Vinsent P.

February 7, 2016

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

Assignment 01 Reproducible Research

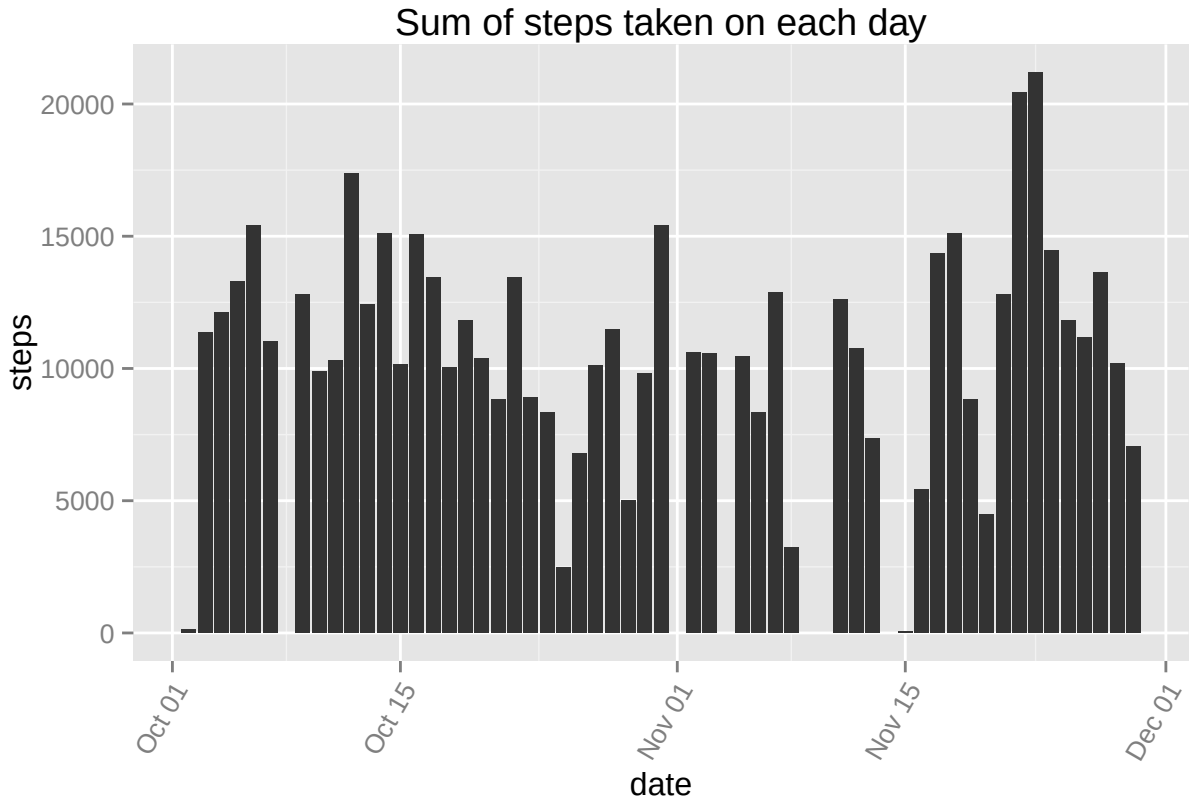
```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.2.2
```

```
## Load the datasets
activity <- read.csv("G:/coursera/DataScience_specialization/RR/repdata_data_activity/activity.csv")
dailySteps <- aggregate(steps ~ date, data=activity, FUN= sum )
dailySteps$date <- as.Date(dailySteps$date)
```

Total Number of Steps Taken Daily

```
ggplot(dailySteps, aes(x=date, y=steps), xlab="Total Number of Steps Taken Daily",main = "Sum of steps
```




```
mean(dailySteps$steps)
```

