Project

m.tagorti

8 juin 2017

This project describes the experiments that have been carried out with a group of 30 volunteers. Each person performed 6 activities (walking, walking_upstairs, walking_downstairs, sitting, standing, laying) wearing a smartphone (Samsung Galaxy S II) on the waist. Using its embedded accelerometer and gyroscope, they captured 3-axial(X,Y,Z) linear acceleration(Body_acc, Body-Gyro, Total_acc) and 3-axial angular velocity. We summarize in the following dataframe all the collected data contained in "Inertial Signals" from both the "test" and "train" files, for each subject and performed activity.

```
library(plyr)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:plyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

directory1<-"C:/Users/HP/downloads/getdata_projectfiles_UCI HAR Dataset/UCI
HAR Dataset/test/Inertial Signals"#setting directory
directory2<-"C:/Users/HP/downloads/getdata_projectfiles_UCI HAR Dataset/UCI
HAR Dataset/train/Inertial Signals"
files_list1<-list.files(directory1, full.names =TRUE)#listing files
data1<-list()#putting all files into a list
for (i in 1:length(files_list1)) {
data1[[i+2]] <- tbl_df(read.table(files_list1[i]))#loops through the files,
rbinding them together
}
data1[[2]]<-tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI
HAR Dataset/UCI HAR Dataset/test/y_test.txt"))#adding the subject column
data1[[1]]<-tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI
HAR Dataset/UCI HAR Dataset/test/subject_test.txt"))#adding the activity
column
dataframe_test<-cbind.data.frame(data1)# collecting all the datas
```

```r
#In the following we will reproduce the same steps
files_list2 <- list.files(directory2, full.names =TRUE)   #creates a list of
files
data2<-list()
for (i in 1:length(files_list2)) {
data2[[i+2]] <- tbl_df(read.table(files_list2[i]))#loops through the files,
rbinding them together
}

data2[[2]]<-tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI
HAR Dataset/UCI HAR Dataset/train/y_train.txt"))
data2[[1]]<-tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI
HAR Dataset/UCI HAR Dataset/train/subject_train.txt"))
dataframe_train<-cbind.data.frame(data2)
dataframe_final<-rbind.data.frame(dataframe_train,dataframe_test)# collecting
both train and test datas
names(dataframe_final)<-
c("subject","activity",paste(replicate(128,"Body_acc_X"),1:128, sep=""),
paste(replicate(128,"Body_acc_Y"),1:128,
sep=""),paste(replicate(128,"Body_acc_Z"),1:128, sep=""),
paste(replicate(128,"Body_gyro_X"),1:128,
sep=""),paste(replicate(128,"Body_gyro_Y"),1:128, sep=""),
paste(replicate(128,"Body_gyro_Z"),1:128,
sep=""),paste(replicate(128,"Total_acc_X"),1:128, sep=""),
paste(replicate(128,"Total_acc_Y"),1:128,
sep=""),paste(replicate(128,"Total_acc_Z"),1:128, sep=""))#naming the columns
dataframe_final<-tbl_df(dataframe_final)
activity_label<-read.table("activity_labels.txt")
dataframe_final<-mutate(dataframe_final,
activity=activity_label[activity,2])#putting labels to the columns
dataframe_final
```

```
## # A tibble: 10,299 × 1,154
##    subject activity   Body_acc_X1   Body_acc_X2   Body_acc_X3
##      <int>   <fctr>         <dbl>         <dbl>         <dbl>
## 1        1 STANDING  1.808515e-04  0.0101385600  0.0092755740
## 2        1 STANDING  1.093752e-03  0.0045500770  0.0028791730
## 3        1 STANDING  3.531266e-03  0.0022850630 -0.0004197538
## 4        1 STANDING -1.772352e-03 -0.0013114480  0.0003876795
## 5        1 STANDING  8.747685e-05 -0.0002719175  0.0010221030
## 6        1 STANDING  5.251613e-04 -0.0018632820 -0.0015102310
## 7        1 STANDING -3.654856e-03 -0.0040257420 -0.0012593590
## 8        1 STANDING  3.775996e-03  0.0042627720  0.0048963610
## 9        1 STANDING  6.200004e-03  0.0026326370  0.0013077080
## 10       1 STANDING -1.649368e-03 -0.0016844210 -0.0010734670
## # ... with 10,289 more rows, and 1149 more variables: Body_acc_X4 <dbl>,
## #   Body_acc_X5 <dbl>, Body_acc_X6 <dbl>, Body_acc_X7 <dbl>,
## #   Body_acc_X8 <dbl>, Body_acc_X9 <dbl>, Body_acc_X10 <dbl>,
## #   Body_acc_X11 <dbl>, Body_acc_X12 <dbl>, Body_acc_X13 <dbl>,
## #   Body_acc_X14 <dbl>, Body_acc_X15 <dbl>, Body_acc_X16 <dbl>,
```

