project report

First download the files and load relevant library

```
library(caret)
```

```
## Loading required package: lattice
## Loading required package: ggplot2
```

```
library(rpart)
library(randomForest)
```

```
## randomForest 4.6-12
## Type rfNews() to see new features/changes/bug fixes.
##
## Attaching package: 'randomForest'
##
## The following object is masked from 'package:ggplot2':
##
##     margin
```

```
set.seed(12345)
trainingfile<-"https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv"
testingfile<-"https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv"
df<-read.csv(trainingfile,header = T,na.strings = c("","NA","#DIV/0!"))
```

Partition the dataset into training and testing set.

```
inTrain<-createDataPartition(df$classe,p=0.8,list=F)
training<-df[inTrain,]
testing<-df[-inTrain,]
```

Quick look at the data

```
summary(training)
```

```
##        X            user_name    raw_timestamp_part_1 raw_timestamp_part_2
##  Min.   :    2   adelmo  :3138   Min.   :1.322e+09    Min.   :   294
##  1st Qu.: 4900   carlitos:2512   1st Qu.:1.323e+09    1st Qu.:254380
##  Median : 9812   charles :2820   Median :1.323e+09    Median :500344
##  Mean   : 9813   eurico  :2429   Mean   :1.323e+09    Mean   :501554
##  3rd Qu.:14706   jeremy  :2718   3rd Qu.:1.323e+09    3rd Qu.:752309
##  Max.   :19621   pedro   :2082   Max.   :1.323e+09    Max.   :998801
##
##           cvtd_timestamp new_window    num_window      roll_belt
##  05/12/2011 11:24:1206   no :15384   Min.   :  1.0   Min.   :-28.90
##  28/11/2011 14:14:1192   yes:  315   1st Qu.:221.0   1st Qu.:  1.10
##  05/12/2011 11:25:1148               Median :424.0   Median :113.00
##  30/11/2011 17:11:1136               Mean   :430.9   Mean   : 64.48
```

```
##  02/12/2011 13:34:1117               3rd Qu.:645.0   3rd Qu.:123.00
##  02/12/2011 14:57:1098               Max.   :864.0   Max.   :162.00
##  (Other)         :8802
##    pitch_belt          yaw_belt       total_accel_belt kurtosis_roll_belt
##  Min.   :-55.8000   Min.   :-180.00   Min.   : 0.00    Min.   :-2.121
##  1st Qu.:  1.7100   1st Qu.: -88.30   1st Qu.: 3.00    1st Qu.:-1.328
##  Median :  5.2800   Median : -13.10   Median :17.00    Median :-0.884
##  Mean   :  0.2289   Mean   : -11.02   Mean   :11.32    Mean   :-0.282
##  3rd Qu.: 14.9000   3rd Qu.:  13.55   3rd Qu.:18.00    3rd Qu.:-0.221
##  Max.   : 60.3000   Max.   : 179.00   Max.   :29.00    Max.   :33.000
##                                                        NA's   :15390
##  kurtosis_picth_belt kurtosis_yaw_belt skewness_roll_belt
##  Min.   :-2.190      Mode:logical      Min.   :-5.745
##  1st Qu.:-1.124      NA's:15699        1st Qu.:-0.457
##  Median :-0.151                        Median : 0.000
##  Mean   : 4.515                        Mean   :-0.020
##  3rd Qu.: 3.663                        3rd Qu.: 0.422
##  Max.   :58.000                        Max.   : 2.713
##  NA's   :15411                         NA's   :15389
##  skewness_roll_belt.1 skewness_yaw_belt max_roll_belt     max_picth_belt
##  Min.   :-7.616       Mode:logical      Min.   :-94.300   Min.   : 3.00
##  1st Qu.:-0.987       NA's:15699        1st Qu.:-88.100   1st Qu.: 5.00
##  Median :-0.044                         Median : -9.700   Median :18.00
##  Mean   :-0.194                         Mean   : -9.502   Mean   :12.45
##  3rd Qu.: 0.758                         3rd Qu.: 11.000   3rd Qu.:19.00
##  Max.   : 7.348                         Max.   :180.000   Max.   :30.00
##  NA's   :15411                          NA's   :15384     NA's   :15384
##   max_yaw_belt   min_roll_belt     min_pitch_belt   min_yaw_belt
##  Min.   :-2.10   Min.   :-180.00   Min.   : 0.00   Min.   :-2.10
##  1st Qu.:-1.30   1st Qu.: -88.45   1st Qu.: 3.00   1st Qu.:-1.30
##  Median :-0.90   Median : -15.50   Median :16.00   Median :-0.90
##  Mean   :-0.28   Mean   : -13.75   Mean   :10.37   Mean   :-0.28
##  3rd Qu.:-0.20   3rd Qu.:   2.40   3rd Qu.:17.00   3rd Qu.:-0.20
##  Max.   :33.00   Max.   : 173.00   Max.   :22.00   Max.   :33.00
##  NA's   :15390   NA's   :15384     NA's   :15384   NA's   :15390
##  amplitude_roll_belt amplitude_pitch_belt amplitude_yaw_belt
##  Min.   :  0.000     Min.   : 0.000       Min.   :0
##  1st Qu.:  0.300     1st Qu.: 1.000       1st Qu.:0
##  Median :  1.000     Median : 1.000       Median :0
##  Mean   :  4.243     Mean   : 2.076       Mean   :0
##  3rd Qu.:  2.000     3rd Qu.: 2.000       3rd Qu.:0
##  Max.   :360.000     Max.   :12.000       Max.   :0
##  NA's   :15384       NA's   :15384        NA's   :15390
##  var_total_accel_belt avg_roll_belt    stddev_roll_belt var_roll_belt
##  Min.   : 0.000       Min.   :-27.40   Min.   : 0.000   Min.   :  0.000
##  1st Qu.: 0.100       1st Qu.:  1.10   1st Qu.: 0.157   1st Qu.:  0.000
##  Median : 0.200       Median :115.00   Median : 0.400   Median :  0.100
##  Mean   : 0.875       Mean   : 64.77   Mean   : 1.231   Mean   :  6.684
##  3rd Qu.: 0.300       3rd Qu.:123.00   3rd Qu.: 0.632   3rd Qu.:  0.400
##  Max.   :16.500       Max.   :157.00   Max.   :12.700   Max.   :160.200
##  NA's   :15384        NA's   :15384    NA's   :15384    NA's   :15384
##  avg_pitch_belt    stddev_pitch_belt var_pitch_belt    avg_yaw_belt
##  Min.   :-51.400   Min.   :0.000     Min.   : 0.000   Min.   :-138.30
##  1st Qu.:  2.200   1st Qu.:0.200     1st Qu.: 0.000   1st Qu.: -88.20
```

```
##  Median :  5.100   Median :0.400     Median : 0.100   Median : -12.20
##  Mean   :  0.375   Mean   :0.587     Mean   : 0.711   Mean   : -11.98
##  3rd Qu.: 14.200   3rd Qu.:0.700     3rd Qu.: 0.500   3rd Qu.:   3.85
##  Max.   : 59.700   Max.   :4.000     Max.   :16.200   Max.   : 173.50
##  NA's   :15384     NA's   :15384     NA's   :15384    NA's   :15384
##  stddev_yaw_belt    var_yaw_belt       gyros_belt_x
##  Min.   :  0.000   Min.   :    0.00   Min.   :-1.040000
##  1st Qu.:  0.100   1st Qu.:    0.01   1st Qu.:-0.030000
##  Median :  0.300   Median :    0.09   Median : 0.030000
##  Mean   :  1.529   Mean   :  138.19   Mean   :-0.005179
##  3rd Qu.:  0.650   3rd Qu.:    0.42   3rd Qu.: 0.110000
##  Max.   :176.600   Max.   :31183.24   Max.   : 2.220000
##  NA's   :15384     NA's   :15384
##   gyros_belt_y      gyros_belt_z      accel_belt_x       accel_belt_y
##  Min.   :-0.6400   Min.   :-1.4600   Min.   :-120.000   Min.   :-69.00
##  1st Qu.: 0.0000   1st Qu.:-0.2000   1st Qu.: -21.000   1st Qu.:  3.00
##  Median : 0.0200   Median :-0.1000   Median : -15.000   Median : 35.00
##  Mean   : 0.0394   Mean   :-0.1314   Mean   :  -5.498   Mean   : 30.16
##  3rd Qu.: 0.1100   3rd Qu.:-0.0200   3rd Qu.:  -5.000   3rd Qu.: 61.00
##  Max.   : 0.6400   Max.   : 1.6200   Max.   :  83.000   Max.   :164.00
##
##   accel_belt_z     magnet_belt_x    magnet_belt_y   magnet_belt_z
##  Min.   :-275.00   Min.   :-52.00   Min.   :354.0   Min.   :-620.0
##  1st Qu.:-162.00   1st Qu.:  9.00   1st Qu.:582.0   1st Qu.:-375.0
##  Median :-152.00   Median : 35.00   Median :601.0   Median :-319.0
##  Mean   : -72.71   Mean   : 55.77   Mean   :593.8   Mean   :-345.3
##  3rd Qu.:  27.00   3rd Qu.: 60.00   3rd Qu.:610.0   3rd Qu.:-306.0
##  Max.   : 105.00   Max.   :485.00   Max.   :673.0   Max.   : 293.0
##
##     roll_arm         pitch_arm          yaw_arm          total_accel_arm
##  Min.   :-180.00   Min.   :-88.200   Min.   :-180.0000   Min.   : 1.0
##  1st Qu.: -31.80   1st Qu.:-25.800   1st Qu.: -42.9000   1st Qu.:17.0
##  Median :   0.00   Median :  0.000   Median :   0.0000   Median :27.0
##  Mean   :  17.86   Mean   : -4.595   Mean   :  -0.6361   Mean   :25.5
##  3rd Qu.:  77.20   3rd Qu.: 11.150   3rd Qu.:  45.7500   3rd Qu.:33.0
##  Max.   : 180.00   Max.   : 88.500   Max.   : 180.0000   Max.   :66.0
##
##  var_accel_arm      avg_roll_arm     stddev_roll_arm   var_roll_arm
##  Min.   :  0.000   Min.   :-166.67   Min.   : 0.000   Min.   :   0.000
##  1st Qu.:  9.247   1st Qu.: -40.86   1st Qu.: 1.117   1st Qu.:   1.248
##  Median : 40.562   Median :   0.00   Median : 5.455   Median :  29.754
##  Mean   : 52.361   Mean   :  12.58   Mean   :10.144   Mean   : 270.296
##  3rd Qu.: 74.166   3rd Qu.:  78.23   3rd Qu.:14.919   3rd Qu.: 222.588
##  Max.   :253.010   Max.   : 163.33   Max.   :90.813   Max.   :8246.941
##  NA's   :15384     NA's   :15384     NA's   :15384    NA's   :15384
##  avg_pitch_arm     stddev_pitch_arm var_pitch_arm       avg_yaw_arm
##  Min.   :-81.773   Min.   : 0.000   Min.   :   0.000   Min.   :-173.440
##  1st Qu.:-20.594   1st Qu.: 1.218   1st Qu.:   1.487   1st Qu.: -25.953
##  Median :  0.000   Median : 7.934   Median :  62.952   Median :   0.000
##  Mean   : -3.906   Mean   :10.143   Mean   : 189.643   Mean   :   2.465
##  3rd Qu.:  9.559   3rd Qu.:16.244   3rd Qu.: 263.856   3rd Qu.:  40.558
##  Max.   : 75.659   Max.   :43.412   Max.   :1884.565   Max.   : 152.000
##  NA's   :15384     NA's   :15384    NA's   :15384      NA's   :15384
##  stddev_yaw_arm     var_yaw_arm         gyros_arm_x
```

```
##  Min.   :  0.000   Min.   :    0.000   Min.   :-6.37000
##  1st Qu.:  1.699   1st Qu.:    2.888   1st Qu.:-1.35000
##  Median : 16.389   Median :  268.588   Median : 0.08000
##  Mean   : 22.535   Mean   : 1132.497   Mean   : 0.03818
##  3rd Qu.: 36.270   3rd Qu.: 1315.545   3rd Qu.: 1.56000
##  Max.   :177.044   Max.   :31344.568   Max.   : 4.87000
##  NA's   :15384     NA's   :15384
##   gyros_arm_y       gyros_arm_z       accel_arm_x       accel_arm_y
##  Min.   :-3.4400   Min.   :-2.3300   Min.   :-404.00   Min.   :-318.00
##  1st Qu.:-0.8000   1st Qu.:-0.0700   1st Qu.:-241.50   1st Qu.: -54.00
##  Median :-0.2400   Median : 0.2300   Median : -43.00   Median :  13.00
##  Mean   :-0.2564   Mean   : 0.2702   Mean   : -60.14   Mean   :  32.25
##  3rd Qu.: 0.1400   3rd Qu.: 0.7200   3rd Qu.:  83.00   3rd Qu.: 138.00
##  Max.   : 2.8400   Max.   : 3.0200   Max.   : 437.00   Max.   : 308.00
##
##   accel_arm_z       magnet_arm_x     magnet_arm_y     magnet_arm_z
##  Min.   :-636.00   Min.   :-584.0   Min.   :-392.0   Min.   :-596.0
##  1st Qu.:-144.00   1st Qu.:-299.0   1st Qu.: -10.0   1st Qu.: 131.0
##  Median : -47.00   Median : 292.0   Median : 201.0   Median : 443.0
##  Mean   : -71.57   Mean   : 192.6   Mean   : 156.2   Mean   : 305.8
##  3rd Qu.:  23.00   3rd Qu.: 637.0   3rd Qu.: 323.0   3rd Qu.: 544.0
##  Max.   : 292.00   Max.   : 782.0   Max.   : 583.0   Max.   : 694.0
##
##  kurtosis_roll_arm kurtosis_picth_arm kurtosis_yaw_arm skewness_roll_arm
##  Min.   :-1.809    Min.   :-2.084     Min.   :-2.103   Min.   :-2.541
##  1st Qu.:-1.351    1st Qu.:-1.286     1st Qu.:-1.220   1st Qu.:-0.559
##  Median :-0.903    Median :-1.013     Median :-0.745   Median : 0.024
##  Mean   :-0.310    Mean   :-0.582     Mean   : 0.351   Mean   : 0.052
##  3rd Qu.: 0.014    3rd Qu.:-0.460     3rd Qu.: 0.070   3rd Qu.: 0.649
##  Max.   :21.456    Max.   :19.751     Max.   :56.000   Max.   : 4.394
##  NA's   :15446     NA's   :15448      NA's   :15394    NA's   :15445
##  skewness_pitch_arm skewness_yaw_arm  max_roll_arm    max_picth_arm
##  Min.   :-4.565     Min.   :-6.557   Min.   :-73.10   Min.   :-173.00
##  1st Qu.:-0.589     1st Qu.:-0.747   1st Qu.:  0.00   1st Qu.:  -0.40
##  Median :-0.045     Median :-0.141   Median :  5.50   Median :  24.40
##  Mean   :-0.068     Mean   :-0.253   Mean   : 11.85   Mean   :  35.47
##  3rd Qu.: 0.451     3rd Qu.: 0.275   3rd Qu.: 27.35   3rd Qu.:  96.70
##  Max.   : 2.655     Max.   : 7.483   Max.   : 85.50   Max.   : 180.00
##  NA's   :15448      NA's   :15394    NA's   :15384    NA's   :15384
##   max_yaw_arm     min_roll_arm    min_pitch_arm      min_yaw_arm
##  Min.   : 4.00   Min.   :-89.10   Min.   :-180.00   Min.   : 2.00
##  1st Qu.:30.00   1st Qu.:-41.35   1st Qu.: -72.35   1st Qu.: 8.00
##  Median :34.00   Median :-22.10   Median : -33.30   Median :13.00
##  Mean   :35.54   Mean   :-20.05   Mean   : -34.99   Mean   :14.75
##  3rd Qu.:41.00   3rd Qu.:  0.00   3rd Qu.:   0.00   3rd Qu.:19.00
##  Max.   :62.00   Max.   : 66.40   Max.   : 152.00   Max.   :38.00
##  NA's   :15384   NA's   :15384    NA's   :15384     NA's   :15384
##  amplitude_roll_arm amplitude_pitch_arm amplitude_yaw_arm
##  Min.   :  0.00     Min.   :  0.00      Min.   : 0.00
##  1st Qu.:  4.25     1st Qu.:  7.90      1st Qu.:13.00
##  Median : 28.31     Median : 54.00      Median :22.00
##  Mean   : 31.90     Mean   : 70.46      Mean   :20.79
##  3rd Qu.: 51.30     3rd Qu.:117.95      3rd Qu.:29.00
##  Max.   :119.50     Max.   :360.00      Max.   :52.00
```

```
##  NA's   :15384      NA's   :15384       NA's   :15384
##  roll_dumbbell     pitch_dumbbell     yaw_dumbbell
##  Min.   :-153.71   Min.   :-149.59   Min.   :-150.871
##  1st Qu.: -17.93   1st Qu.: -40.43   1st Qu.: -77.772
##  Median :  48.33   Median : -20.90   Median :  -5.400
##  Mean   :  24.10   Mean   : -10.79   Mean   :   1.167
##  3rd Qu.:  67.99   3rd Qu.:  17.19   3rd Qu.:  78.913
##  Max.   : 153.55   Max.   : 137.03   Max.   : 154.952
##
##  kurtosis_roll_dumbbell kurtosis_picth_dumbbell kurtosis_yaw_dumbbell
##  Min.   :-2.174         Min.   :-2.200          Mode:logical
##  1st Qu.:-0.718         1st Qu.:-0.695          NA's:15699
##  Median :-0.045         Median :-0.108
##  Mean   : 0.512         Mean   : 0.396
##  3rd Qu.: 1.064         3rd Qu.: 0.654
##  Max.   :54.998         Max.   :55.628
##  NA's   :15388          NA's   :15386
##  skewness_roll_dumbbell skewness_pitch_dumbbell skewness_yaw_dumbbell
##  Min.   :-7.384         Min.   :-7.447          Mode:logical
##  1st Qu.:-0.610         1st Qu.:-0.498          NA's:15699
##  Median :-0.111         Median :-0.081
##  Mean   :-0.132         Mean   :-0.007
##  3rd Qu.: 0.400         3rd Qu.: 0.535
##  Max.   : 1.958         Max.   : 3.769
##  NA's   :15387          NA's   :15385
##  max_roll_dumbbell max_picth_dumbbell max_yaw_dumbbell min_roll_dumbbell
##  Min.   :-70.10    Min.   :-104.50    Min.   :-2.20    Min.   :-134.90
##  1st Qu.:-27.30    1st Qu.: -65.85    1st Qu.:-0.70    1st Qu.: -60.90
##  Median : 12.30    Median :  43.60    Median : 0.00    Median : -45.10
##  Mean   : 13.45    Mean   :  32.35    Mean   : 0.51    Mean   : -42.23
##  3rd Qu.: 50.80    3rd Qu.: 129.80    3rd Qu.: 1.10    3rd Qu.: -26.55
##  Max.   :137.00    Max.   : 155.00    Max.   :55.00    Max.   :  73.20
##  NA's   :15384     NA's   :15384      NA's   :15388    NA's   :15384
##  min_pitch_dumbbell min_yaw_dumbbell amplitude_roll_dumbbell
##  Min.   :-147.00    Min.   :-2.20    Min.   :  0.00
##  1st Qu.: -92.05    1st Qu.:-0.70    1st Qu.: 16.07
##  Median : -69.20    Median : 0.00    Median : 35.80
##  Mean   : -35.30    Mean   : 0.51    Mean   : 55.68
##  3rd Qu.:  15.50    3rd Qu.: 1.10    3rd Qu.: 84.38
##  Max.   : 120.90    Max.   :55.00    Max.   :256.48
##  NA's   :15384      NA's   :15388    NA's   :15384
##  amplitude_pitch_dumbbell amplitude_yaw_dumbbell total_accel_dumbbell
##  Min.   :  0.00           Min.   :0              Min.   : 0.00
##  1st Qu.: 17.39           1st Qu.:0              1st Qu.: 4.00
##  Median : 42.99           Median :0              Median :10.00
##  Mean   : 67.65           Mean   :0              Mean   :13.72
##  3rd Qu.:101.41           3rd Qu.:0              3rd Qu.:19.00
##  Max.   :273.59           Max.   :0              Max.   :58.00
##  NA's   :15384            NA's   :15388
##  var_accel_dumbbell avg_roll_dumbbell  stddev_roll_dumbbell
##  Min.   :  0.000    Min.   :-128.963   Min.   :  0.000
##  1st Qu.:  0.400    1st Qu.:  -8.566   1st Qu.:  4.679
##  Median :  1.068    Median :  50.918   Median : 12.297
##  Mean   :  4.677    Mean   :  25.651   Mean   : 21.236
```

```
##  3rd Qu.:  3.469    3rd Qu.:  64.860   3rd Qu.: 26.494
##  Max.   :230.428    Max.   : 125.993   Max.   :123.778
##  NA's   :15384      NA's   :15384      NA's   :15384
##  var_roll_dumbbell  avg_pitch_dumbbell stddev_pitch_dumbbell
##  Min.   :    0.00   Min.   :-70.73     Min.   : 0.000
##  1st Qu.:   21.89   1st Qu.:-42.73     1st Qu.: 3.626
##  Median :  151.21   Median :-20.80     Median : 8.285
##  Mean   : 1068.52   Mean   :-12.71     Mean   :13.425
##  3rd Qu.:  701.92   3rd Qu.: 13.41     3rd Qu.:19.316
##  Max.   :15321.01   Max.   : 94.28     Max.   :82.680
##  NA's   :15384      NA's   :15384      NA's   :15384
##  var_pitch_dumbbell avg_yaw_dumbbell   stddev_yaw_dumbbell
##  Min.   :   0.00    Min.   :-114.242   Min.   : 0.000
##  1st Qu.:  13.15    1st Qu.: -76.606   1st Qu.: 3.931
##  Median :  68.64    Median :  -2.694   Median :10.687
##  Mean   : 364.19    Mean   :  -1.300   Mean   :16.841
##  3rd Qu.: 373.10    3rd Qu.:  66.187   3rd Qu.:25.177
##  Max.   :6836.02    Max.   : 134.905   Max.   :99.563
##  NA's   :15384      NA's   :15384      NA's   :15384
##  var_yaw_dumbbell  gyros_dumbbell_x    gyros_dumbbell_y
##  Min.   :   0.00   Min.   :-204.0000   Min.   :-2.10000
##  1st Qu.:  15.45   1st Qu.:  -0.0300   1st Qu.:-0.14000
##  Median : 114.21   Median :   0.1300   Median : 0.03000
##  Mean   : 594.19   Mean   :   0.1579   Mean   : 0.04543
##  3rd Qu.: 633.86   3rd Qu.:   0.3500   3rd Qu.: 0.21000
##  Max.   :9912.85   Max.   :   2.2200   Max.   :52.00000
##  NA's   :15384
##  gyros_dumbbell_z   accel_dumbbell_x  accel_dumbbell_y  accel_dumbbell_z
##  Min.   : -2.3800   Min.   :-419.00   Min.   :-182.00   Min.   :-334.00
##  1st Qu.: -0.3100   1st Qu.: -51.00   1st Qu.:  -8.00   1st Qu.:-142.00
##  Median : -0.1300   Median :  -8.00   Median :  42.00   Median :  -2.00
##  Mean   : -0.1249   Mean   : -28.67   Mean   :  52.75   Mean   : -38.82
##  3rd Qu.:  0.0300   3rd Qu.:  11.00   3rd Qu.: 110.00   3rd Qu.:  37.00
##  Max.   :317.0000   Max.   : 235.00   Max.   : 315.00   Max.   : 318.00
##
##  magnet_dumbbell_x magnet_dumbbell_y magnet_dumbbell_z  roll_forearm
##  Min.   :-643.0    Min.   :-744.0    Min.   :-249.00   Min.   :-180.000
##  1st Qu.:-536.0    1st Qu.: 231.0    1st Qu.: -45.00   1st Qu.:  -0.665
##  Median :-480.0    Median : 311.0    Median :  13.00   Median :  21.100
##  Mean   :-329.3    Mean   : 221.4    Mean   :  45.42   Mean   :  33.875
##  3rd Qu.:-306.0    3rd Qu.: 390.0    3rd Qu.:  95.00   3rd Qu.: 140.000
##  Max.   : 584.0    Max.   : 632.0    Max.   : 452.00   Max.   : 180.000
##
##  pitch_forearm     yaw_forearm      kurtosis_roll_forearm
##  Min.   :-72.50   Min.   :-180.00   Min.   :-1.879
##  1st Qu.:  0.00   1st Qu.: -68.90   1st Qu.:-1.384
##  Median :  9.15   Median :   0.00   Median :-1.127
##  Mean   : 10.76   Mean   :  19.12   Mean   :-0.625
##  3rd Qu.: 28.40   3rd Qu.: 110.00   3rd Qu.:-0.595
##  Max.   : 89.80   Max.   : 180.00   Max.   :40.060
##                                     NA's   :15448
##  kurtosis_picth_forearm kurtosis_yaw_forearm skewness_roll_forearm
##  Min.   :-2.082         Mode:logical         Min.   :-2.297
##  1st Qu.:-1.386         NA's:15699           1st Qu.:-0.392
```

```
##  Median :-0.893                                Median : 0.003
##  Mean   : 0.371                                Mean   : 0.013
##  3rd Qu.: 0.127                                3rd Qu.: 0.373
##  Max.   :33.626                                Max.   : 5.856
##  NA's   :15449                                 NA's   :15447
##  skewness_pitch_forearm skewness_yaw_forearm max_roll_forearm
##  Min.   :-5.241         Mode:logical         Min.   :-66.60
##  1st Qu.:-0.870         NA's:15699           1st Qu.:  0.00
##  Median :-0.193                              Median : 27.70
##  Mean   :-0.256                              Mean   : 24.28
##  3rd Qu.: 0.482                              3rd Qu.: 45.10
##  Max.   : 3.600                              Max.   : 87.90
##  NA's   :15449                               NA's   :15384
##  max_picth_forearm max_yaw_forearm  min_roll_forearm  min_pitch_forearm
##  Min.   :-151.00   Min.   :-1.900   Min.   :-72.500   Min.   :-180.00
##  1st Qu.:   0.00   1st Qu.:-1.400   1st Qu.: -6.550   1st Qu.:-175.00
##  Median : 112.00   Median :-1.100   Median :  0.000   Median : -64.60
##  Mean   :  82.51   Mean   :-0.625   Mean   : -0.975   Mean   : -58.63
##  3rd Qu.: 175.00   3rd Qu.:-0.600   3rd Qu.: 12.550   3rd Qu.:   0.00
##  Max.   : 180.00   Max.   :40.100   Max.   : 60.400   Max.   : 167.00
##  NA's   :15384     NA's   :15448    NA's   :15384     NA's   :15384
##  min_yaw_forearm  amplitude_roll_forearm amplitude_pitch_forearm
##  Min.   :-1.900   Min.   :  0.000       Min.   :  0.0
##  1st Qu.:-1.400   1st Qu.:  1.505       1st Qu.:  2.0
##  Median :-1.100   Median : 19.100       Median : 85.5
##  Mean   :-0.625   Mean   : 25.251       Mean   :141.1
##  3rd Qu.:-0.600   3rd Qu.: 40.355       3rd Qu.:350.0
##  Max.   :40.100   Max.   :126.000       Max.   :360.0
##  NA's   :15448    NA's   :15384         NA's   :15384
##  amplitude_yaw_forearm total_accel_forearm var_accel_forearm
##  Min.   :0             Min.   :  0.00      Min.   :  0.000
##  1st Qu.:0             1st Qu.: 29.00      1st Qu.:  6.876
##  Median :0             Median : 36.00      Median : 22.765
##  Mean   :0             Mean   : 34.75      Mean   : 33.485
##  3rd Qu.:0             3rd Qu.: 41.00      3rd Qu.: 50.342
##  Max.   :0             Max.   :108.00      Max.   :172.606
##  NA's   :15448                             NA's   :15384
##  avg_roll_forearm   stddev_roll_forearm var_roll_forearm
##  Min.   :-177.130   Min.   :  0.000     Min.   :    0.00
##  1st Qu.:  -0.382   1st Qu.:  0.456     1st Qu.:    0.21
##  Median :  11.537   Median :  8.230     Median :   67.74
##  Mean   :  33.369   Mean   : 42.747     Mean   : 5375.68
##  3rd Qu.: 111.059   3rd Qu.: 88.626     3rd Qu.: 7856.94
##  Max.   : 177.256   Max.   :179.171     Max.   :32102.24
##  NA's   :15384      NA's   :15384       NA's   :15384
##  avg_pitch_forearm stddev_pitch_forearm var_pitch_forearm
##  Min.   :-68.17    Min.   : 0.000       Min.   :   0.000
##  1st Qu.:  0.00    1st Qu.: 0.426       1st Qu.:   0.182
##  Median : 11.94    Median : 5.971       Median :  35.649
##  Mean   : 11.16    Mean   : 8.110       Mean   : 141.373
##  3rd Qu.: 27.77    3rd Qu.:12.685       3rd Qu.: 160.919
##  Max.   : 70.15    Max.   :47.745       Max.   :2279.617
##  NA's   :15384     NA's   :15384        NA's   :15384
##  avg_yaw_forearm   stddev_yaw_forearm var_yaw_forearm
```

```
##  Min.   :-155.06   Min.   :  0.000    Min.   :    0.00
##  1st Qu.: -25.32   1st Qu.:  0.602    1st Qu.:    0.36
##  Median :   0.00   Median : 25.061    Median :  628.06
##  Mean   :  18.13   Mean   : 45.426    Mean   : 4709.58
##  3rd Qu.:  85.58   3rd Qu.: 87.787    3rd Qu.: 7707.29
##  Max.   : 169.24   Max.   :197.508    Max.   :39009.33
##  NA's   :15384     NA's   :15384      NA's   :15384
##  gyros_forearm_x    gyros_forearm_y     gyros_forearm_z
##  Min.   :-22.0000   Min.   : -7.02000   Min.   : -7.9400
##  1st Qu.: -0.2150   1st Qu.: -1.48000   1st Qu.: -0.1800
##  Median :  0.0500   Median :  0.03000   Median :  0.0800
##  Mean   :  0.1584   Mean   :  0.08375   Mean   :  0.1522
##  3rd Qu.:  0.5800   3rd Qu.:  1.62000   3rd Qu.:  0.4900
##  Max.   :  3.9700   Max.   :311.00000   Max.   :231.0000
##
##  accel_forearm_x   accel_forearm_y  accel_forearm_z   magnet_forearm_x
##  Min.   :-498.00   Min.   :-632.0   Min.   :-446.00   Min.   :-1280.0
##  1st Qu.:-179.00   1st Qu.:  55.0   1st Qu.:-182.00   1st Qu.: -616.0
##  Median : -57.00   Median : 201.0   Median : -40.00   Median : -382.0
##  Mean   : -62.63   Mean   : 163.5   Mean   : -55.79   Mean   : -314.2
##  3rd Qu.:  76.00   3rd Qu.: 312.0   3rd Qu.:  26.00   3rd Qu.:  -78.0
##  Max.   : 389.00   Max.   : 923.0   Max.   : 291.00   Max.   :  672.0
##
##  magnet_forearm_y magnet_forearm_z classe
##  Min.   :-890.0   Min.   :-973.0   A:4464
##  1st Qu.:   0.5   1st Qu.: 199.0   B:3038
##  Median : 589.0   Median : 511.0   C:2738
##  Mean   : 379.0   Mean   : 395.4   D:2573
##  3rd Qu.: 737.0   3rd Qu.: 653.0   E:2886
##  Max.   :1480.0   Max.   :1080.0
##
```

There are many colnums with na values or values that are not relevant. Remove columns with na values and columnes that are not relevant.

```
nafilter<-!(colSums(is.na(training))>0) #& !colSums(training=="NA")>0

training<-training[,nafilter]
training<-training[-c(1,3,4,5)]

testing<-testing[,nafilter]
testing<-testing[-c(1,3,4,5)]
```

look at the dimmision with the data

```
dim(training)
```

```
## [1] 15699    56
```

There are only 56 columnes reserved.

Firstly try a decion tree model, which hopefully will give some insights.

