

Assignement Report

Zine Eddine Bouzennoune

Decemner 2nd, 2019

Human Activity Recognition

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, our goal will be to use data from accelerometers on the belt, forearm, arm, and dumbbell of 6 participants, to classify into 5 categories, the type of exercise people are doing.

First, we download the data that is available in these two links:

Test set:

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

Training set:

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

The whole data for this project come from this source:

<http://web.archive.org/web/20161224072740/http://groupware.les.inf.puc-rio.br/har>.

Discovering the data set

To do so, we should first load the necessary packages for our program.

```
library(skimr)
```

```
## Warning: package 'skimr' was built under R version 3.5.3
```

```
##
```

```
## Attaching package: 'skimr'
```

```
## The following object is masked from 'package:stats':
```

```
##
```

```
##      filter
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.2.1 --
```

```
## v ggplot2 3.0.0      v purrr   0.2.5
```

```
## v tibble  2.1.3      v dplyr  0.8.3
```

```
## v tidyr   1.0.0      v stringr 1.3.1
```

```
## v readr   1.1.1      v forcats 0.3.0
```

```
## Warning: package 'tibble' was built under R version 3.5.3
```

```
## Warning: package 'tidyr' was built under R version 3.5.3
```

```
## Warning: package 'dplyr' was built under R version 3.5.3
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks skimr::filter(), stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 3.5.3
## Loading required package: lattice
##
## Attaching package: 'caret'
## The following object is masked from 'package:purrr':
##
##     lift
```

The package 'skimr' was used to get a good summary of the dataset. The command that shows this is: `skim(dataset)`. Tidyverse was loaded in order to use the different packages and functions like `ggplot`, the pipe operator, and many other useful commands. Finally, the `caret` package was used to train the machine learning model.

loading the test and training datasets and setting the seed for reproducible results

```
set.seed(seed = 123)

setwd(dir='C:\\Users\\zineeddine.bouzenou\\OneDrive\\Coursera\\JHDSS\\Practical Machine Learning')

training <- read.csv('pml-training.csv', stringsAsFactors = F)

testing <- read.csv('pml-testing.csv', stringsAsFactors = F)
```