```
## #   Body_acc_X17 <dbl>, Body_acc_X18 <dbl>, Body_acc_X19 <dbl>,
## #   Body_acc_X20 <dbl>, Body_acc_X21 <dbl>, Body_acc_X22 <dbl>,
## #   Body_acc_X23 <dbl>, Body_acc_X24 <dbl>, Body_acc_X25 <dbl>,
## #   Body_acc_X26 <dbl>, Body_acc_X27 <dbl>, Body_acc_X28 <dbl>,
## #   Body_acc_X29 <dbl>, Body_acc_X30 <dbl>, Body_acc_X31 <dbl>,
## #   Body_acc_X32 <dbl>, Body_acc_X33 <dbl>, Body_acc_X34 <dbl>,
## #   Body_acc_X35 <dbl>, Body_acc_X36 <dbl>, Body_acc_X37 <dbl>,
## #   Body_acc_X38 <dbl>, Body_acc_X39 <dbl>, Body_acc_X40 <dbl>,
## #   Body_acc_X41 <dbl>, Body_acc_X42 <dbl>, Body_acc_X43 <dbl>,
## #   Body_acc_X44 <dbl>, Body_acc_X45 <dbl>, Body_acc_X46 <dbl>,
## #   Body_acc_X47 <dbl>, Body_acc_X48 <dbl>, Body_acc_X49 <dbl>,
## #   Body_acc_X50 <dbl>, Body_acc_X51 <dbl>, Body_acc_X52 <dbl>,
## #   Body_acc_X53 <dbl>, Body_acc_X54 <dbl>, Body_acc_X55 <dbl>,
## #   Body_acc_X56 <dbl>, Body_acc_X57 <dbl>, Body_acc_X58 <dbl>,
## #   Body_acc_X59 <dbl>, Body_acc_X60 <dbl>, Body_acc_X61 <dbl>,
## #   Body_acc_X62 <dbl>, Body_acc_X63 <dbl>, Body_acc_X64 <dbl>,
## #   Body_acc_X65 <dbl>, Body_acc_X66 <dbl>, Body_acc_X67 <dbl>,
## #   Body_acc_X68 <dbl>, Body_acc_X69 <dbl>, Body_acc_X70 <dbl>,
## #   Body_acc_X71 <dbl>, Body_acc_X72 <dbl>, Body_acc_X73 <dbl>,
## #   Body_acc_X74 <dbl>, Body_acc_X75 <dbl>, Body_acc_X76 <dbl>,
## #   Body_acc_X77 <dbl>, Body_acc_X78 <dbl>, Body_acc_X79 <dbl>,
## #   Body_acc_X80 <dbl>, Body_acc_X81 <dbl>, Body_acc_X82 <dbl>,
## #   Body_acc_X83 <dbl>, Body_acc_X84 <dbl>, Body_acc_X85 <dbl>,
## #   Body_acc_X86 <dbl>, Body_acc_X87 <dbl>, Body_acc_X88 <dbl>,
## #   Body_acc_X89 <dbl>, Body_acc_X90 <dbl>, Body_acc_X91 <dbl>,
## #   Body_acc_X92 <dbl>, Body_acc_X93 <dbl>, Body_acc_X94 <dbl>,
## #   Body_acc_X95 <dbl>, Body_acc_X96 <dbl>, Body_acc_X97 <dbl>,
## #   Body_acc_X98 <dbl>, Body_acc_X99 <dbl>, Body_acc_X100 <dbl>,
## #   Body_acc_X101 <dbl>, Body_acc_X102 <dbl>, Body_acc_X103 <dbl>, ...
```

The variable "Body_acc_gravityXYZ"" is obtained by substracting the variable "Body_accXYZ"" from the variable "Total_accXYZ". The variable Body_gyroXYZ is the measurement we get from the gyroscope. By manipulating these three variables ("Body_acc_gravityXYZ"", "BodygyroXYZ", "Body_accXYZ"") with the angular velocity signals, we extract the set of features described in the file "features.txt". For instance, the feature "tBody-acc_(mean)-XYZ"" is obtained by computing the mean over all the 128 variables "(tBody-acc-(mean)-XYZ)"" shown in the dataframe above.

Part2

In this part we will try to collect all the features data from both the train and test sets. Then we will attribute to each suject for each activity the corresponding features. We will reproduce the same kind of code steps that we have performed in part 1.

```
X_test<-tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI HAR
Dataset/UCI HAR Dataset/test/X_test.txt"))
X_test_subject<-
tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI HAR
Dataset/UCI HAR Dataset/test/subject_test.txt"))
X_test_activity<-
```

```
tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI HAR
Dataset/UCI HAR Dataset/test/y_test.txt"))
X_test_final<-cbind.data.frame(X_test_subject,X_test_activity,X_test)
X_train<-tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI
HAR Dataset/UCI HAR Dataset/train/X_train.txt",header=FALSE))
X_train_subject<-
tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI HAR
Dataset/UCI HAR Dataset/train/subject_train.txt",header=FALSE))
X_train_activity<-
tbl_df(read.table("C:/Users/HP/downloads/getdata_projectfiles_UCI HAR
Dataset/UCI HAR Dataset/train/y_train.txt",header=FALSE))
X_train_final<-cbind.data.frame(X_train_subject,X_train_activity,X_train)
X_subject_activity<-rbind.data.frame(X_train_final,X_test_final)
str(X_subject_activity)
```

```
## 'data.frame':    10299 obs. of  563 variables:
##  $ V1  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ V1  : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ V1  : num  0.289 0.278 0.28 0.279 0.277 ...
