The codebook to accompany “MyTidyData.txt” dataset produced by “run_analysis.R” script.

This Data Set is produced from an original Data Set “Human Activity Recognition Using Smartphones Data Set [1].”

Feature Selection
==========
The features in this dataset are selected 79 features from the original “Human Activity Recognition Using Smartphones Data Set”. Only the features describing means and standard deviation were selected.

The current data set contains 180 observations (30 subjects, each performing 6 activities) and 81 variables.
To obtain the values for observations, the means of original observations for each subject and each activity across 79 features were calculated.

Note that the features in the original dataset were normalized and bounded within [-1,1], therefore the values for observation in current datasets are also bounded within [-1,1].

The current data set was produced in accordance with “tidy data” principles [2], in particular: (i) each variable forms a column, (ii) each observation forms a row.

The variables names from the original dataset were modified for current dataset to produce descriptive names containing only legal characters which can be read by R.

The variables in the original data set came from the measurements of accelerometer and gyroscope 3-axial raw signals tAcc-XYZ and tGyro-XYZ signals on Samsung smartphones [3].
These signals were used to estimate variables of the feature vector for each pattern:
'-XYZ' is used to denote 3-axial signals in the X, Y and Z directions:

tBodyAcc-XYZ
tGravityAcc-XYZ
tBodyAccJerk-XYZ
tBodyGyro-XYZ
tBodyGyroJerk-XYZ
tBodyAccMag
tGravityAccMag
tBodyAccJerkMag
tBodyGyroMag
tBodyGyroJerkMag
fBodyAcc-XYZ
fBodyAccJerk-XYZ
fBodyGyro-XYZ
fBodyAccMag
fBodyAccJerkMag
fBodyGyroMag
fBodyGyroJerkMag

The set of variables that were estimated from these signals and which were selected for current dataset are:

mean(): Mean value
std(): Standard deviation
meanFreq(): Weighted average of the frequency components to obtain a mean frequency

The full variable names are produced by adding suffixes ‘mean’, ‘std’, and ‘meanFreq’ to each signal measurement name.

In the descriptive names of current dataset the following abbreviations are used:
t – time, f – frequency, Acc – accelerometer measurement, Gyro – gyroscope measurement, Mag – magnitude.

THE VARIABLES IN CURRENT DATASET ARE:

[1] "subject", 30 possible values ranging from 1 to 30,
where each number represent a single subject.

[2] "activity", represent 6 activities performed, with 6 possible values:
walking
walkingUpstairs
walkingDownstairs
sitting
standing
laying

The following 79 variables represent averages for each subject and each activity corresponding to normalized measurements as described above.
Numerical values in range [-1, 1].

[3] "tBodyAccmeanX"
[4] "tBodyAccmeanY"
[5] "tBodyAccmeanZ"
[6] "tBodyAccstdX"
[7] "tBodyAccstdY"
[8] "tBodyAccstdZ"
[9] "tGravityAccmeanX"
[10] "tGravityAccmeanY"
[11] "tGravityAccmeanZ"
[12] "tGravityAccstdX"
[13] "tGravityAccstdY"
[14] "tGravityAccstdZ"
[15] "tBodyAccJerkmeanX"
[16] "tBodyAccJerkmeanY"
[17] "tBodyAccJerkmeanZ"
[18] "tBodyAccJerkstdX"
[19] "tBodyAccJerkstdY"
[20] "tBodyAccJerkstdZ"
[21] "tBodyGyromeanX"
[22] "tBodyGyromeanY"
[23] "tBodyGyromeanZ"
[24] "tBodyGyrostdX"
[25] "tBodyGyrostdY"
[26] "tBodyGyrostdZ"
[27] "tBodyGyroJerkmeanX"
[28] "tBodyGyroJerkmeanY"
[29] "tBodyGyroJerkmeanZ"
[30] "tBodyGyroJerkstdX"
[31] "tBodyGyroJerkstdY"
[32] "tBodyGyroJerkstdZ"
[33] "tBodyAccMagmean"
[34] "tBodyAccMagstd"

```
[35] "tGravityAccMagmean"
[36] "tGravityAccMagstd"
[37] "tBodyAccJerkMagmean"
[38] "tBodyAccJerkMagstd"
[39] "tBodyGyroMagmean"
[40] "tBodyGyroMagstd"
[41] "tBodyGyroJerkMagmean"
[42] "tBodyGyroJerkMagstd"
[43] "fBodyAccmeanX"
[44] "fBodyAccmeanY"
[45] "fBodyAccmeanZ"
[46] "fBodyAccstdX"
[47] "fBodyAccstdY"
[48] "fBodyAccstdZ"
[49] "fBodyAccmeanFreqX"
[50] "fBodyAccmeanFreqY"
[51] "fBodyAccmeanFreqZ"
[52] "fBodyAccJerkmeanX"
[53] "fBodyAccJerkmeanY"
[54] "fBodyAccJerkmeanZ"
[55] "fBodyAccJerkstdX"
[56] "fBodyAccJerkstdY"
[57] "fBodyAccJerkstdZ"
[58] "fBodyAccJerkmeanFreqX"
[59] "fBodyAccJerkmeanFreqY"
[60] "fBodyAccJerkmeanFreqZ"
[61] "fBodyGyromeanX"
[62] "fBodyGyromeanY"
[63] "fBodyGyromeanZ"
[64] "fBodyGyrostdX"
[65] "fBodyGyrostdY"
[66] "fBodyGyrostdZ"
[67] "fBodyGyromeanFreqX"
[68] "fBodyGyromeanFreqY"
[69] "fBodyGyromeanFreqZ"
[70] "fBodyAccMagmean"
[71] "fBodyAccMagstd"
[72] "fBodyAccMagmeanFreq"
[73] "fBodyAccJerkMagmean"
[74] "fBodyAccJerkMagstd"
[75] "fBodyAccJerkMagmeanFreq"
[76] "fBodyGyroMagmean"
[77] "fBodyGyroMagstd"
[78] "fBodyGyroMagmeanFreq"
[79] "fBodyGyroJerkMagmean"
[80] "fBodyGyroJerkMagstd"
[81] "fBodyGyroJerkMagmeanFreq"
```

