Practical Machine Learning - Week 4 Project

By Jeff Gauzza

Background

Background

Using devices such as Jawbone Up, Nike FuelBand, and Fitbit it is now possible to collect a large amount of data about personal activity relatively inexpensively. These type of devices are part of the quantified self movement – a group of enthusiasts who take measurements about themselves regularly to improve their health, to find patterns in their behavior, or because they are tech geeks. One thing that people regularly do is quantify how much of a particular activity they do, but they rarely quantify how well they do it. In this project, your goal will be to use data from accelerometers on the belt, forearm, arm, and dumbell of 6 participants. They were asked to perform barbell lifts correctly and incorrectly in 5 different ways. More information is available from the website here: http://web.archive.org/web/20161224072740/http:/groupware.les.inf.puc-rio.br/har (see the section on the Weight Lifting Exercise Dataset).

Data

The training data for this project are available here:

https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv

The test data are available here:

https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv

The data for this project come from this source: http://web.archive.org/web/20161224072740/http:/groupware.les.inf.puc-rio.br/har

Load packages

```
library(caret)
library(randomForest)
```

Load the data

First I will set the working directory, download the data and handle any NA or NULL values in the data

```r
#Set working directory
setwd("C:/Users/Jeff/Desktop/Coursera/ML")

#delete old data
rm(list = ls())

#Download files & get date

URLtrain <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv"

trainFile <- "pml-training.csv"

if (!file.exists(trainFile)){download.file(URLtrain, destfile = trainFile, mode='wb')}

URLtest <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv"

testFile <- "pml-testing.csv"

if (!file.exists(testFile)){download.file(URLtest, destfile = testFile, mode='wb')}

dataTrain = read.csv("pml-training.csv", header=TRUE, na.strings=c("NA","#DIV/0!", ""))
dataTest = read.csv("pml-testing.csv", header=TRUE, na.strings=c("NA","#DIV/0!", ""))

dateDownloaded <- date()

head(dataTrain)
```

```
##   X user_name raw_timestamp_part_1 raw_timestamp_part_2   cvtd_timestamp
## 1 1  carlitos           1323084231               788290 05/12/2011 11:23
## 2 2  carlitos           1323084231               808298 05/12/2011 11:23
## 3 3  carlitos           1323084231               820366 05/12/2011 11:23
## 4 4  carlitos           1323084232               120339 05/12/2011 11:23
## 5 5  carlitos           1323084232               196328 05/12/2011 11:23
## 6 6  carlitos           1323084232               304277 05/12/2011 11:23
##   new_window num_window roll_belt pitch_belt yaw_belt total_accel_belt
## 1         no         11      1.41       8.07    -94.4                3
## 2         no         11      1.41       8.07    -94.4                3
## 3         no         11      1.42       8.07    -94.4                3
## 4         no         12      1.48       8.05    -94.4                3
## 5         no         12      1.48       8.07    -94.4                3
## 6         no         12      1.45       8.06    -94.4                3
##   kurtosis_roll_belt kurtosis_picth_belt kurtosis_yaw_belt skewness_roll_belt
## 1                 NA                  NA                NA                 NA
## 2                 NA                  NA                NA                 NA
## 3                 NA                  NA                NA                 NA
## 4                 NA                  NA                NA                 NA
## 5                 NA                  NA                NA                 NA
## 6                 NA                  NA                NA                 NA
##   skewness_roll_belt.1 skewness_yaw_belt max_roll_belt max_picth_belt
## 1                   NA                NA            NA             NA
## 2                   NA                NA            NA             NA
## 3                   NA                NA            NA             NA
## 4                   NA                NA            NA             NA
```

```
## 5                   NA                NA            NA             NA
## 6                   NA                NA            NA             NA
##   max_yaw_belt min_roll_belt min_pitch_belt min_yaw_belt amplitude_roll_belt
## 1           NA            NA             NA           NA                  NA
## 2           NA            NA             NA           NA                  NA
## 3           NA            NA             NA           NA                  NA
## 4           NA            NA             NA           NA                  NA
## 5           NA            NA             NA           NA                  NA
## 6           NA            NA             NA           NA                  NA
##   amplitude_pitch_belt amplitude_yaw_belt var_total_accel_belt avg_roll_belt
## 1                   NA                NA                   NA            NA
## 2                   NA                NA                   NA            NA
## 3                   NA                NA                   NA            NA
## 4                   NA                NA                   NA            NA
## 5                   NA                NA                   NA            NA
## 6                   NA                NA                   NA            NA
##   stddev_roll_belt var_roll_belt avg_pitch_belt stddev_pitch_belt
## 1               NA            NA             NA                NA
## 2               NA            NA             NA                NA
## 3               NA            NA             NA                NA
## 4               NA            NA             NA                NA
## 5               NA            NA             NA                NA
## 6               NA            