##  $ V2  : num  -0.0203 -0.0164 -0.0195 -0.0262 -0.0166 ...
##  $ V3  : num  -0.133 -0.124 -0.113 -0.123 -0.115 ...
##  $ V4  : num  -0.995 -0.998 -0.995 -0.996 -0.998 ...
##  $ V5  : num  -0.983 -0.975 -0.967 -0.983 -0.981 ...
##  $ V6  : num  -0.914 -0.96 -0.979 -0.991 -0.99 ...
##  $ V7  : num  -0.995 -0.999 -0.997 -0.997 -0.998 ...
##  $ V8  : num  -0.983 -0.975 -0.964 -0.983 -0.98 ...
##  $ V9  : num  -0.924 -0.958 -0.977 -0.989 -0.99 ...
##  $ V10 : num  -0.935 -0.943 -0.939 -0.939 -0.942 ...
##  $ V11 : num  -0.567 -0.558 -0.558 -0.576 -0.569 ...
##  $ V12 : num  -0.744 -0.818 -0.818 -0.83 -0.825 ...
##  $ V13 : num  0.853 0.849 0.844 0.844 0.849 ...
##  $ V14 : num  0.686 0.686 0.682 0.682 0.683 ...
##  $ V15 : num  0.814 0.823 0.839 0.838 0.838 ...
##  $ V16 : num  -0.966 -0.982 -0.983 -0.986 -0.993 ...
##  $ V17 : num  -1 -1 -1 -1 -1 ...
##  $ V18 : num  -1 -1 -1 -1 -1 ...
##  $ V19 : num  -0.995 -0.998 -0.999 -1 -1 ...
##  $ V20 : num  -0.994 -0.999 -0.997 -0.997 -0.998 ...
##  $ V21 : num  -0.988 -0.978 -0.965 -0.984 -0.981 ...
##  $ V22 : num  -0.943 -0.948 -0.975 -0.986 -0.991 ...
##  $ V23 : num  -0.408 -0.715 -0.592 -0.627 -0.787 ...
##  $ V24 : num  -0.679 -0.501 -0.486 -0.851 -0.559 ...
##  $ V25 : num  -0.602 -0.571 -0.571 -0.912 -0.761 ...
##  $ V26 : num  0.9293 0.6116 0.273 0.0614 0.3133 ...
##  $ V27 : num  -0.853 -0.3295 -0.0863 0.0748 -0.1312 ...
##  $ V28 : num  0.36 0.284 0.337 0.198 0.191 ...
##  $ V29 : num  -0.0585 0.2846 -0.1647 -0.2643 0.0869 ...
##  $ V30 : num  0.2569 0.1157 0.0172 0.0725 0.2576 ...
##  $ V31 : num  -0.2248 -0.091 -0.0745 -0.1553 -0.2725 ...
##  $ V32 : num  0.264 0.294 0.342 0.323 0.435 ...
```