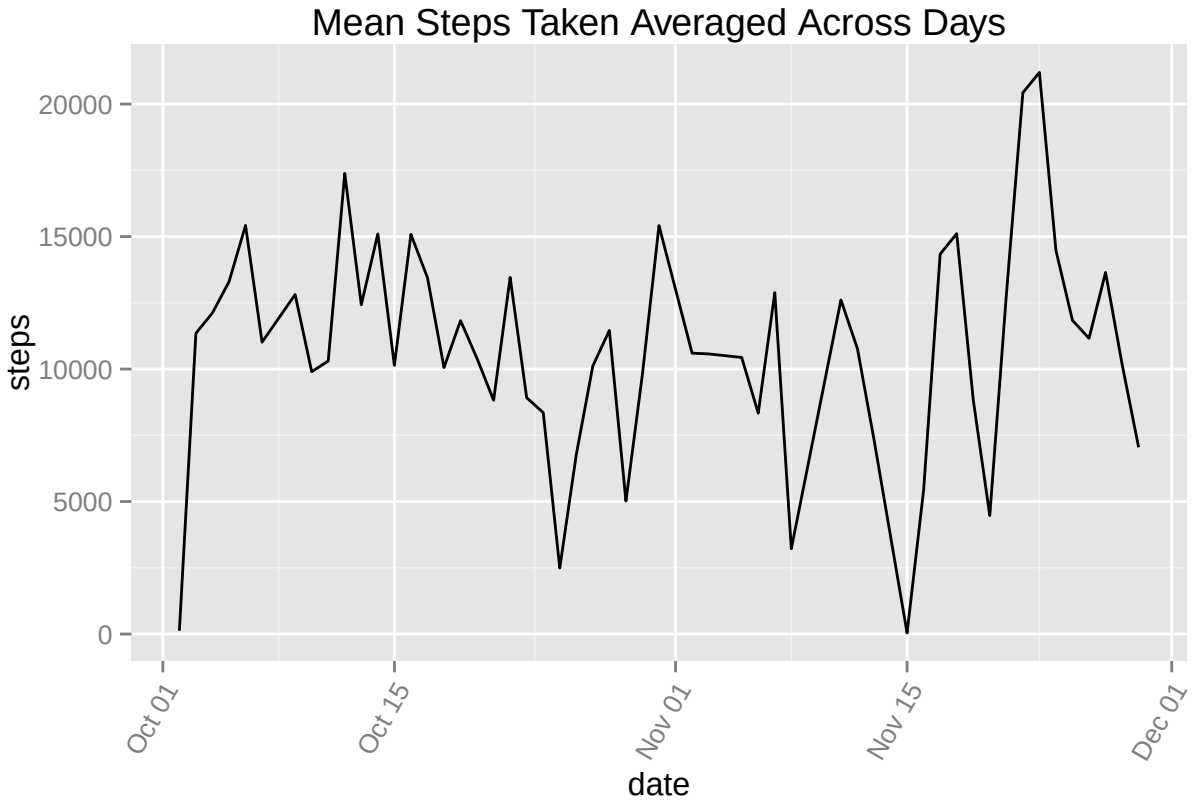
```
## [1] 10766.19
```

```
median(dailySteps$steps)
```

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

Time Series Plot

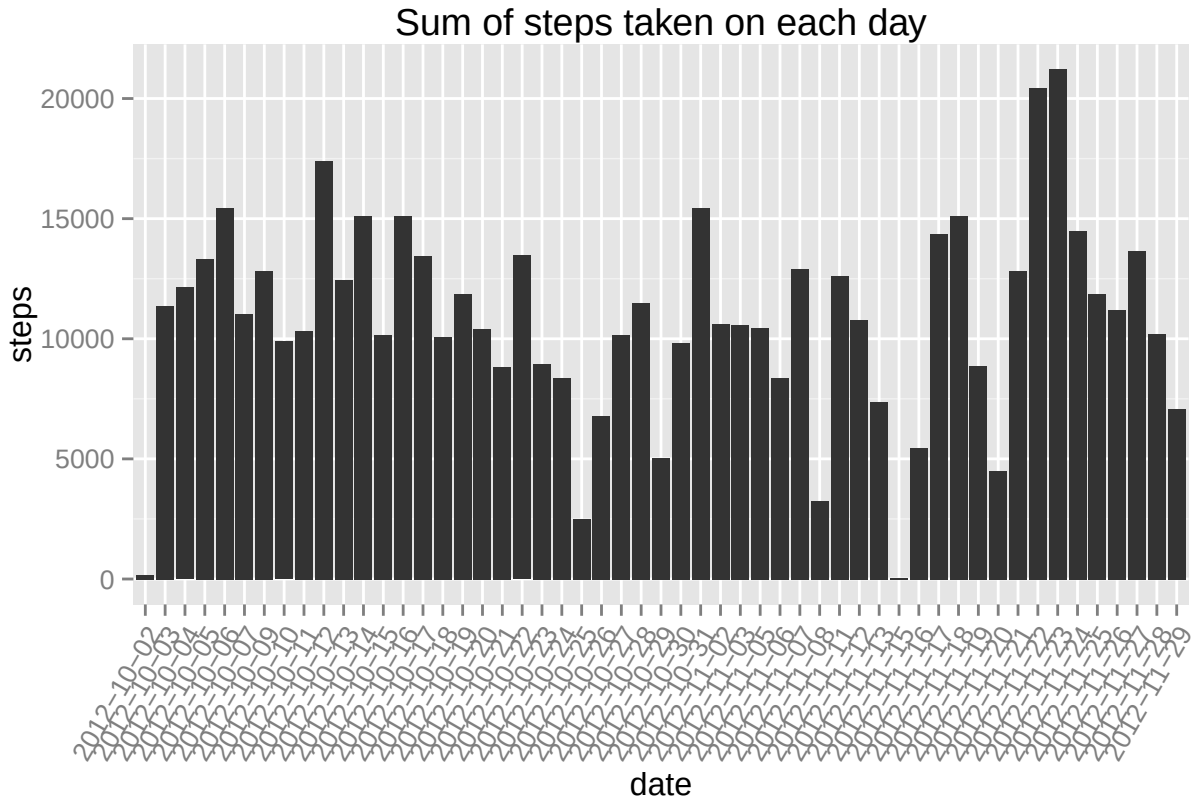
```
dailyStepsMean <- aggregate(steps ~ date, data=activity, FUN= mean);

ggplot(dailySteps, aes(x=date, y=steps), xlab="5-Minute Interval mean steps ") + geom_line(stat="identi
```