First, we use `skim` to have a good summary of the dataset:

```
skim(training)
```

```
## Skim summary statistics
##   n obs: 19622
##   n variables: 160
##
## -- Variable type:character -----
##           variable missing complete      n min max empty n_unique
##      amplitude_yaw_belt      0    19622 19622      0  7 19216         4
##      amplitude_yaw_dumbbell    0    19622 19622      0  7 19216         3
##      amplitude_yaw_forearm     0    19622 19622      0  7 19216         3
##           classe      0    19622 19622      1  1      0         5
##      cvtd_timestamp     0    19622 19622     16 16      0        20
##      kurtosis_picth_arm      0    19622 19622      0  8 19216       328
##      kurtosis_picth_belt     0    19622 19622      0  9 19216       317
##      kurtosis_picth_dumbbell  0    19622 19622      0  7 19216       401
##      kurtosis_picth_forearm   0    19622 19622      0  7 19216       323
##      kurtosis_roll_arm       0    19622 19622      0  8 19216       330
##      kurtosis_roll_belt      0    19622 19622      0  9 19216       397
##      kurtosis_roll_dumbbell   0    19622 19622      0  7 19216       398
##      kurtosis_roll_forearm    0    19622 19622      0  7 19216       322
##      kurtosis_yaw_arm        0    19622 19622      0  8 19216       395
##      kurtosis_yaw_belt       0    19622 19622      0  7 19216         2
##      kurtosis_yaw_dumbbell    0    19622 19622      0  7 19216         2
##      kurtosis_yaw_forearm     0    19622 19622      0  7 19216         2
##      max_yaw_belt           0    19622 19622      0  7 19216        68
```

```

##      max_yaw_dumbbell      0      19622 19622      0      7 19216      73
##      max_yaw_forearm      0      19622 19622      0      7 19216      45
##      min_yaw_belt      0      19622 19622      0      7 19216      68
##      min_yaw_dumbbell      0      19622 19622      0      7 19216      73
##      min_yaw_forearm      0      19622 19622      0      7 19216      45
##      new_window      0      19622 19622      2      3      0      2
##      skewness_pitch_arm      0      19622 19622      0      8 19216      328
##      skewness_pitch_dumbbell      0      19622 19622      0      7 19216      402
##      skewness_pitch_forearm      0      19622 19622      0      7 19216      319
##      skewness_roll_arm      0      19622 19622      0      8 19216      331
##      skewness_roll_belt      0      19622 19622      0      9 19216      395
##      skewness_roll_belt.1      0      19622 19622      0      9 19216      338
##      skewness_roll_dumbbell      0      19622 19622      0      7 19216      401
##      skewness_roll_forearm      0      19622 19622      0      7 19216      323
##      skewness_yaw_arm      0      19622 19622      0      8 19216      395
##      skewness_yaw_belt      0      19622 19622      0      7 19216      2
##      skewness_yaw_dumbbell      0      19622 19622      0      7 19216      2
##      skewness_yaw_forearm      0      19622 19622      0      7 19216      2
##      user_name      0      19622 19622      5      8      0      6
##
## -- Variable type:integer -----
##      variable missing complete      n      mean      sd
##      accel_arm_x      0      19622 19622     -60.24     182.04
##      accel_arm_y      0      19622 19622      32.6      109.87
##      accel_arm_z      0      19622 19622    -71.25     134.65
##      accel_belt_x      0      19622 19622     -5.59      29.64
##      accel_belt_y      0      19622 19622     30.15      28.58
##      accel_belt_z      0      19622 19622    -72.59     100.45
##      accel_dumbbell_x      0      19622 19622    -28.62      67.32
##      accel_dumbbell_y      0      19622 19622     52.63      80.75
##      accel_dumbbell_z      0      19622 19622    -38.32     109.47
##      accel_forearm_x      0      19622 19622    -61.65     180.59
##      accel_forearm_y      0      19622 19622    163.66     200.13
##      accel_forearm_z      0      19622 19622    -55.29     138.4
##      amplitude_pitch_belt 19216      406 19622      2.17      2.36
##      amplitude_yaw_arm 19216      406 19622     20.79     12.28
##      magnet_arm_x      0      19622 19622    191.72     443.64
##      magnet_arm_y      0      19622 19622    156.61     201.91
##      magnet_arm_z      0      19622 19622    306.49     326.62
##      magnet_belt_x      0      19622 19622     55.6      64.18
##      magnet_belt_y      0      19622 19622    593.68      35.68
##      magnet_belt_z      0      19622 19622   -345.48      65.21
##      magnet_dumbbell_x      0      19622 19622   -328.48     339.72
##      magnet_dumbbell_y      0      19622 19622    220.97     326.87
##      magnet_forearm_x      0      19622 19622   -312.58     346.96
##      max_pitch_belt 19216      406 19622     12.92      8.01
##      max_yaw_arm 19216      406 19622     35.46     10.45
##      min_pitch_belt 19216      406 19622     10.76      7.47
##      min_yaw_arm 19216      406 19622     14.67      9.11
##      num_window      0      19622 19622    430.64     247.91
##      raw_timestamp_part_1      0      19622 19622    1.3e+09    2e+05
##      raw_timestamp_part_2      0      19622 19622    5e+05    288222.88
##      total_accel_arm      0      19622 19622     25.51     10.52
##      total_accel_belt      0      19622 19622     11.31      7.74