```
##  $ V33 : num  -0.0952 -0.2812 -0.3326 -0.1708 -0.3154 ...
##  $ V34 : num  0.279 0.086 0.239 0.295 0.44 ...
##  $ V35 : num  -0.4651 -0.0222 -0.1362 -0.3061 -0.2691 ...
##  $ V36 : num  0.4919 -0.0167 0.1739 0.4821 0.1794 ...
##  $ V37 : num  -0.191 -0.221 -0.299 -0.47 -0.089 ...
##  $ V38 : num  0.3763 -0.0134 -0.1247 -0.3057 -0.1558 ...
##  $ V39 : num  0.4351 -0.0727 -0.1811 -0.3627 -0.1898 ...
##  $ V40 : num  0.661 0.579 0.609 0.507 0.599 ...
##  $ V41 : num  0.963 0.967 0.967 0.968 0.968 ...
##  $ V42 : num  -0.141 -0.142 -0.142 -0.144 -0.149 ...
##  $ V43 : num  0.1154 0.1094 0.1019 0.0999 0.0945 ...
##  $ V44 : num  -0.985 -0.997 -1 -0.997 -0.998 ...
##  $ V45 : num  -0.982 -0.989 -0.993 -0.981 -0.988 ...
##  $ V46 : num  -0.878 -0.932 -0.993 -0.978 -0.979 ...
##  $ V47 : num  -0.985 -0.998 -1 -0.996 -0.998 ...
##  $ V48 : num  -0.984 -0.99 -0.993 -0.981 -0.989 ...
##  $ V49 : num  -0.895 -0.933 -0.993 -0.978 -0.979 ...
##  $ V50 : num  0.892 0.892 0.892 0.894 0.894 ...
##  $ V51 : num  -0.161 -0.161 -0.164 -0.164 -0.167 ...
##  $ V52 : num  0.1247 0.1226 0.0946 0.0934 0.0917 ...
##  $ V53 : num  0.977 0.985 0.987 0.987 0.987 ...
##  $ V54 : num  -0.123 -0.115 -0.115 -0.121 -0.122 ...
##  $ V55 : num  0.0565 0.1028 0.1028 0.0958 0.0941 ...
##  $ V56 : num  -0.375 -0.383 -0.402 -0.4 -0.4 ...
##  $ V57 : num  0.899 0.908 0.909 0.911 0.912 ...
##  $ V58 : num  -0.971 -0.971 -0.97 -0.969 -0.967 ...
##  $ V59 : num  -0.976 -0.979 -0.982 -0.982 -0.984 ...
##  $ V60 : num  -0.984 -0.999 -1 -0.996 -0.998 ...
##  $ V61 : num  -0.989 -0.99 -0.992 -0.981 -0.991 ...
##  $ V62 : num  -0.918 -0.942 -0.993 -0.98 -0.98 ...
##  $ V63 : num  -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 ...
##  $ V64 : num  -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 ...
##  $ V65 : num  0.114 -0.21 -0.927 -0.596 -0.617 ...
##  $ V66 : num  -0.59042 -0.41006 0.00223 -0.06493 -0.25727 ...
##  $ V67 : num  0.5911 0.4139 0.0275 0.0754 0.2689 ...
##  $ V68 : num  -0.5918 -0.4176 -0.0567 -0.0858 -0.2807 ...
##  $ V69 : num  0.5925 0.4213 0.0855 0.0962 0.2926 ...
##  $ V70 : num  -0.745 -0.196 -0.329 -0.295 -0.167 ...
##  $ V71 : num  0.7209 0.1253 0.2705 0.2283 0.0899 ...
##  $ V72 : num  -0.7124 -0.1056 -0.2545 -0.2063 -0.0663 ...
##  $ V73 : num  0.7113 0.1091 0.2576 0.2048 0.0671 ...
##  $ V74 : num  -0.995 -0.834 -0.705 -0.385 -0.237 ...
##  $ V75 : num  0.996 0.834 0.714 0.386 0.239 ...
##  $ V76 : num  -0.996 -0.834 -0.723 -0.387 -0.241 ...
##  $ V77 : num  0.992 0.83 0.729 0.385 0.241 ...
##  $ V78 : num  0.57 -0.831 -0.181 -0.991 -0.408 ...
##  $ V79 : num  0.439 -0.866 0.338 -0.969 -0.185 ...
##  $ V80 : num  0.987 0.974 0.643 0.984 0.965 ...
##  $ V81 : num  0.078 0.074 0.0736 0.0773 0.0734 ...
##  $ V82 : num  0.005 0.00577 0.0031 0.02006 0.01912 ...
```

```
##  $ V83 : num  -0.06783 0.02938 -0.00905 -0.00986 0.01678 ...
##  $ V84 : num  -0.994 -0.996 -0.991 -0.993 -0.996 ...
##  $ V85 : num  -0.988 -0.981 -0.981 -0.988 -0.988 ...
##  $ V86 : num  -0.994 -0.992 -0.99 -0.993 -0.992 ...
##  $ V87 : num  -0.994 -0.996 -0.991 -0.994 -0.997 ...
##  $ V88 : num  -0.986 -0.979 -0.979 -0.986 -0.987 ...
##  $ V89 : num  -0.993 -0.991 -0.987 -0.991 -0.991 ...
##  $ V90 : num  -0.985 -0.995 -0.987 -0.987 -0.997 ...
##  $ V91 : num  -0.992 -0.979 -0.979 -0.992 -0.992 ...
##  $ V92 : num  -0.993 -0.992 -0.992 -0.99 -0.99 ...
##  $ V93 : num  0.99 0.993 0.988 0.988 0.994 ...
##  $ V94 : num  0.992 0.992 0.992 0.993 0.993 ...
##  $ V95 : num  0.991 0.989 0.989 0.993 0.986 ...
##  $ V96 : num  -0.994 -0.991 -0.988 -0.993 -0.994 ...
##  $ V97 : num  -1 -1 -1 -1 -1 ...
##   [list output truncated]
```

Part3

In this section we will attribute to our dataframe its column names which are "subject", "activity" and the features labels in "features.txt". Then we extract the features that compute the mean and the std of the measurments. We do this, by looking for the terms "mean", "Mean" and "std" in the features labels'.

