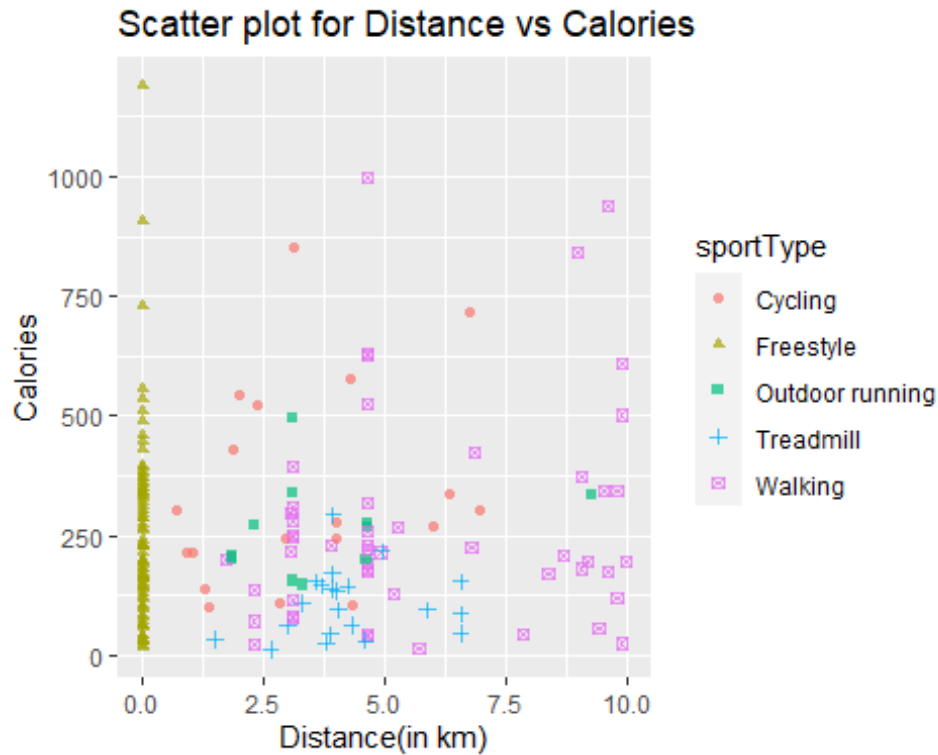
Scatter plot of calories vs sport time





[6]

The more time spent in each sport, the more calories are burned. Therefore, time and calories show a linear relationship. The above facet shows the maximum sessions I have done are walking and freestyle. Treadmill sessions were minimum in number.



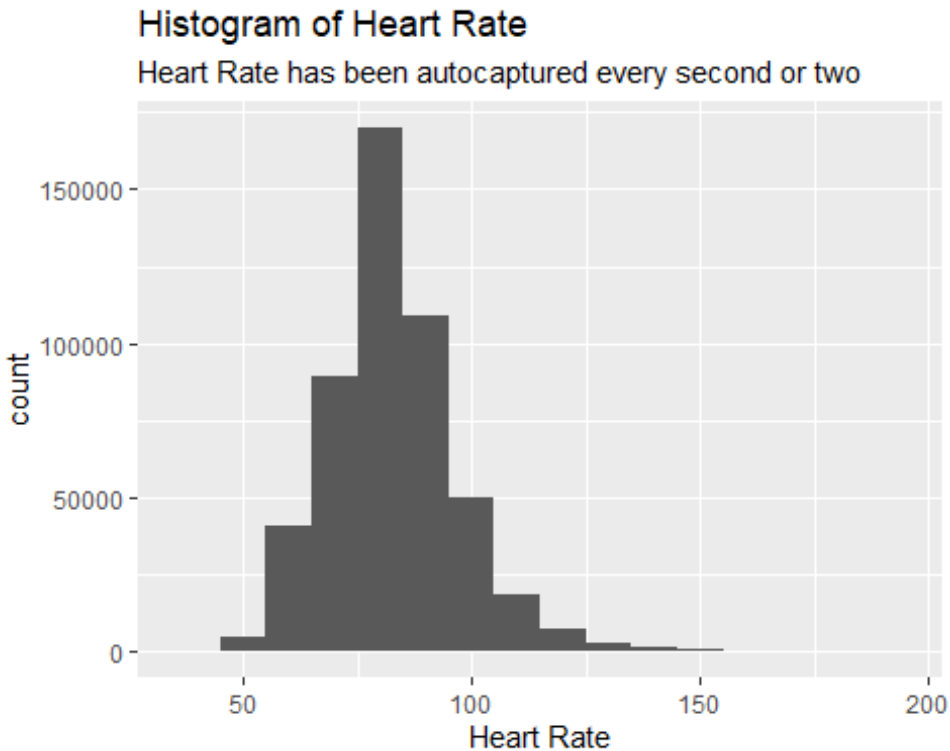
A Scatter plot of distance and calories tell us that distance is recorded in all sports except freestyle.

Heart Rate Analysis

Summary of heart rate data is

##	date	time	heartRate
##	Min. :2019-11-12	Length:493647	Min. : 40.00
##	1st Qu.:2020-03-13	Class :character	1st Qu.: 75.00
##	Median :2020-07-20	Mode :character	Median : 81.00
##	Mean :2020-07-13		Mean : 83.28
##	3rd Qu.:2020-11-11		3rd Qu.: 91.00
##	Max. :2021-03-04		Max. :194.00

A histogram and boxplot for heart rate are given below.



The range of heart rate is 40 to 194 BPM (beats per minute). The average heart rate is 83.28 and the median is 81. There are some points (outliers) that are greater than 91. Those heart rates must be recorded while performing any sports activities. Heart rate between 40 to 50 is very low. It is recorded while sleeping and it is normal.

Limitations

MI band 4 doesn't have in-built GPS but it uses the phone's GPS to record the location. Locations are captured mostly in sports such as walking, running, and cycling. While exporting the data, MI didn't provide the longitude and latitude of those events. Otherwise, it could have been better to display geographical maps.

Conclusion

The analysis of steps, distance, calories, sleep, and heart rate from the data recorded in MI band 4 has been done successfully.

References

- [1] "r - Convert UNIX epoch to Date object," *Stack Overflow*.
<https://stackoverflow.com/questions/13456241/convert-unix-epoch-to-date-object> (accessed Mar. 01, 2021).

- [2] “r - Find the day of a week,” *Stack Overflow*.
<https://stackoverflow.com/questions/9216138/find-the-day-of-a-week> (accessed Mar. 08, 2021).
- [3] “plot - How do I group my date variable into month/year in R?,” *Stack Overflow*.
<https://stackoverflow.com/questions/33221425/how-do-i-group-my-date-variable-into-month-year-in-r> (accessed Mar. 01, 2021).
- [4] Y. Holtz, “Grouped, stacked and percent stacked barplot in ggplot2.” <https://www.r-graph-gallery.com/48-grouped-barplot-with-ggplot2.html> (accessed Mar. 01, 2021).
- [5] “r - Showing data values on stacked bar chart in ggplot2,” *Stack Overflow*.
<https://stackoverflow.com/questions/6644997/showing-data-values-on-stacked-bar-chart-in-ggplot2> (accessed Mar. 01, 2021).
- [6] R. Kabacoff, *Data Visualization with R* .