dailyStepsMean[which.max(dailyStepsMean$steps),]

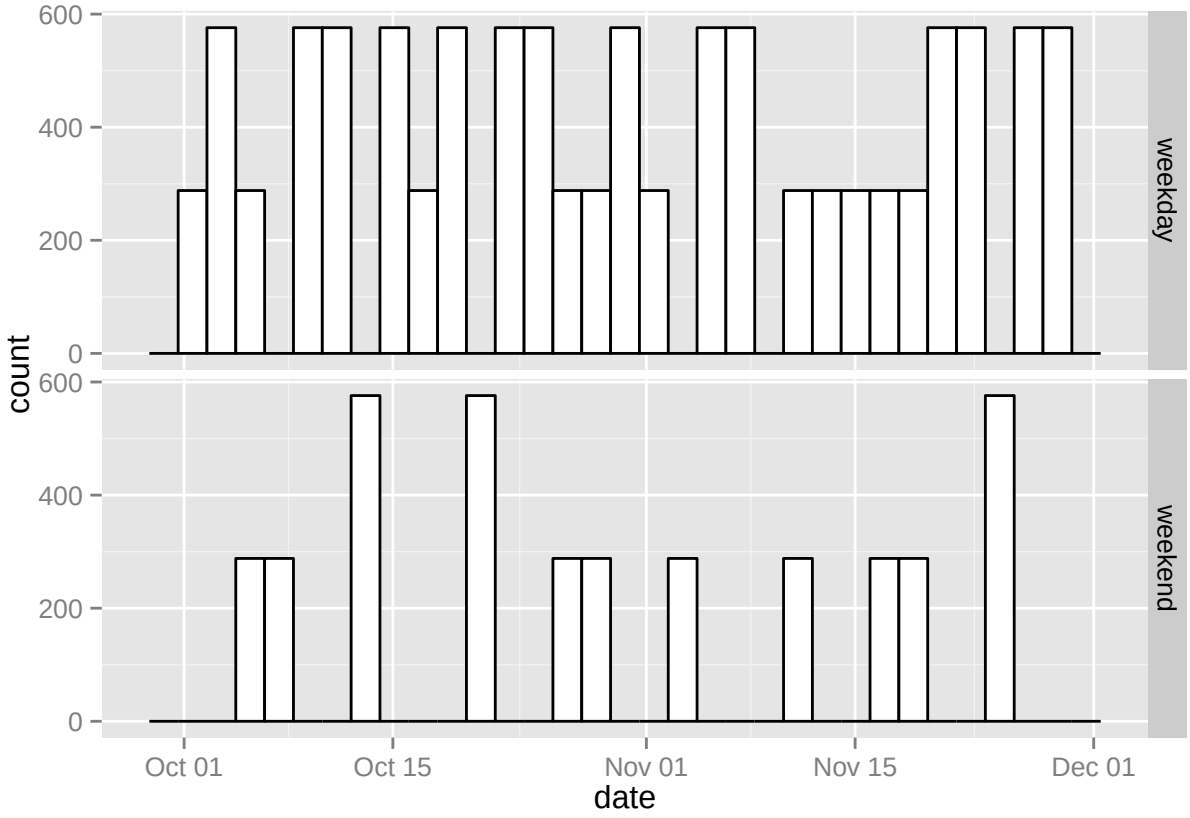
strategy for imputing missing data

```
activity.not.na <- na.omit(activity)

ggplot(activity.not.na, aes(x=date, y=steps), xlab="Total Number of Steps Taken Daily",main = "Sum of s
```