NA             NA                NA
##   var_pitch_belt avg_yaw_belt stddev_yaw_belt var_yaw_belt gyros_belt_x
## 1             NA           NA              NA           NA         0.00
## 2             NA           NA              NA           NA         0.02
## 3             NA           NA              NA           NA         0.00
## 4             NA           NA              NA           NA         0.02
## 5             NA           NA              NA           NA         0.02
## 6             NA           NA              NA           NA         0.02
##   gyros_belt_y gyros_belt_z accel_belt_x accel_belt_y accel_belt_z
## 1         0.00        -0.02          -21            4           22
## 2         0.00        -0.02          -22            4           22
## 3         0.00        -0.02          -20            5           23
## 4         0.00        -0.03          -22            3           21
## 5         0.02        -0.02          -21            2           24
## 6         0.00        -0.02          -21            4           21
##   magnet_belt_x magnet_belt_y magnet_belt_z roll_arm pitch_arm yaw_arm
## 1            -3           599          -313     -128      22.5    -161
## 2            -7           608          -311     -128      22.5    -161
## 3            -2           600          -305     -128      22.5    -161
## 4            -6           604          -310     -128      22.1    -161
## 5            -6           600          -302     -128      22.1    -161
## 6             0           603          -312     -128      22.0    -161
##   total_accel_arm var_accel_arm avg_roll_arm stddev_roll_arm var_roll_arm
## 1              34            NA           NA              NA           NA
## 2              34            NA           NA              NA           NA
## 3              34            NA           NA              NA           NA
## 4              34            NA           NA              NA           NA
## 5              34            NA           NA              NA           NA
## 6              34            NA           NA              NA           NA
##   avg_pitch_arm stddev_pitch_arm var_pitch_arm avg_yaw_arm stddev_yaw_arm
## 1            NA               NA            NA          NA             NA
## 2            NA               NA            NA          NA             NA
```

```
## 3            NA               NA            NA          NA             NA
## 4            NA               NA            NA          NA             NA
## 5            NA               NA            NA          NA             NA
## 6            NA               NA            NA          NA             NA
##   var_yaw_arm gyros_arm_x gyros_arm_y gyros_arm_z accel_arm_x accel_arm_y
## 1          NA        0.00        0.00       -0.02        -288         109
## 2          NA        0.02       -0.02       -0.02        -290         110
## 3          NA        0.02       -0.02       -0.02        -289         110
## 4          NA        0.02       -0.03        0.02        -289         111
## 5          NA        0.00       -0.03        0.00        -289         111
## 6          NA        0.02       -0.03        0.00        -289         111
##   accel_arm_z magnet_arm_x magnet_arm_y magnet_arm_z kurtosis_roll_arm
## 1        -123         -368          337          516                NA
## 2        -125         -369          337          513                NA
## 3        -126         -368          344          513                NA
## 4        -123         -372          344          512                NA
## 5        -123         -374          337          506                NA
## 6        -122         -369          342          513                NA
##   kurtosis_picth_arm kurtosis_yaw_arm skewness_roll_arm skewness_pitch_arm
## 1                 NA               NA                NA                 NA
## 2                 NA               NA                NA                 NA
## 3                 NA               NA                NA                 NA
## 4                 NA               NA                NA                 NA
## 5                 NA               NA                NA                 NA
## 6                 NA               NA                NA                 NA
##   skewness_yaw_arm max_roll_arm max_picth_arm max_yaw_arm min_roll_arm
## 1               NA           NA            NA          NA           NA
## 2               NA           NA            NA          NA           NA
## 3               NA           NA            NA          NA           NA
## 4               NA           NA            NA          NA           NA
## 5               NA           NA            NA          NA           NA
## 6               NA           NA            NA          NA           NA
##   min_pitch_arm min_yaw_arm amplitude_roll_arm amplitude_pitch_arm
## 1            NA          NA                NA                  NA
## 2            NA          NA                NA                  NA
## 3            NA          NA                NA                  NA
## 4            NA          NA                NA                  NA
## 5            NA          NA                NA                  NA
## 6            NA          NA                NA                  NA
##   amplitude_yaw_arm roll_dumbbell pitch_dumbbell yaw_dumbbell
## 1                NA      13.05217      -70.49400    -84.87394
## 2                NA      13.13074      -70.63751    -84.71065
## 3                NA      12.85075      -70.27812    -85.14078
## 4                NA      13.43120      -70.39379    -84.87363
## 5                NA      13.37872      -70.42856    -84.85306
## 6                NA      13.38246      -70.81759    -84.46500
##   kurtosis_roll_dumbbell kurtosis_picth_dumbbell kurtosis_yaw_dumbbell
## 1                     NA                      NA                    NA
## 2                     NA                      NA                    NA
## 3                     NA                      NA                    NA
## 4                     NA                      NA                    NA
## 5                     NA                      NA                    NA
## 6                     NA                      NA                    NA
##   skewness_roll_dumbbell skewness_pitch_dumbbell skewness_yaw_dumbbell
```