```

```

## total_accel_dumbbell      0  19622 19622    13.72    10.23
## total_accel_forearm      0  19622 19622    34.72    10.06
##              X            0  19622 19622  9811.5    5664.53
##      p0      p25      p50      p75      p100      hist
## -404      -242      -44      84      437      <U+2581><U+2587><U+2582><U+2583><U+258
## -318      -54      14      139      308      <U+2581><U+2581><U+2583><U+2587><U+258
## -636      -143      -47      23      292      <U+2581><U+2581><U+2581><U+2582><U+258
## -120      -21      -15      -5      85      <U+2581><U+2581><U+2581><U+2587><U+258
## -69      3      35      61      164      <U+2581><U+2581><U+2587><U+2583><U+258
## -275      -162      -152      27      105      <U+2581><U+2581><U+2587><U+2581><U+258
## -419      -50      -8      11      235      <U+2581><U+2581><U+2581><U+2583><U+258
## -189      -8      41.5      111      315      <U+2581><U+2581><U+2587><U+2587><U+258
## -334      -142      -1      38      318      <U+2581><U+2583><U+2585><U+2582><U+258
## -498      -178      -57      76      477      <U+2582><U+2583><U+2585><U+2587><U+258
## -632      57      201      312      923      <U+2581><U+2581><U+2582><U+2586><U+258
## -446      -182      -39      26      291      <U+2581><U+2581><U+2587><U+2583><U+258
## 0      1      1      2      12      <U+2587><U+2583><U+2581><U+2581><U+258
## 0      13      22      28.75      52      <U+2585><U+2582><U+2583><U+2587><U+258
## -584      -300      289      637      782      <U+2582><U+2586><U+2582><U+2582><U+258
## -392      -9      202      323      583      <U+2581><U+2582><U+2583><U+2583><U+258
## -597      131.25      444      545      694      <U+2581><U+2581><U+2582><U+2582><U+258
## -52      9      35      59      485      <U+2586><U+2587><U+2581><U+2583><U+258
## 354      581      601      610      673      <U+2581><U+2581><U+2581><U+2581><U+258
## -623      -375      -320      -306      293      <U+2581><U+2582><U+2587><U+2581><U+258
## -643      -535      -479      -304      592      <U+2587><U+2585><U+2582><U+2581><U+258
## -3600      231      311      390      633      <U+2581><U+2581><U+2581><U+2581><U+258
## -1280      -616      -378      -73      672      <U+2581><U+2581><U+2587><U+2585><U+258
## 3      5      18      19      30      <U+2587><U+2581><U+2581><U+2581><U+258
## 4      29      34      41      65      <U+2581><U+2581><U+2582><U+2587><U+258
## 0      3      16      17      23      <U+2585><U+2586><U+2581><U+2581><U+258
## 1      8      13      19      38      <U+2585><U+2587><U+2586><U+2585><U+258
## 1      222      424      644      864      <U+2587><U+2587><U+2587><U+2587><U+258
## 1.3e+09      1.3e+09      1.3e+09      1.3e+09      1.3e+09      <U+2583><U+2581><U+2583><U+2581><U+258
## 294      252912.25      5e+05      751890.75      1e+06      <U+2587><U+2587><U+2587><U+2587><U+258
## 1      17      27      33      66      <U+2582><U+2585><U+2586><U+2587><U+258
## 0      3      17      18      29      <U+2587><U+2585><U+2581><U+2581><U+258
## 0      4      10      19      58      <U+2587><U+2585><U+2585><U+2582><U+258
## 0      29      36      41      108      <U+2581><U+2582><U+2587><U+2585><U+258
## 1      4906.25      9811.5      14716.75      19622      <U+2587><U+2587><U+2587><U+2587><U+258
##
## -- Variable type:numeric -----
##      variable missing complete      n      mean      sd
##      amplitude_pitch_arm      19216      406 19622      69.68      66.98
##      amplitude_pitch_dumbbell      19216      406 19622      65.93      65.23
##      amplitude_pitch_forearm      19216      406 19622      139.06      147.86
##      amplitude_roll_arm      19216      406 19622      32.45      27.39
##      amplitude_roll_belt      19216      406 19622      3.77      25.26
##      amplitude_roll_dumbbell      19216      406 19622      55      54.94
##      amplitude_roll_forearm      19216      406 19622      24.65      25.88
##      avg_pitch_arm      19216      406 19622      -4.9      26.83
##      avg_pitch_belt      19216      406 19622      0.52      22.41
##      avg_pitch_dumbbell      19216      406 19622      -12.33      32.06
##      avg_pitch_forearm      19216      406 19622      11.8      24.83
##      avg_roll_arm      19216      406 19622      12.68      68.58