```
model1<-rpart(classe~., data=training,method="class")
rtrain<-predict(model1,newdata=training,type="class")
confusionMatrix(rtrain,training$classe)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 4010  657  109  282  208
##          B   97 1721  178   91  255
##          C   45  203 2228  370  228
##          D  255  365  149 1702  334
##          E   57   92   74  128 1861
##
## Overall Statistics
##
##                Accuracy : 0.7339
##                  95% CI : (0.7269, 0.7408)
##     No Information Rate : 0.2843
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.6617
##  Mcnemar's Test P-Value : < 2.2e-16
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            0.8983   0.5665   0.8137   0.6615   0.6448
## Specificity            0.8882   0.9510   0.9347   0.9160   0.9726
## Pos Pred Value         0.7615   0.7348   0.7248   0.6068   0.8413
## Neg Pred Value         0.9565   0.9014   0.9596   0.9324   0.9240
## Prevalence             0.2843   0.1935   0.1744   0.1639   0.1838
## Detection Rate         0.2554   0.1096   0.1419   0.1084   0.1185
## Detection Prevalence   0.3354   0.1492   0.1958   0.1787   0.1409
## Balanced Accuracy      0.8933   0.7587   0.8742   0.7887   0.8087
```

```
rtest<-predict(model1,newdata=testing,type="class")
confusionMatrix(rtest,testing$classe)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1012  173   24   79   44
##          B   26  421   60   27   70
##          C   14   41  544   99   60
##          D   45   97   40  419   91
##          E   19   27   16   19  456
##
## Overall Statistics
##
##                Accuracy : 0.727
```

```
##                  95% CI : (0.7128, 0.7409)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.6526
##  Mcnemar's Test P-Value : < 2.2e-16
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            0.9068   0.5547   0.7953   0.6516   0.6325
## Specificity            0.8860   0.9422   0.9339   0.9168   0.9747
## Pos Pred Value         0.7598   0.6970   0.7177   0.6055   0.8492
## Neg Pred Value         0.9599   0.8982   0.9558   0.9307   0.9217
## Prevalence             0.2845   0.1935   0.1744   0.1639   0.1838
## Detection Rate         0.2580   0.1073   0.1387   0.1068   0.1162
## Detection Prevalence   0.3395   0.1540   0.1932   0.1764   0.1369
## Balanced Accuracy      0.8964   0.7484   0.8646   0.7842   0.8036
```

The results are pretty ugly even on the training set. try with a more complicated random forrest model.

```
#model2<-train(classe~., method="rf",data=training,tcControl=tc)
print("random forest result#########")
```

```
## [1] "random forest result#########"
```

```
model2<-randomForest(classe~., data=training,ntree=2000)
confusionMatrix(predict(model2,newdata=training),training$classe)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 4464    0    0    0    0
##          B    0 3038    0    0    0
##          C    0    0 2738    0    0
##          D    0    0    0 2573    0
##          E    0    0    0    0 2886
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.9998, 1)
##     No Information Rate : 0.2843
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 1
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
```

```
## Sensitivity            1.0000   1.0000   1.0000   1.0000   1.0000
## Specificity            1.0000   1.0000   1.0000   1.0000   1.0000
## Pos Pred Value         1.0000   1.0000   1.0000   1.0000   1.0000
## Neg Pred Value         1.0000   1.0000   1.0000   1.0000   1.0000
## Prevalence             0.2843   0.1935   0.1744   0.1639   0.1838
## Detection Rate         0.2843   0.1935   0.1744   0.1639   0.1838
## Detection Prevalence   0.2843   0.1935   0.1744   0.1639   0.1838
## Balanced Accuracy      1.0000   1.0000   1.0000   1.0000   1.0000
```

```
confusionMatrix(predict(model2,newdata=testing),testing$classe)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1116    0    0    0    0
##          B    0  759    7    0    0
##          C    0    0  677    7    0
##          D    0    0    0  636    2
##          E    0    0    0    0  719
##
## Overall Statistics
##
##                Accuracy : 0.9959
##                  95% CI : (0.9934, 0.9977)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9948
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            1.0000   1.0000   0.9898   0.9891   0.9972
## Specificity            1.0000   0.9978   0.9978   0.9994   1.0000
## Pos Pred Value         1.0000   0.9909   0.9898   0.9969   1.0000
## Neg Pred Value         1.0000   1.0000   0.9978   0.9979   0.9994
## Prevalence             0.2845   0.1935   0.1744   0.1639   0.1838
## Detection Rate         0.2845   0.1935   0.1726   0.1621   0.1833
## Detection Prevalence   0.2845   0.1953   0.1744   0.1626   0.1833
## Balanced Accuracy      1.0000   0.9989   0.9938   0.9943   0.9986
```

results look promising and because it is random forrest, the expected out of sample result should be quite similar. load test data and make predicion.

```
va<-read.csv(testingfile,header=T)
levels(va$new_window) <- levels(training$new_window)
predictions<-predict(model2,newdata=va)

answers<-as.character(predictions)

pml_write_files = function(x){
```

```
  n = length(x)
  for(i in 1:n){
    filename = paste0("problem_id_",i,".txt")
    write.table(x[i],file=filename,quote=FALSE,row.names=FALSE,col.names=FALSE)
  }
}
```

pml_write_files(answers)