```
## 1                    NA                    NA                    NA
## 2                    NA                    NA                    NA
## 3                    NA                    NA                    NA
## 4                    NA                    NA                    NA
## 5                    NA                    NA                    NA
## 6                    NA                    NA                    NA
##   max_roll_dumbbell max_picth_dumbbell max_yaw_dumbbell min_roll_dumbbell
## 1                NA                 NA               NA                NA
## 2                NA                 NA               NA                NA
## 3                NA                 NA               NA                NA
## 4                NA                 NA               NA                NA
## 5                NA                 NA               NA                NA
## 6                NA                 NA               NA                NA
##   min_pitch_dumbbell min_yaw_dumbbell amplitude_roll_dumbbell
## 1                 NA               NA                      NA
## 2                 NA               NA                      NA
## 3                 NA               NA                      NA
## 4                 NA               NA                      NA
## 5                 NA               NA                      NA
## 6                 NA               NA                      NA
##   amplitude_pitch_dumbbell amplitude_yaw_dumbbell total_accel_dumbbell
## 1                       NA                     NA                   37
## 2                       NA                     NA                   37
## 3                       NA                     NA                   37
## 4                       NA                     NA                   37
## 5                       NA                     NA                   37
## 6                       NA                     NA                   37
##   var_accel_dumbbell avg_roll_dumbbell stddev_roll_dumbbell var_roll_dumbbell
## 1                 NA                NA                   NA                NA
## 2                 NA                NA                   NA                NA
## 3                 NA                NA                   NA                NA
## 4                 NA                NA                   NA                NA
## 5                 NA                NA                   NA                NA
## 6                 NA                NA                   NA                NA
##   avg_pitch_dumbbell stddev_pitch_dumbbell var_pitch_dumbbell avg_yaw_dumbbell
## 1                 NA                    NA                 NA                NA
## 2                 NA                    NA                 NA                NA
## 3                 NA                    NA                 NA                NA
## 4                 NA                    NA                 NA                NA
## 5                 NA                    NA                 NA                NA
## 6                 NA                    NA                 NA                NA
##   stddev_yaw_dumbbell var_yaw_dumbbell gyros_dumbbell_x gyros_dumbbell_y
## 1                  NA               NA                0            -0.02
## 2                  NA               NA                0            -0.02
## 3                  NA               NA                0            -0.02
## 4                  NA               NA                0            -0.02
## 5                  NA               NA                0            -0.02
## 6                  NA               NA                0            -0.02
##   gyros_dumbbell_z accel_dumbbell_x accel_dumbbell_y accel_dumbbell_z
## 1             0.00             -234               47             -271
## 2             0.00             -233               47             -269
## 3             0.00             -232               46             -270
## 4            -0.02             -232               48             -269
## 5             0.00             -233               48             -270
```

```
## 6             0.00             -234               48             -269
##   magnet_dumbbell_x magnet_dumbbell_y magnet_dumbbell_z roll_forearm
## 1              -559               293               -65         28.4
## 2              -555               296               -64         28.3
## 3              -561               298               -63         28.3
## 4              -552               303               -60         28.1
## 5              -554               292               -68         28.0
## 6              -558               294               -66         27.9
##   pitch_forearm yaw_forearm kurtosis_roll_forearm kurtosis_picth_forearm
## 1         -63.9        -153                    NA                     NA
## 2         -63.9        -153                    NA                     NA
## 3         -63.9        -152                    NA                     NA
## 4         -63.9        -152                    NA                     NA
## 5         -63.9        -152                    NA                     NA
## 6         -63.9        -152                    NA                     NA
##   kurtosis_yaw_forearm skewness_roll_forearm skewness_pitch_forearm
## 1                   NA                    NA                     NA
## 2                   NA                    NA                     NA
## 3                   NA                    NA                     NA
## 4                   NA                    NA                     NA
## 5                   NA                    NA                     NA
## 6                   NA                    NA                     