References:

(1) Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine. International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012.
(2) Tidy Data" by Hadley Wickham, http://vita.had.co.nz/papers/tidy-data.pdf
(3) http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones

The example of the obtained tidy data set is attached on the next page:

Here is a partial sample from “MyTidyData” tidy data set:

subject	activity	tBodyAccmeanX	tBodyAccmeanY	tBodyAccmeanZ	tBodyAccstdX	tBodyAccstdY	tBodyAccstdZ
1	walking	0.2773308	-0.017383819	-0.11114810	-0.283740259	0.114461337	-0.26002790
1	walkingUpstairs	0.2554617	-0.023953149	-0.09730200	-0.354708025	-0.002320265	-0.01947924
1	walkingDownstairs	0.2891883	-0.009918505	-0.10756619	0.030035338	-0.031935943	-0.23043421
1	sitting	0.2612376	-0.001308288	-0.10454418	-0.977229008	-0.922618642	-0.93958629
1	standing	0.2789176	-0.016137590	-0.11060182	-0.995759902	-0.973190056	-0.97977588
1	laying	0.2215982	-0.040513953	-0.11320355	-0.928056469	-0.836827406	-0.82606140
2	walking	0.2764266	-0.018594920	-0.10550036	-0.423642838	-0.078091253	-0.42525752
2	walkingUpstairs	0.2471648	-0.021412113	-0.15251390	-0.304376406	0.108027280	-0.11212102
2	walkingDownstairs	0.2776153	-0.022661416	-0.11681294	0.046366681	0.262881789	-0.10283791
2	sitting	0.2770874	-0.015687994	-0.10921827	-0.986822280	-0.950704499	-0.95982817
2	standing	0.2779115	-0.018420827	-0.10590854	-0.987271889	-0.957304989	-0.94974185
2	laying	0.2813734	-0.018158740	-0.10724561	-0.974059465	-0.980277399	-0.98423330
3	walking	0.2755675	-0.017176784	-0.11267486	-0.360356726	-0.069914065	-0.38741199
3	walkingUpstairs	0.2608199	-0.032410941	-0.11006486	-0.313123438	0.011628081	-0.36975460
3	walkingDownstairs	0.2924235	-0.019355408	-0.11613984	-0.057410048	-0.033150372	-0.36224025
3	sitting	0.2571976	-0.003502998	-0.09835792	-0.971010121	-0.856617812	-0.87511018
3	standing	0.2800465	-0.014337656	-0.10162172	-0.966742543	-0.893449307	-0.91141902
3	laying	0.2755169	-0.018955679	-0.10130048	-0.982776639	-0.962057545	-0.96369103
4	walking	0.2785820	-0.014839948	-0.11140306	-0.440829971	-0.078826737	-0.58625281
4	walkingUpstairs	0.2708767	-0.031980430	-0.11421946	-0.204933043	-0.066689911	-0.37213778
4	walkingDownstairs	0.2799653	-0.009802009	-0.10677752	0.011193550	-0.218598325	-0.47918598
4	sitting	0.2715383	-0.007163065	-0.10587460	-0.980309931	-0.890223981	-0.93220301
4	standing	0.2804997	-0.009489111	-0.09615749	-0.976920580	-0.861597379	-0.89688054
4	laying	0.2635592	-0.015003184	-0.11068815	-0.954193739	-0.941714042	-0.96266731