```
features<-read.table("features.txt",sep=" ",header=FALSE)#reading the
features labels
features<-paste(features[,2], features[,1], sep="")
names(X_subject_activity)<-
c("subject","activity",as.character(unique(features)))# attributing
columnnames to the dataframe
X_subject_activity<-tbl_df(X_subject_activity)
X_subject_activity<-select(X_subject_activity,subject,activity,grep("mean",
names(X_subject_activity)),grep("Mean",names(X_subject_activity)),grep("std",
names(X_subject_activity)))#looking for the key terms to extract the data
X_subject_activity
```

```
## # A tibble: 10,299 × 88
##    subject activity `tBodyAcc-mean()-X1` `tBodyAcc-mean()-Y2`
##      <int>    <int>                <dbl>                <dbl>
## 1        1        5            0.2885845        -0.020294171
## 2        1        5            0.2784188        -0.016410568
## 3        1        5            0.2796531        -0.019467156
## 4        1        5            0.2791739        -0.026200646
## 5        1        5            0.2766288        -0.016569655
## 6        1        5            0.2771988        -0.010097850
## 7        1        5            0.2794539        -0.019640776
## 8        1        5            0.2774325        -0.030488303
## 9        1        5            0.2772934        -0.021750698
## 10       1        5            0.2805857        -0.009960298
## # ... with 10,289 more rows, and 84 more variables:
```

```
## #   `tBodyAcc-mean()-Z3` <dbl>, `tGravityAcc-mean()-X41` <dbl>,
## #   `tGravityAcc-mean()-Y42` <dbl>, `tGravityAcc-mean()-Z43` <dbl>,
## #   `tBodyAccJerk-mean()-X81` <dbl>, `tBodyAccJerk-mean()-Y82` <dbl>,
## #   `tBodyAccJerk-mean()-Z83` <dbl>, `tBodyGyro-mean()-X121` <dbl>,
## #   `tBodyGyro-mean()-Y122` <dbl>, `tBodyGyro-mean()-Z123` <dbl>,
## #   `tBodyGyroJerk-mean()-X161` <dbl>, `tBodyGyroJerk-mean()-Y162` <dbl>,
## #   `tBodyGyroJerk-mean()-Z163` <dbl>, `tBodyAccMag-mean()201` <dbl>,
## #   `tGravityAccMag-mean()214` <dbl>, `tBodyAccJerkMag-mean()227` <dbl>,
## #   `tBodyGyroMag-mean()240` <dbl>, `tBodyGyroJerkMag-mean()253` <dbl>,
## #   `fBodyAcc-mean()-X266` <dbl>, `fBodyAcc-mean()-Y267` <dbl>,
## #   `fBodyAcc-mean()-Z268` <dbl>, `fBodyAcc-meanFreq()-X294` <dbl>,
## #   `fBodyAcc-meanFreq()-Y295` <dbl>, `fBodyAcc-meanFreq()-Z296` <dbl>,
## #   `fBodyAccJerk-mean()-X345` <dbl>, `fBodyAccJerk-mean()-Y346` <dbl>,
## #   `fBodyAccJerk-mean()-Z347` <dbl>,
## #   `fBodyAccJerk-meanFreq()-X373` <dbl>,
## #   `fBodyAccJerk-meanFreq()-Y374` <dbl>,
## #   `fBodyAccJerk-meanFreq()-Z375` <dbl>, `fBodyGyro-mean()-X424` <dbl>,
## #   `fBodyGyro-mean()-Y425` <dbl>, `fBodyGyro-mean()-Z426` <dbl>,
## #   `fBodyGyro-meanFreq()-X452` <dbl>, `fBodyGyro-meanFreq()-Y453` <dbl>,
## #   `fBodyGyro-meanFreq()-Z454` <dbl>, `fBodyAccMag-mean()503` <dbl>,
## #   `fBodyAccMag-meanFreq()513` <dbl>,
## #   `fBodyBodyAccJerkMag-mean()516` <dbl>,
## #   `fBodyBodyAccJerkMag-meanFreq()526` <dbl>,
## #   `fBodyBodyGyroMag-mean()529` <dbl>,
## #   `fBodyBodyGyroMag-meanFreq()539` <dbl>,
## #   `fBodyBodyGyroJerkMag-mean()542` <dbl>,
## #   `fBodyBodyGyroJerkMag-meanFreq()552` <dbl>,
## #   `angle(tBodyAccMean,gravity)555` <dbl>,
## #   `angle(tBodyAccJerkMean),gravityMean)556` <dbl>,
## #   `angle(tBodyGyroMean,gravityMean)557` <dbl>,
## #   `angle(tBodyGyroJerkMean,gravityMean)558` <dbl>,
## #   `angle(X,gravityMean)559` <dbl>, `angle(Y,gravityMean)560` <dbl>,
## #   `angle(Z,gravityMean)561` <dbl>, `tBodyAcc-std()-X4` <dbl>,
## #   `tBodyAcc-std()-Y5` <dbl>, `tBodyAcc-std()-Z6` <dbl>,
## #   `tGravityAcc-std()-X44` <dbl>, `tGravityAcc-std()-Y45` <dbl>,
## #   `tGravityAcc-std()-Z46` <dbl>, `tBodyAccJerk-std()-X84` <dbl>,
## #   `tBodyAccJerk-std()-Y85` <dbl>, `tBodyAccJerk-std()-Z86` <dbl>,
## #   `tBodyGyro-std()-X124` <dbl>, `tBodyGyro-std()-Y125` <dbl>,
## #   `tBodyGyro-std()-Z126` <dbl>, `tBodyGyroJerk-std()-X164` <dbl>,
## #   `tBodyGyroJerk-std()-Y165` <dbl>, `tBodyGyroJerk-std()-Z166` <dbl>,
## #   `tBodyAccMag-std()202` <dbl>, `tGravityAccMag-std()215` <dbl>,
## #   `tBodyAccJerkMag-std()228` <dbl>, `tBodyGyroMag-std()241` <dbl>,
## #   `tBodyGyroJerkMag-std()254` <dbl>, `fBodyAcc-std()-X269` <dbl>,
## #   `fBodyAcc-std()-Y270` <dbl>, `fBodyAcc-std()-Z271` <dbl>,
## #   `fBodyAccJerk-std()-X348` <dbl>, `fBodyAccJerk-std()-Y349` <dbl>,
## #   `fBodyAccJerk-std()-Z350` <dbl>, `fBodyGyro-std()-X427` <dbl>,
## #   `fBodyGyro-std()-Y428` <dbl>, `fBodyGyro-std()-Z429` <dbl>,
## #   `fBodyAccMag-std()504` <dbl>, `fBodyBodyAccJerkMag-std()517` <dbl>,
## #   `fBodyBodyGyroMag-std()530` <dbl>,
## #   `fBodyBodyGyroJerkMag-std()543` <dbl>
```

We get a 10299*88 dataframe with its attributed names, after having extracted the appropriate features.