NA
##   skewness_yaw_forearm max_roll_forearm max_picth_forearm max_yaw_forearm
## 1                   NA               NA                NA              NA
## 2                   NA               NA                NA              NA
## 3                   NA               NA                NA              NA
## 4                   NA               NA                NA              NA
## 5                   NA               NA                NA              NA
## 6                   NA               NA                NA              NA
##   min_roll_forearm min_pitch_forearm min_yaw_forearm amplitude_roll_forearm
## 1               NA                NA              NA                     NA
## 2               NA                NA              NA                     NA
## 3               NA                NA              NA                     NA
## 4               NA                NA              NA                     NA
## 5               NA                NA              NA                     NA
## 6               NA                NA              NA                     NA
##   amplitude_pitch_forearm amplitude_yaw_forearm total_accel_forearm
## 1                      NA                    NA                 36
## 2                      NA                    NA                 36
## 3                      NA                    NA                 36
## 4                      NA                    NA                 36
## 5                      NA                    NA                 36
## 6                      NA                    NA                 36
##   var_accel_forearm avg_roll_forearm stddev_roll_forearm var_roll_forearm
## 1                NA               NA                  NA               NA
## 2                NA               NA                  NA               NA
## 3                NA               NA                  NA               NA
## 4                NA               NA                  NA               NA
## 5                NA               NA                  NA               NA
## 6                NA               NA                  NA               NA
##   avg_pitch_forearm stddev_pitch_forearm var_pitch_forearm avg_yaw_forearm
## 1                NA                  NA                NA              NA
## 2                NA                  NA                NA              NA
## 3                NA                  NA                NA              NA
```

```
## 4                NA                   NA               NA             NA
## 5                NA                   NA               NA             NA
## 6                NA                   NA               NA             NA
##   stddev_yaw_forearm var_yaw_forearm gyros_forearm_x gyros_forearm_y
## 1                 NA              NA            0.03            0.00
## 2                 NA              NA            0.02            0.00
## 3                 NA              NA            0.03           -0.02
## 4                 NA              NA            0.02           -0.02
## 5                 NA              NA            0.02            0.00
## 6                 NA              NA            0.02           -0.02
##   gyros_forearm_z accel_forearm_x accel_forearm_y accel_forearm_z
## 1           -0.02             192             203            -215
## 2           -0.02             192             203            -216
## 3            0.00             196             204            -213
## 4            0.00             189             206            -214
## 5           -0.02             189             206            -214
## 6           -0.03             193             203            -215
##   magnet_forearm_x magnet_forearm_y magnet_forearm_z classe
## 1              -17              654              476      A
## 2              -18              661              473      A
## 3              -18              658              469      A
## 4              -16              658              469      A
## 5              -17              655              473      A
## 6               -9              660              478      A
```

```
head(dataTest)
```

```
##   X user_name raw_timestamp_part_1 raw_timestamp_part_2   cvtd_timestamp
## 1 1     pedro           1323095002               868349 05/12/2011 14:23
## 2 2    jeremy           1322673067               778725 30/11/2011 17:11
## 3 3    jeremy           1322673075               342967 30/11/2011 17:11
## 4 4    adelmo           1322832789               560311 02/12/2011 13:33
## 5 5    eurico           1322489635               814776 28/11/2011 14:13
## 6 6    jeremy           1322673149               510661 30/11/2011 17:12
##   new_window num_window roll_belt pitch_belt yaw_belt total_accel_belt
## 1         no         74    123.00      27.00    -4.75               20
## 2         no        431      1.02       4.87   -88.90                4
## 3         no        439      0.87       1.82   -88.50                5
## 4         no        194    125.00     -41.60   162.00               17
## 5         no        235      1.35       3.33   -88.60                3
## 6         no        504     -5.92       1.59   -87.70                4
##   kurtosis_roll_belt kurtosis_picth_belt kurtosis_yaw_belt skewness_roll_belt
## 1                 NA                  NA                NA                 NA
## 2                 NA                  NA                NA                 NA
## 3                 NA                  NA                NA                 NA
## 4                 NA                  NA                NA                 NA
## 5                 NA                  NA                NA                 NA
## 6                 NA                  NA                NA                 NA
##   skewness_roll_belt.1 skewness_yaw_belt max_roll_belt max_picth_belt
## 1                   NA                NA            NA             NA
## 2                   NA                NA            NA             NA
## 3                   NA                NA            NA             NA
## 4                   NA                NA            NA             NA
## 5                   NA                NA            NA             NA
```