```

##	avg_roll_belt	19216	406	19622	68.06	63.14
##	avg_roll_dumbbell	19216	406	19622	23.86	62.9
##	avg_roll_forearm	19216	406	19622	33.17	79.52
##	avg_yaw_arm	19216	406	19622	2.36	61.33
##	avg_yaw_belt	19216	406	19622	-8.83	93.48
##	avg_yaw_dumbbell	19216	406	19622	0.2	78.21
##	avg_yaw_forearm	19216	406	19622	18	77.56
##	gyros_arm_x	0	19622	19622	0.043	1.99
##	gyros_arm_y	0	19622	19622	-0.26	0.85
##	gyros_arm_z	0	19622	19622	0.27	0.55
##	gyros_belt_x	0	19622	19622	-0.0056	0.21
##	gyros_belt_y	0	19622	19622	0.04	0.078
##	gyros_belt_z	0	19622	19622	-0.13	0.24
##	gyros_dumbbell_x	0	19622	19622	0.16	1.51
##	gyros_dumbbell_y	0	19622	19622	0.046	0.61
##	gyros_dumbbell_z	0	19622	19622	-0.13	2.29
##	gyros_forearm_x	0	19622	19622	0.16	0.65
##	gyros_forearm_y	0	19622	19622	0.075	3.1
##	gyros_forearm_z	0	19622	19622	0.15	1.75
##	magnet_dumbbell_z	0	19622	19622	46.05	139.96
##	magnet_forearm_y	0	19622	19622	380.12	509.37
##	magnet_forearm_z	0	19622	19622	393.61	369.27
##	max_picth_arm	19216	406	19622	35.75	69.62
##	max_picth_dumbbell	19216	406	19622	32.75	93.37
##	max_picth_forearm	19216	406	19622	81.49	95.54
##	max_roll_arm	19216	406	19622	11.24	26.93
##	max_roll_belt	19216	406	19622	-6.67	94.59
##	max_roll_dumbbell	19216	406	19622	13.76	48.3
##	max_roll_forearm	19216	406	19622	24.49	31.04
##	min_pitch_arm	19216	406	19622	-33.92	60.83
##	min_pitch_dumbbell	19216	406	19622	-33.18	74.28
##	min_pitch_forearm	19216	406	19622	-57.57	110.74
##	min_roll_arm	19216	406	19622	-21.22	28.72
##	min_roll_belt	19216	406	19622	-10.44	93.62
##	min_roll_dumbbell	19216	406	19622	-41.24	34.71
##	min_roll_forearm	19216	406	19622	-0.17	22.59
##	pitch_arm	0	19622	19622	-4.61	30.68
##	pitch_belt	0	19622	19622	0.31	22.35
##	pitch_dumbbell	0	19622	19622	-10.78	36.99
##	pitch_forearm	0	19622	19622	10.71	28.15
##	roll_arm	0	19622	19622	17.83	72.74
##	roll_belt	0	19622	19622	64.41	62.75
##	roll_dumbbell	0	19622	19622	23.84	69.93
##	roll_forearm	0	19622	19622	33.83	108.04
##	stddev_pitch_arm	19216	406	19622	10.38	9.4
##	stddev_pitch_belt	19216	406	19622	0.6	0.64
##	stddev_pitch_dumbbell	19216	406	19622	13.15	13.34
##	stddev_pitch_forearm	19216	406	19622	7.98	8.73
##	stddev_roll_arm	19216	406	19622	11.2	17.1
##	stddev_roll_belt	19216	406	19622	1.34	2.44
##	stddev_roll_dumbbell	19216	406	19622	20.76	24.3
##	stddev_roll_forearm	19216	406	19622	41.99	59.33
##	stddev_yaw_arm	19216	406	19622	22.27	23.69
##	stddev_yaw_belt	19216	406	19622	1.34	10.29