Part4

In this part we will match each activity symbols (1,2,3,4,5,6) with its labels in the file "activity_labels.txt".

```
activity_label<-read.table("activity_labels.txt")
X_subject_activity<-mutate(X_subject_activity,
activity=activity_label[activity,2])
X_subject_activity
```

```
## # A tibble: 10,299 × 88
##    subject activity `tBodyAcc-mean()-X1` `tBodyAcc-mean()-Y2`
##      <int>   <fctr>                <dbl>                <dbl>
## 1        1 STANDING            0.2885845         -0.020294171
## 2        1 STANDING            0.2784188         -0.016410568
## 3        1 STANDING            0.2796531         -0.019467156
## 4        1 STANDING            0.2791739         -0.026200646
## 5        1 STANDING            0.2766288         -0.016569655
## 6        1 STANDING            0.2771988         -0.010097850
## 7        1 STANDING            0.2794539         -0.019640776
## 8        1 STANDING            0.2774325         -0.030488303
## 9        1 STANDING            0.2772934         -0.021750698
## 10       1 STANDING            0.2805857         -0.009960298
## # ... with 10,289 more rows, and 84 more variables:
## #   `tBodyAcc-mean()-Z3` <dbl>, `tGravityAcc-mean()-X41` <dbl>,
## #   `tGravityAcc-mean()-Y42` <dbl>, `tGravityAcc-mean()-Z43` <dbl>,
## #   `tBodyAccJerk-mean()-X81` <dbl>, `tBodyAccJerk-mean()-Y82` <dbl>,
## #   `tBodyAccJerk-mean()-Z83` <dbl>, `tBodyGyro-mean()-X121` <dbl>,
## #   `tBodyGyro-mean()-Y122` <dbl>, `tBodyGyro-mean()-Z123` <dbl>,
## #   `tBodyGyroJerk-mean()-X161` <dbl>, `tBodyGyroJerk-mean()-Y162` <dbl>,
## #   `tBodyGyroJerk-mean()-Z163` <dbl>, `tBodyAccMag-mean()201` <dbl>,
## #   `tGravityAccMag-mean()214` <dbl>, `tBodyAccJerkMag-mean()227` <dbl>,
## #   `tBodyGyroMag-mean()240` <dbl>, `tBodyGyroJerkMag-mean()253` <dbl>,
## #   `fBodyAcc-mean()-X266` <dbl>, `fBodyAcc-mean()-Y267` <dbl>,
## #   `fBodyAcc-mean()-Z268` <dbl>, `fBodyAcc-meanFreq()-X294` <dbl>,
## #   `fBodyAcc-meanFreq()-Y295` <dbl>, `fBodyAcc-meanFreq()-Z296` <dbl>,
## #   `fBodyAccJerk-mean()-X345` <dbl>, `fBodyAccJerk-mean()-Y346` <dbl>,
## #   `fBodyAccJerk-mean()-Z347` <dbl>,
## #   `fBodyAccJerk-meanFreq()-X373` <dbl>,
## #   `fBodyAccJerk-meanFreq()-Y374` <dbl>,
## #   `fBodyAccJerk-meanFreq()-Z375` <dbl>, `fBodyGyro-mean()-X424` <dbl>,
## #   `fBodyGyro-mean()-Y425` <dbl>, `fBodyGyro-mean()-Z426` <dbl>,
## #   `fBodyGyro-meanFreq()-X452` <dbl>, `fBodyGyro-meanFreq()-Y453` <dbl>,
## #   `fBodyGyro-meanFreq()-Z454` <dbl>, `fBodyAccMag-mean()503` <dbl>,
## #   `fBodyAccMag-meanFreq()513` <dbl>,
## #   `fBodyBodyAccJerkMag-mean()516` <dbl>,
## #   `fBodyBodyAccJerkMag-meanFreq()526` <dbl>,
```

```
## #   `fBodyBodyGyroMag-mean()529` <dbl>,
## #   `fBodyBodyGyroMag-meanFreq()539` <dbl>,
## #   `fBodyBodyGyroJerkMag-mean()542` <dbl>,
## #   `fBodyBodyGyroJerkMag-meanFreq()552` <dbl>,
## #   `angle(tBodyAccMean,gravity)555` <dbl>,
## #   `angle(tBodyAccJerkMean),gravityMean)556` <dbl>,
## #   `angle(tBodyGyroMean,gravityMean)557` <dbl>,
## #   `angle(tBodyGyroJerkMean,gravityMean)558` <dbl>,
## #   `angle(X,gravityMean)559` <dbl>, `angle(Y,gravityMean)560` <dbl>,
## #   `angle(Z,gravityMean)561` <dbl>, `tBodyAcc-std()-X4` <dbl>,
## #   `tBodyAcc-std()-Y5` <dbl>, `tBodyAcc-std()-Z6` <dbl>,
## #   `tGravityAcc-std()-X44` <dbl>, `tGravityAcc-std()-Y45` <dbl>,
## #   `tGravityAcc-std()-Z46` <dbl>, `tBodyAccJerk-std()-X84` <dbl>,
## #   `tBodyAccJerk-std()-Y85` <dbl>, `tBodyAccJerk-std()-Z86` <dbl>,
## #   `tBodyGyro-std()-X124` <dbl>, `tBodyGyro-std()-Y125` <dbl>,
## #   `tBodyGyro-std()-Z126` <dbl>, `tBodyGyroJerk-std()-X164` <dbl>,
## #   `tBodyGyroJerk-std()-Y165` <dbl>, `tBodyGyroJerk-std()-Z166` <dbl>,
## #   `tBodyAccMag-std()202` <dbl>, `tGravityAccMag-std()215` <dbl>,
## #   `tBodyAccJerkMag-std()228` <dbl>, `tBodyGyroMag-std()241` <dbl>,
## #   `tBodyGyroJerkMag-std()254` <dbl>, `fBodyAcc-std()-X269` <dbl>,
## #   `fBodyAcc-std()-Y270` <dbl>, `fBodyAcc-std()-Z271` <dbl>,
## #   `fBodyAccJerk-std()-X348` <dbl>, `fBodyAccJerk-std()-Y349` <dbl>,
## #   `fBodyAccJerk-std()-Z350` <dbl>, `fBodyGyro-std()-X427` <dbl>,
## #   `fBodyGyro-std()-Y428` <dbl>, `fBodyGyro-std()-Z429` <dbl>,
## #   `fBodyAccMag-std()504` <dbl>, `fBodyBodyAccJerkMag-std()517` <dbl>,
## #   `fBodyBodyGyroMag-std()530` <dbl>,
## #   `fBodyBodyGyroJerkMag-std()543` <dbl>
```

We can see the term "standing" that appears instead of the symbol "5" shown in the previous dataframe.

Part5

In this last part we will proceed to the final step of the process of cleaning data. That is we will try to match for each subject belonging to the 30 volunteers and each activity in the six possible activities, the mean of the extracted features. We expect to get a dataframe with 180 rows and 88 columns.

```
Gr_by<-group_by(X_subject_activity,subject,activity)#grouping the data by
subject and activity
List<-list()
for (i in 3:88){
name<-names(Gr_by)[[i]]#name is the true name of the dataframe column
names(Gr_by)[[i]]<-"h"  # h is a temporary variable that we use only in the
#function summarize
List[[i-2]]<-as.data.frame(summarize(Gr_by, mean(h)))
names(Gr_by)[[i]]<-name#we reattribute to the dataframe its true name
names(List[[i-2]])<-c("subject","activity",names(Gr_by[,i]))}
tidy_data<-join_all(List)#we join all the dataframes by subject and activity
```

```
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
```

```
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
## Joining by: subject, activity
```

```
tbl_df(tidy_data)
```

```
## # A tibble: 180 × 88
##    subject           activity `tBodyAcc-mean()-X1` `tBodyAcc-mean()-Y2`
##      <int>             <fctr>                <dbl>                <dbl>
## 1        1             LAYING            0.2215982         -0.040513953
## 2        1            SITTING            0.2612376         -0.001308288
## 3        1           STANDING            0.2789176         -0.016137590
## 4        1            WALKING            0.2773308         -0.017383819
## 5        1 WALKING_DOWNSTAIRS            0.2891883         -0.009918505
## 6        1   WALKING_UPSTAIRS            0.2554617         -0.023953149
## 7        2             LAYING            0.2813734         -0.018158740
## 8        2            SITTING            0.2770874         -0.015687994
## 9        2           STANDING            0.2779115         -0.018420827
```

```