```
## 6                   NA                NA            NA             NA
##   max_yaw_belt min_roll_belt min_pitch_belt min_yaw_belt amplitude_roll_belt
## 1           NA            NA             NA           NA                  NA
## 2           NA            NA             NA           NA                  NA
## 3           NA            NA             NA           NA                  NA
## 4           NA            NA             NA           NA                  NA
## 5           NA            NA             NA           NA                  NA
## 6           NA            NA             NA           NA                  NA
##   amplitude_pitch_belt amplitude_yaw_belt var_total_accel_belt avg_roll_belt
## 1                   NA                 NA                   NA            NA
## 2                   NA                 NA                   NA            NA
## 3                   NA                 NA                   NA            NA
## 4                   NA                 NA                   NA            NA
## 5                   NA                 NA                   NA            NA
## 6                   NA                 NA                   NA            NA
##   stddev_roll_belt var_roll_belt avg_pitch_belt stddev_pitch_belt
## 1               NA            NA             NA                NA
## 2               NA            NA             NA                NA
## 3               NA            NA             NA                NA
## 4               NA            NA             NA                NA
## 5               NA            NA             NA                NA
## 6               NA            NA             NA                NA
##   var_pitch_belt avg_yaw_belt stddev_yaw_belt var_yaw_belt gyros_belt_x
## 1             NA           NA              NA           NA        -0.50
## 2             NA           NA              NA           NA        -0.06
## 3             NA           NA              NA           NA         0.05
## 4             NA           NA              NA           NA         0.11
## 5             NA           NA              NA           NA         0.03
## 6             NA           NA              NA           NA         0.10
##   gyros_belt_y gyros_belt_z accel_belt_x accel_belt_y accel_belt_z
## 1        -0.02        -0.46          -38           69         -179
## 2        -0.02        -0.07          -13           11           39
## 3         0.02         0.03            1           -1           49
## 4         0.11        -0.16           46           45         -156
## 5         0.02         0.00           -8            4           27
## 6         0.05        -0.13          -11          -16           38
##   magnet_belt_x magnet_belt_y magnet_belt_z roll_arm pitch_arm yaw_arm
## 1           -13           581          -382     40.7    -27.80     178
## 2            43           636          -309      0.0      0.00       0
## 3            29           631          -312      0.0      0.00       0
## 4           169           608          -304   -109.0     55.00    -142
## 5            33           566          -418     76.1      2.76     102
## 6            31           638          -291      0.0      0.00       0
##   total_accel_arm var_accel_arm avg_roll_arm stddev_roll_arm var_roll_arm
## 1              10            NA           NA              NA           NA
## 2              38            NA           NA              NA           NA
## 3              44            NA           NA              NA           NA
## 4              25            NA           NA              NA           NA
## 5              29            NA           NA              NA           NA
## 6              14            NA           NA              NA           NA
##   avg_pitch_arm stddev_pitch_arm var_pitch_arm avg_yaw_arm stddev_yaw_arm
## 1            NA               NA            NA          NA             NA
## 2            NA               NA            NA          NA             NA
## 3            NA               NA            NA          NA             NA
```

```
## 4            NA               NA            NA          NA             NA
## 5            NA               NA            NA          NA             NA
## 6            NA               NA            NA          NA             NA
##   var_yaw_arm gyros_arm_x gyros_arm_y gyros_arm_z accel_arm_x accel_arm_y
## 1          NA       -1.65        0.48       -0.18          16          38
## 2          NA       -1.17        0.85       -0.43        -290         215
## 3          NA        2.10       -1.36        1.13        -341         245
## 4          NA        0.22       -0.51        0.92        -238         -57
## 5          NA       -1.96        0.79       -0.54        -197         200
## 6          NA        0.02        0.05       -0.07         -26         130
##   accel_arm_z magnet_arm_x magnet_arm_y magnet_arm_z kurtosis_roll_arm
## 1          93         -326          385          481                NA
## 2         -90         -325          447          434                NA
## 3         -87         -264          474          413                NA
## 4           6         -173          257          633                NA
## 5         -30         -170          275          617                NA
## 6         -19          396          176          516                NA
##   kurtosis_picth_arm kurtosis_yaw_arm skewness_roll_arm skewness_pitch_arm
## 1                 NA               NA                NA                 NA
## 2                 NA               NA                NA                 NA
## 3                 NA               NA                NA                 NA
## 4                 NA               NA                NA                 NA
## 5                 