##	stddev_yaw_dumbbell	19216	406	19622	16.65	17.71
##	stddev_yaw_forearm	19216	406	19622	44.85	51.33
##	var_accel_arm	19216	406	19622	53.24	53.98
##	var_accel_dumbbell	19216	406	19622	4.39	13.51
##	var_accel_forearm	19216	406	19622	33.5	33.95
##	var_pitch_arm	19216	406	19622	195.86	292.6
##	var_pitch_belt	19216	406	19622	0.77	1.76
##	var_pitch_dumbbell	19216	406	19622	350.31	673.96
##	var_pitch_forearm	19216	406	19622	139.59	266.49
##	var_roll_arm	19216	406	19622	417.26	2007.16
##	var_roll_belt	19216	406	19622	7.7	23.16
##	var_roll_dumbbell	19216	406	19622	1020.27	2262.56
##	var_roll_forearm	19216	406	19622	5274.1	9177.18
##	var_total_accel_belt	19216	406	19622	0.93	2.22
##	var_yaw_arm	19216	406	19622	1055.93	2722.17
##	var_yaw_belt	19216	406	19622	107.49	1655.52
##	var_yaw_dumbbell	19216	406	19622	589.84	1244.59
##	var_yaw_forearm	19216	406	19622	4639.85	7284.97
##	yaw_arm	0	19622	19622	-0.62	71.36
##	yaw_belt	0	19622	19622	-11.21	95.19
##	yaw_dumbbell	0	19622	19622	1.67	82.52
##	yaw_forearm	0	19622	19622	19.21	103.22
##	p0	p25	p50	p75	p100	hist
##	0	9.93	54.9	115.17	360	<U+2587><U+2585><U+2583><U+2582><U+2581><U+2581><U+2581><U+
##	0	17.06	41.72	99.54	273.59	<U+2587><U+2585><U+2582><U+2582><U+2582><U+2581><U+2581><U+
##	0	2	83.7	350	360	<U+2587><U+2582><U+2582><U+2581><U+2581><U+2581><U+2581><U+
##	0	5.43	28.45	50.96	119.5	<U+2587><U+2586><U+2585><U+2585><U+2582><U+2581><U+2581><U+
##	0	0.3	1	2.08	360	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	14.97	35.05	81.04	256.48	<U+2587><U+2583><U+2582><U+2581><U+2581><U+2581><U+2581><U+
##	0	1.12	17.77	39.88	126	<U+2587><U+2583><U+2583><U+2582><U+2581><U+2581><U+2581><U+
##	-81.77	-22.77	0	8.28	75.66	<U+2581><U+2581><U+2585><U+2585><U+2587><U+2582><U+2581><U+
##	-51.4	2.02	5.2	15.78	59.7	<U+2582><U+2581><U+2581><U+2583><U+2587><U+2583><U+2581><U+
##	-70.73	-42	-19.9	13.21	94.28	<U+2583><U+2587><U+2587><U+2585><U+2586><U+2582><U+2581><U+
##	-68.17	0	12.02	28.48	72.09	<U+2581><U+2581><U+2581><U+2587><U+2587><U+2586><U+2583><U+
##	-166.67	-38.37	0	76.32	163.33	<U+2581><U+2581><U+2585><U+2585><U+2587><U+2583><U+2585><U+
##	-27.4	1.1	116.35	123.38	157.4	<U+2581><U+2587><U+2581><U+2581><U+2581><U+2581><U+2587><U+
##	-128.96	-12.33	48.23	64.37	125.99	<U+2582><U+2582><U+2582><U+2582><U+2582><U+2587><U+2583><U+
##	-177.23	-0.91	11.17	107.13	177.26	<U+2581><U+2581><U+2582><U+2587><U+2583><U+2582><U+2583><U+
##	-173.44	-29.2	0	38.18	152	<U+2581><U+2581><U+2582><U+2585><U+2587><U+2583><U+2582><U+
##	-138.3	-88.18	-6.55	14.12	173.5	<U+2581><U+2587><U+2581><U+2585><U+2581><U+2581><U+2581><U+
##	-117.95	-76.7	-4.5	71.23	134.91	<U+2585><U+2587><U+2582><U+2582><U+2585><U+2583><U+2583><U+
##	-155.06	-26.26	0	85.79	169.24	<U+2581><U+2582><U+2582><U+2587><U+2582><U+2583><U+2585><U+
##	-6.37	-1.33	0.08	1.57	4.87	<U+2581><U+2581><U+2583><U+2583><U+2587><U+2585><U+2583><U+
##	-3.44	-0.8	-0.24	0.14	2.84	<U+2581><U+2581><U+2583><U+2587><U+2587><U+2583><U+2581><U+
##	-2.33	-0.07	0.23	0.72	3.02	<U+2581><U+2581><U+2582><U+2587><U+2587><U+2582><U+2581><U+
##	-1.04	-0.03	0.03	0.11	2.22	<U+2581><U+2581><U+2587><U+2581><U+2581><U+2581><U+2581><U+
##	-0.64	0	0.02	0.11	0.64	<U+2581><U+2581><U+2581><U+2587><U+2587><U+2581><U+2581><U+
##	-1.46	-0.2	-0.1	-0.02	1.62	<U+2581><U+2581><U+2582><U+2587><U+2581><U+2581><U+2581><U+
##	-204	-0.03	0.13	0.35	2.22	<U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	-2.1	-0.14	0.03	0.21	52	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	-2.38	-0.31	-0.13	0.03	317	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	-22	-0.22	0.05	0.56	3.97	<U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+2587><U+
##	-7.02	-1.46	0.03	1.62	311	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	-8.09	-0.18	0.08	0.49	231	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+