## 10       2            WALKING           0.2764266         -0.018594920
## # ... with 170 more rows, and 84 more variables:
## #   `tBodyAcc-mean()-Z3` <dbl>, `tGravityAcc-mean()-X41` <dbl>,
## #   `tGravityAcc-mean()-Y42` <dbl>, `tGravityAcc-mean()-Z43` <dbl>,
## #   `tBodyAccJerk-mean()-X81` <dbl>, `tBodyAccJerk-mean()-Y82` <dbl>,
## #   `tBodyAccJerk-mean()-Z83` <dbl>, `tBodyGyro-mean()-X121` <dbl>,
## #   `tBodyGyro-mean()-Y122` <dbl>, `tBodyGyro-mean()-Z123` <dbl>,
## #   `tBodyGyroJerk-mean()-X161` <dbl>, `tBodyGyroJerk-mean()-Y162` <dbl>,
## #   `tBodyGyroJerk-mean()-Z163` <dbl>, `tBodyAccMag-mean()201` <dbl>,
## #   `tGravityAccMag-mean()214` <dbl>, `tBodyAccJerkMag-mean()227` <dbl>,
## #   `tBodyGyroMag-mean()240` <dbl>, `tBodyGyroJerkMag-mean()253` <dbl>,
## #   `fBodyAcc-mean()-X266` <dbl>, `fBodyAcc-mean()-Y267` <dbl>,
## #   `fBodyAcc-mean()-Z268` <dbl>, `fBodyAcc-meanFreq()-X294` <dbl>,
## #   `fBodyAcc-meanFreq()-Y295` <dbl>, `fBodyAcc-meanFreq()-Z296` <dbl>,
## #   `fBodyAccJerk-mean()-X345` <dbl>, `fBodyAccJerk-mean()-Y346` <dbl>,
## #   `fBodyAccJerk-mean()-Z347` <dbl>,
## #   `fBodyAccJerk-meanFreq()-X373` <dbl>,
## #   `fBodyAccJerk-meanFreq()-Y374` <dbl>,
## #   `fBodyAccJerk-meanFreq()-Z375` <dbl>, `fBodyGyro-mean()-X424` <dbl>,
## #   `fBodyGyro-mean()-Y425` <dbl>, `fBodyGyro-mean()-Z426` <dbl>,
## #   `fBodyGyro-meanFreq()-X452` <dbl>, `fBodyGyro-meanFreq()-Y453` <dbl>,
## #   `fBodyGyro-meanFreq()-Z454` <dbl>, `fBodyAccMag-mean()503` <dbl>,
## #   `fBodyAccMag-meanFreq()513` <dbl>,
## #   `fBodyBodyAccJerkMag-mean()516` <dbl>,
## #   `fBodyBodyAccJerkMag-meanFreq()526` <dbl>,
## #   `fBodyBodyGyroMag-mean()529` <dbl>,
## #   `fBodyBodyGyroMag-meanFreq()539` <dbl>,
## #   `fBodyBodyGyroJerkMag-mean()542` <dbl>,
## #   `fBodyBodyGyroJerkMag-meanFreq()552` <dbl>,
## #   `angle(tBodyAccMean,gravity)555` <dbl>,
## #   `angle(tBodyAccJerkMean),gravityMean)556` <dbl>,
## #   `angle(tBodyGyroMean,gravityMean)557` <dbl>,
## #   `angle(tBodyGyroJerkMean,gravityMean)558` <dbl>,
## #   `angle(X,gravityMean)559` <dbl>, `angle(Y,gravityMean)560` <dbl>,
## #   `angle(Z,gravityMean)561` <dbl>, `tBodyAcc-std()-X4` <dbl>,
## #   `tBodyAcc-std()-Y5` <dbl>, `tBodyAcc-std()-Z6` <dbl>,
## #   `tGravityAcc-std()-X44` <dbl>, `tGravityAcc-std()-Y45` <dbl>,
## #   `tGravityAcc-std()-Z46` <dbl>, `tBodyAccJerk-std()-X84` <dbl>,
## #   `tBodyAccJerk-std()-Y85` <dbl>, `tBodyAccJerk-std()-Z86` <dbl>,
## #   `tBodyGyro-std()-X124` <dbl>, `tBodyGyro-std()-Y125` <dbl>,
## #   `tBodyGyro-std()-Z126` <dbl>, `tBodyGyroJerk-std()-X164` <dbl>,
## #   `tBodyGyroJerk-std()-Y165` <dbl>, `tBodyGyroJerk-std()-Z166` <dbl>,
## #   `tBodyAccMag-std()202` <dbl>, `tGravityAccMag-std()215` <dbl>,
## #   `tBodyAccJerkMag-std()228` <dbl>, `tBodyGyroMag-std()241` <dbl>,
## #   `tBodyGyroJerkMag-std()254` <dbl>, `fBodyAcc-std()-X269` <dbl>,
## #   `fBodyAcc-std()-Y270` <dbl>, `fBodyAcc-std()-Z271` <dbl>,
## #   `fBodyAccJerk-std()-X348` <dbl>, `fBodyAccJerk-std()-Y349` <dbl>,
## #   `fBodyAccJerk-std()-Z350` <dbl>, `fBodyGyro-std()-X427` <dbl>,
## #   `fBodyGyro-std()-Y428` <dbl>, `fBodyGyro-std()-Z429` <dbl>,
## #   `fBodyAccMag-std()504` <dbl>, `fBodyBodyAccJerkMag-std()517` <dbl>,
```

```
## #   `fBodyBodyGyroMag-std()530` <dbl>,
## #   `fBodyBodyGyroJerkMag-std()543` <dbl>
```