NA               NA                NA                 NA
## 6                 NA               NA                NA                 NA
##   skewness_yaw_arm max_roll_arm max_picth_arm max_yaw_arm min_roll_arm
## 1               NA           NA            NA          NA           NA
## 2               NA           NA            NA          NA           NA
## 3               NA           NA            NA          NA           NA
## 4               NA           NA            NA          NA           NA
## 5               NA           NA            NA          NA           NA
## 6               NA           NA            NA          NA           NA
##   min_pitch_arm min_yaw_arm amplitude_roll_arm amplitude_pitch_arm
## 1            NA          NA                NA                 NA
## 2            NA          NA                NA                 NA
## 3            NA          NA                NA                 NA
## 4            NA          NA                NA                 NA
## 5            NA          NA                NA                 NA
## 6            NA          NA                NA                 NA
##   amplitude_yaw_arm roll_dumbbell pitch_dumbbell yaw_dumbbell
## 1                NA     -17.73748       24.96085    126.23596
## 2                NA      54.47761      -53.69758    -75.51480
## 3                NA      57.07031      -51.37303    -75.20287
## 4                NA      43.10927      -30.04885   -103.32003
## 5                NA    -101.38396      -53.43952    -14.19542
## 6                NA      62.18750      -50.55595    -71.12063
##   kurtosis_roll_dumbbell kurtosis_picth_dumbbell kurtosis_yaw_dumbbell
## 1                     NA                      NA                    NA
## 2                     NA                      NA                    NA
## 3                     NA                      NA                    NA
## 4                     NA                      NA                    NA
## 5                     NA                      NA                    NA
## 6                     NA                      NA                    NA
##   skewness_roll_dumbbell skewness_pitch_dumbbell skewness_yaw_dumbbell
## 1                     NA                      NA                    NA
```

```
## 2                    NA                    NA                    NA
## 3                    NA                    NA                    NA
## 4                    NA                    NA                    NA
## 5                    NA                    NA                    NA
## 6                    NA                    NA                    NA
##   max_roll_dumbbell max_picth_dumbbell max_yaw_dumbbell min_roll_dumbbell
## 1                NA                 NA               NA                NA
## 2                NA                 NA               NA                NA
## 3                NA                 NA               NA                NA
## 4                NA                 NA               NA                NA
## 5                NA                 NA               NA                NA
## 6                NA                 NA               NA                NA
##   min_pitch_dumbbell min_yaw_dumbbell amplitude_roll_dumbbell
## 1                 NA               NA                      NA
## 2                 NA               NA                      NA
## 3                 NA               NA                      NA
## 4                 NA               NA                      NA
## 5                 NA               NA                      NA
## 6                 NA               NA                      NA
##   amplitude_pitch_dumbbell amplitude_yaw_dumbbell total_accel_dumbbell
## 1                       NA                     NA                    9
## 2                       NA                     NA                   31
## 3                       NA                     NA                   29
## 4                       NA                     NA                   18
## 5                       NA                     NA                    4
## 6                       NA                     NA                   29
##   var_accel_dumbbell avg_roll_dumbbell stddev_roll_dumbbell var_roll_dumbbell
## 1                 NA                NA                   NA                NA
## 2                 NA                NA                   NA                NA
## 3                 NA                NA                   NA                NA
## 4                 NA                NA                   NA                NA
## 5                 NA                NA                   NA                NA
## 6                 NA                NA                   NA                NA
##   avg_pitch_dumbbell stddev_pitch_dumbbell var_pitch_dumbbell avg_yaw_dumbbell
## 1                 NA                    NA                 NA                NA
## 2                 NA                    NA                 NA                NA
## 3                 NA                    NA                 NA                NA
## 4                 NA                    NA                 NA                NA
## 5                 NA                    NA                 NA                NA
## 6                 NA                    NA                 NA                NA
##   stddev_yaw_dumbbell var_yaw_dumbbell gyros_dumbbell_x gyros_dumbbell_y
## 1                  NA               NA             0.64             0.06
## 2                  NA               NA             0.34             0.05
## 3                  NA               NA             0.39             0.14
## 4                  NA               NA             0.10            -0.02
## 5                  NA               NA             0.29            -0.47
## 6                  NA               NA            -0.59             0.80
##   gyros_dumbbell_z accel_dumbbell_x accel_dumbbell_y accel_dumbbell_z
## 1            -0.61              21              -15               81
## 2            -0.71            -153              155             -205
## 3            -0.34            -141              155             -196
## 4             0.05             -51               72             -148
## 5            -0.46             -18              -30               -5
## 6             1.10            -138              166             -186
```