##	-262	-45	13	95	452	<U+2581><U+2582><U+2587><U+2586><U+2582><U+2581><U+2582><U+
##	-896	2	591	737	1480	<U+2581><U+2582><U+2582><U+2582><U+2583><U+2587><U+2581><U+
##	-973	191	511	653	1090	<U+2581><U+2581><U+2581><U+2582><U+2582><U+2587><U+2587><U+
##	-173	-1.98	23.25	95.97	180	<U+2581><U+2581><U+2582><U+2587><U+2582><U+2583><U+2585><U+
##	-112.9	-66.7	40.05	133.23	155	<U+2583><U+2585><U+2582><U+2581><U+2582><U+2582><U+2582><U+
##	-151	0	113	174.75	180	<U+2581><U+2581><U+2582><U+2585><U+2581><U+2581><U+2583><U+
##	-73.1	-0.18	4.95	26.78	85.5	<U+2581><U+2581><U+2582><U+2587><U+2585><U+2583><U+2581><U+
##	-94.3	-88	-5.1	18.5	180	<U+2587><U+2581><U+2585><U+2581><U+2581><U+2581><U+2581><U+
##	-70.1	-27.15	14.85	50.58	137	<U+2583><U+2587><U+2583><U+2585><U+2585><U+2583><U+2582><U+
##	-66.6	0	26.8	45.95	89.8	<U+2582><U+2581><U+2581><U+2587><U+2587><U+2587><U+2585><U+
##	-180	-72.62	-33.85	0	152	<U+2582><U+2582><U+2586><U+2587><U+2587><U+2582><U+2581><U+
##	-147	-91.8	-66.15	21.2	120.9	<U+2582><U+2587><U+2582><U+2582><U+2583><U+2582><U+2582><U+
##	-180	-175	-61	0	167	<U+2587><U+2581><U+2581><U+2581><U+2585><U+2581><U+2582><U+
##	-89.1	-41.98	-22.45	0	66.4	<U+2581><U+2583><U+2587><U+2585><U+2587><U+2582><U+2581><U+
##	-180	-88.4	-7.85	9.05	173	<U+2581><U+2583><U+2587><U+2583><U+2583><U+2581><U+2581><U+
##	-149.6	-59.68	-43.55	-25.2	73.2	<U+2581><U+2581><U+2582><U+2587><U+2585><U+2583><U+2581><U+
##	-72.5	-6.07	0	12.07	62.1	<U+2581><U+2581><U+2581><U+2582><U+2587><U+2583><U+2581><U+
##	-88.8	-25.9	0	11.2	88.5	<U+2581><U+2582><U+2585><U+2583><U+2587><U+2582><U+2581><U+
##	-55.8	1.76	5.28	14.9	60.3	<U+2582><U+2581><U+2581><U+2581><U+2587><U+2583><U+2581><U+
##	-149.59	-40.89	-20.96	17.5	149.4	<U+2581><U+2581><U+2586><U+2587><U+2587><U+2582><U+2581><U+
##	-72.5	0	9.24	28.4	89.8	<U+2581><U+2581><U+2581><U+2587><U+2586><U+2583><U+2582><U+
##	-180	-31.78	0	77.3	180	<U+2581><U+2581><U+2582><U+2587><U+2582><U+2585><U+2583><U+
##	-28.9	1.1	113	123	162	<U+2581><U+2587><U+2581><U+2581><U+2581><U+2581><U+2587><U+
##	-153.71	-18.49	48.17	67.61	153.55	<U+2581><U+2582><U+2581><U+2582><U+2582><U+2587><U+2583><U+
##	-180	-0.74	21.7	140	180	<U+2583><U+2581><U+2581><U+2587><U+2583><U+2582><U+2582><U+
##	0	1.64	8.13	16.33	43.41	<U+2587><U+2586><U+2583><U+2583><U+2582><U+2581><U+2581><U+
##	0	0.2	0.4	0.7	4	<U+2587><U+2582><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	3.48	8.09	19.24	82.68	<U+2587><U+2583><U+2582><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.34	5.52	12.87	47.75	<U+2587><U+2583><U+2582><U+2582><U+2581><U+2581><U+2581><U+
##	0	1.38	5.7	14.92	161.96	<U+2587><U+2582><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.2	0.4	0.7	14.2	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	4.64	12.2	26.36	123.78	<U+2587><U+2583><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.43	8.03	85.37	179.17	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	2.58	16.68	35.98	177.04	<U+2587><U+2583><U+2582><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.1	0.3	0.7	176.6	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	3.88	10.26	24.67	107.09	<U+2587><U+2582><U+2582><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.52	24.74	85.82	197.51	<U+2587><U+2583><U+2581><U+2581><U+2582><U+2581><U+2581><U+
##	0	9.03	40.61	75.62	331.7	<U+2587><U+2585><U+2582><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.38	1	3.43	230.43	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	6.76	21.16	51.24	172.61	<U+2587><U+2583><U+2582><U+2582><U+2581><U+2581><U+2581><U+
##	0	2.7	66.15	266.58	1884.56	<U+2587><U+2582><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	0	0.1	0.5	16.2	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	12.13	65.43	370.11	6836.02	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.11	30.43	165.53	2279.62	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	1.9	32.52	222.65	26232.21	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	0	0.1	0.5	200.7	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	21.52	148.95	694.65	15321.01	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.18	64.48	7289.08	32102.24	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.1	0.2	0.3	16.5	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	6.64	278.31	1294.85	31344.57	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.01	0.09	0.47	31183.24	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	15.09	105.35	608.79	11467.91	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	0	0.27	612.21	7368.41	39009.33	<U+2587><U+2581><U+2581><U+2581><U+2581><U+2581><U+2581><U+
##	-180	-43.1	0	45.88	180	<U+2581><U+2581><U+2583><U+2587><U+2583><U+2582><U+2582><U+

```
## -180      -88.3  -13      12.9   179    <U+2581><U+2583><U+2587><U+2585><U+2582><U+2581><U+2581><U+
## -150.87   -77.64  -3.32   79.64   154.95 <U+2581><U+2587><U+2583><U+2582><U+2583><U+2583><U+2583><U+
## -180      -68.6   0       110     180    <U+2583><U+2583><U+2583><U+2587><U+2581><U+2583><U+2587><U+
```