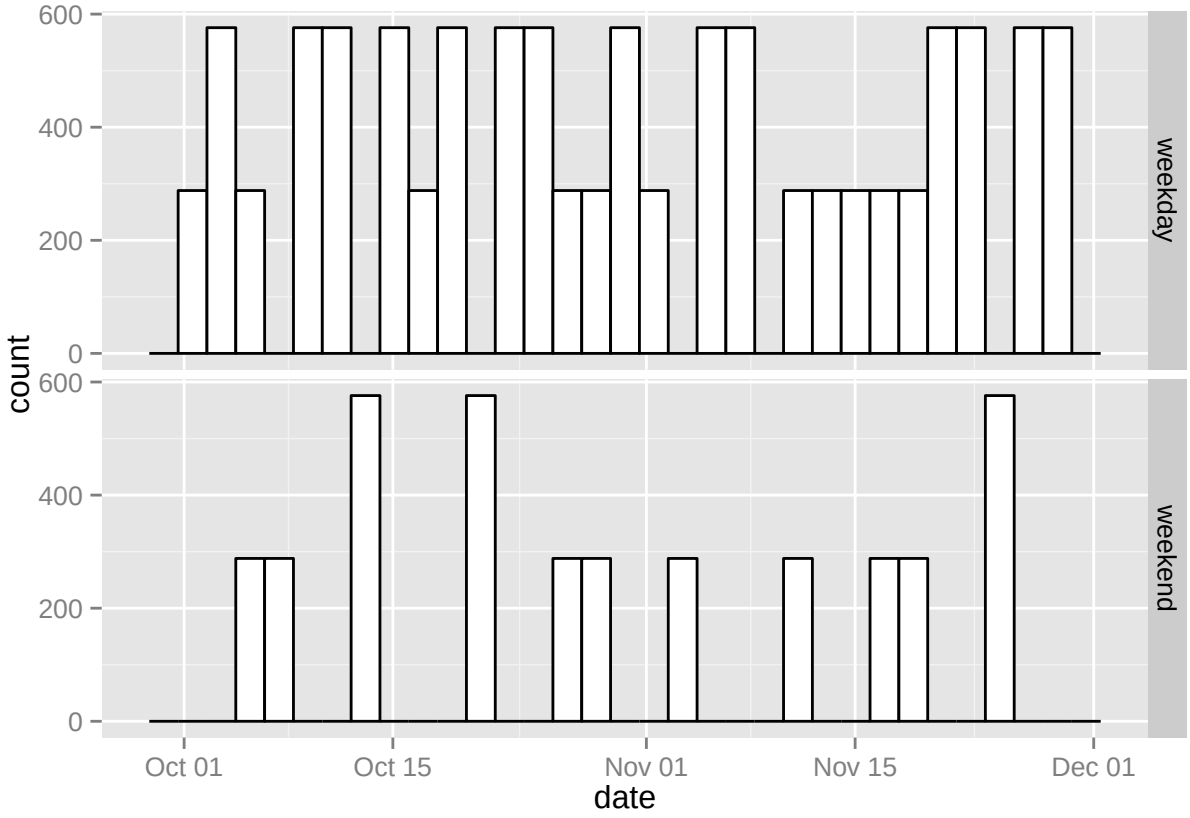
##weekdays and weekends

```
activity.not.na$date <- as.Date(activity.not.na$date)
activity.not.na$weektime <- as.factor(ifelse(weekdays(activity.not.na$date) %in%
                c("Saturday","Sunday"),"weekend", "weekday"))

ggplot(activity.not.na, aes(x=date, y=steps, fill=weektime)) +
geom_bar()
```




```
ggplot(activity.not.na, aes(x=date)) + geom_histogram(fill="white", colour="black") +
facet_grid(weektime ~ .)
```

```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```



Conclusion

Weekend have more consistant activites compared to the weekdays.. Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.