```
##   magnet_dumbbell_x magnet_dumbbell_y magnet_dumbbell_z roll_forearm
## 1               523              -528               -56          141
## 2              -502               388               -36          109
## 3              -506               349                41          131
## 4              -576               238                53            0
## 5              -424               252               312         -176
## 6              -543               262                96          150
##   pitch_forearm yaw_forearm kurtosis_roll_forearm kurtosis_picth_forearm
## 1         49.30       156.0                    NA                     NA
## 2        -17.60       106.0                    NA                     NA
## 3        -32.60        93.0                    NA                     NA
## 4          0.00         0.0                    NA                     NA
## 5         -2.16       -47.9                    NA                     NA
## 6          1.46        89.7                    NA                     NA
##   kurtosis_yaw_forearm skewness_roll_forearm skewness_pitch_forearm
## 1                   NA                    NA                     NA
## 2                   NA                    NA                     NA
## 3                   NA                    NA                     NA
## 4                   NA                    NA                     NA
## 5                   NA                    NA                     NA
## 6                   NA                    NA                     NA
##   skewness_yaw_forearm max_roll_forearm max_picth_forearm max_yaw_forearm
## 1                   NA               NA                NA              NA
## 2                   NA               NA                NA              NA
## 3                   NA               NA                NA              NA
## 4                   NA               NA                NA              NA
## 5                   NA               NA                NA              NA
## 6                   NA               NA                NA              NA
##   min_roll_forearm min_pitch_forearm min_yaw_forearm amplitude_roll_forearm
## 1               NA                NA              NA                     NA
## 2               NA                NA              NA                     NA
## 3               NA                NA              NA                     NA
## 4               NA                NA              NA                     NA
## 5               NA                NA              NA                     NA
## 6               NA                NA              NA                     NA
##   amplitude_pitch_forearm amplitude_yaw_forearm total_accel_forearm
## 1                      NA                   NA                 33
## 2                      NA                   NA                 39
## 3                      NA                   NA                 34
## 4                      NA                   NA                 43
## 5                      NA                   NA                 24
## 6                      NA                   NA                 43
##   var_accel_forearm avg_roll_forearm stddev_roll_forearm var_roll_forearm
## 1                NA               NA                  NA              NA
## 2                NA               NA                  NA              NA
## 3                NA               NA                  NA              NA
## 4                NA               NA                  NA              NA
## 5                NA               NA                  NA              NA
## 6                NA               NA                  NA              NA
##   avg_pitch_forearm stddev_pitch_forearm var_pitch_forearm avg_yaw_forearm
## 1                NA                   NA                NA              NA
## 2                NA                   NA                NA              NA
## 3                NA                   NA                NA              NA
## 4                NA                   NA                NA              NA
```

```
## 5                NA                   NA               NA             NA
## 6                NA                   NA               NA             NA
##   stddev_yaw_forearm var_yaw_forearm gyros_forearm_x gyros_forearm_y
## 1                 NA              NA            0.74           -3.34
## 2                 NA              NA            1.12           -2.78
## 3                 NA              NA            0.18           -0.79
## 4                 NA              NA            1.38            0.69
## 5                 NA              NA           -0.75            3.10
## 6                 NA              NA           -0.88            4.26
##   gyros_forearm_z accel_forearm_x accel_forearm_y accel_forearm_z
## 1           -0.59            -110             267            -149
## 2           -0.18             212             297            -118
## 3            0.28             154             271            -129
## 4            1.80             -92             406             -39
## 5            0.80             131             -93             172
## 6            1.35             230             322            -144
##   magnet_forearm_x magnet_forearm_y magnet_forearm_z problem_id
## 1             -714              419              617          1
## 2             -237              791              873          2
## 3              -51              698              783          3
## 4             -233              783              521          4
## 5              375             -787               91          5
## 6             -300              800              884          6
```