We can see that we have three types of variables: character, integer, and numeric. Usually, character variables have a low number of missing, but most of them are full with empty values. Integer variables are mostly measurements of the x,y,z axis, and numeric ones are mostly aggregations (avg, min, max etc.).

If we dig deeper, into the dataset, we can see that aggregation variables like (min, max,avg) are related to the column: new_window. Each time the new_window is 'yes', those columns contain values, otherwise they are empty. Also, when looking at the data, two values are problematic: the empty character '' and the '#DIV/0!'. We have replaced both values with 'NA' as shown in the next code chunk.

```
training <- replace(training,training==' ' || training=='#DIV/0!',NA)
```

Next thing was to delete the class column from the dataset, along with raw_timestamp_part_1 which we considered not really useful. Then we created two datasets for predictors and output as: x.t1 for predictors and y.t for outcome, as shown below.

```
x.t <- training[, -which(names(training) %in% c("classe","raw_timestamp_part_1"))]
y.t <- training[, which(names(training) == "classe")]
```

Then, because of the high number of NA's in the dataset, we removed all columns that has more than one NA in them, along with those which still has the empty character.

```
x.t1 <- x.t[,colSums(is.na(x.t))<1 ]
x.t1 <- x.t1[,colSums(x.t1=='')<1]
```

then, after checking the dataset, we've found that there are many columns with character type. That's why we have eliminated all of these.

```
x.t1 <- x.t1[, !sapply(x.t1, is.character)]
```

After several iterations training the model, we've found that the best predictors were those variables that are numeric and had 'x','y','z' at the end. That's why we decided to train the final model with only these predictors. We rejoined the outcome y.t to the set of predictors x.t1 before passing it to the train function.

```
x.t1 <- x.t1[,which((str_sub(names(x.t1),-1)%in% c('x','y','z')))]
x.t1 <- cbind(x.t1,y.t)
```

The model used was gbm, because it supports multi-class prediction and can give a very good result with a bunch of good learners.

```
modFit<-train(y.t~., method="gbm",data=x.t1,verbose=F)
```

The result modFit are below:

Stochastic Gradient Boosting

19622 samples 36 predictor 5 classes: 'A', 'B', 'C', 'D', 'E'

No pre-processing

The final values used for the model were n.trees = 150, interaction.depth = 3, shrinkage = 0.1 and n.minobsinnode = 10.

Note that this result on the training set, is very coherent with the one of the testing set:90%. This was the best model we could get.