Clean data

As you can see from display of the first few rows above, the first bunch of columns (user_name, raw_timestamp_part_1, raw_timestamp_part_2, cvtd_timestamp, new_window, and num_window) are not needed and will be removed below from both the training and testing data. I will also remove variables which have a lot of NA values. I know I am going to do a RandomForest for my model and it does not do well with variables with many NA values.

```
dataTrain <- dataTrain[,-c(1:7)]
dataTest <- dataTest[,-c(1:7)]

naIndex <- colSums(is.na(dataTrain))/nrow(dataTrain) < 0.95
dataTrain <- dataTrain[,naIndex]
```

Cross validation

I further split the training data for cross validation purposes using a 70/30 split.

```
subIndex <- createDataPartition(y=dataTrain$classe, p=0.70, list=FALSE)
trainSubTrain <- dataTrain[subIndex, ]
trainSubTest<- dataTrain[-subIndex, ]
```

How I built the model

For the model I am going to use a RandomForest. As I learned in this class, Random Forests often produce good predictive models, even if the choices they make can be hard to logically determine the reasoning behind them. They can also take a long time to run. I will keep the number of trees low (10 in this case) to keep run times down.

```
modelRanFor <- train(classe ~., method='rf', data=trainSubTrain, ntree=10)
predRanFor <- predict(modelRanFor, trainSubTest)

RanForConMatrix <- confusionMatrix(predRanFor,trainSubTest$classe)
RanForConMatrix
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1670   11    0    0    4
##          B    3 1123    6    2    4
##          C    0    4 1010   14    1
##          D    1    0   10  947    6
##          E    0    1    0    1 1067
##
## Overall Statistics
##
##                Accuracy : 0.9884
##                  95% CI : (0.9854, 0.991)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9854
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            0.9976   0.9860   0.9844   0.9824   0.9861
## Specificity            0.9964   0.9968   0.9961   0.9965   0.9996
## Pos Pred Value         0.9911   0.9868   0.9815   0.9824   0.9981
## Neg Pred Value         0.9990   0.9966   0.9967   0.9965   0.9969
## Prevalence             0.2845   0.1935   0.1743   0.1638   0.1839
## Detection Rate         0.2838   0.1908   0.1716   0.1609   0.1813
## Detection Prevalence   0.2863   0.1934   0.1749   0.1638   0.1816
## Balanced Accuracy      0.9970   0.9914   0.9902   0.9895   0.9929
```

Expected sample error

The expected sample error is 1-the accuracy. Even with running a Random Forest with only 10 trees, we got an accuracy of 98.7%. The expected error is then (1-.987)*100 = 1.3%

Results

Overall 98.7% accuracy is very good and would only be improved with a larger number included in the random forest. For explanative purposes, it was easier to show how accurate the method could be with a lower number of trees and limited processing power and understanding the potential for greater accuracy with the change of one variable. ***

Predicting test cases

Below is the predictions for the test set load at the beginning of the project.

```
finalPredictions <- predict(modelRanFor, dataTest)
finalPredictions
```

```
##  [1] B A B A A E D B A A B C B A E E A B B B
## Levels: A B C D E
```