Codebook Accelerometer Analysis from Samsung Galaxy S smartphone

---

title: Codebook Accelerometer Analysis from Samsung Galaxy S smartphone

author: Irene Statkus

date: July 24 2015

output:

html\_document:

keep\_md: yes

**## Project Description**

Clean and prepare a tidy data set from smartphone activity monitoring data collected. The raw data is a set of files from:

<http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

The consumer of this tidy data set requests an observation per

<SUBJECT> <ACTIVITY> <FEATURE> <Mean of the measure>

For all Subjects

For all Activities

Only for Features that represented means and std

In addition, mulitiple observations are to be averaged

**## SOURCE Study design and data processing <From the Source Documentation>**

The Smartlab has developed a new publicly available database of daily human activities that has been recorded using accelerometer and gyroscope data from a waist-mounted Android-OS smartphone (Samsung Galaxy S II).

<SUBJECTS>

The experiments have been carried out with a group of 30 volunteers {Subject\_train and Subject\_test] within an age bracket of 19-48 years.

For more information visit: [http://smartlab.ws/index.php?option=c...](http://smartlab.ws/index.php?option=com_content&view=article&id=60)

Youtube video: http://www.youtube.com/watch?v=XOEN9W05\_4A

Subject names are not used just a number 1:30

<ACTIVITIES>

Contained the ID for each activity, 1 row per observation matching the length of the X\_train and X-\_test

<FEATURES>

There are 561 distinct measures termed as “Features”

The assumption is that the consumer of the data set is familiar with the technology

And can acquire additional information on the features from the source

The Format of the feature names: Not every combination of prefix and summary measure is valid

Only those features containing mean, MEAN, std were chosen for this tidy data set

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| tBodyAcc | mean | fBodyAccMag | mean | fBodyAcc | mean() |
| tGravityAcc | std | fBodyBodyAccJerkMag | std | fBodyAccJerk | std() |
| tBodyAccJerk | mad | fBodyBodyGyroJerkMag | mad | fBodyGyro | mad() |
| tBodyGyro | max | fBodyBodyGyroMag | max |  | max() |
| tBodyGyroJerk | min | tBodyAccJerkMag | min |  | min() |
|  | sma | tBodyAccMag | sma |  | sma() |
|  | energy | tBodyGyroJerkMag | energy |  | energy() |
|  | iqr | tBodyGyroMag | iqr |  | iqr() |
|  | entropy | tGravityAccMag | entropy |  | entropy() |
|  | arCoeff |  | arCoeff |  | maxInds |
|  | correlation |  |  |  | meanFreq() |
|  |  |  |  |  | skewness() |
|  |  |  |  |  | kurtosis() |
|  |  |  |  |  | bandsEnergy() |
|  |  |  |  |  |  |
| angle |  |  |  |  |  |

**###Description of Collection of the raw data from source documents**

Each person performed six 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, we captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The experiments have been video-recorded to label the data manually. The obtained dataset has been randomly partitioned into two sets, where 70% of the volunteers was selected for generating the training data and 30% the test data.   
  
For more Information: <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

**###Notes on the original (raw) data tables**

activity\_labels – Includes the code and Name of the Activity

1 WALKING

2 WALKING\_UPSTAIRS

3 WALKING\_DOWNSTAIRS

4 SITTING

5 STANDING

6 LAYING

features – Lists the column number and activity measure for columns in X\_train and X\_test. There are 561

X\_train – The training measures

y\_train – The participants, each row represents a participant in X\_train

X\_test – the test measures

y\_test – the participants in X\_test

**##Creating the tidy datafile**

**Set the environment and working directory**

library(dplyr)

library(tidyr)

library(stringr)

library(reshape2)

Step 1: Get the data by Downloading and unzipping

Step 2: Prepare the table for Activity Dimension to use later to swap out the Activity id

Step 3: Prepare the table that will be used to assign the column headers of the measures

Tidied the column names by removing the braces ()

Selected only the columns with measures that represented means and std

Step 4: Read the Train and Test tables Each has 561 columns of measures

Rows are observations per subject/activity id /measure as a column

Step 5: Use the feature vector to add column headers to the measures tables

Step 6: Read the activity vectors for test and train tells me for each row in the measures table

What the activity was for that row

Step 7: Read the subject vector ids who was subject in each row of the measures (X\_)

Step 8: Combine the Subject Vector, Activity and Measures into one table

Step 9: Combine test and train into one table

Step 10 : Switch out Activity id for its List of Value Name and convert to factor

Step11: Gather the wide format into a narrow one so the column becomes a row

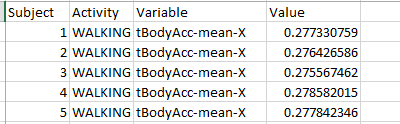
Step 11 summarizing to a subject, activity a single measure and

calculate its mean for all observations

Step 12 write it out

**##Description of the variables in the tidydata.txt file**

Sample:



> str(tidydata)

'data.frame': 15480 obs. of 4 variables:

$ Subject : Factor w/ 30 levels "1","2","3","4",..: 1 2 3 4 5 6 7 8 9 10 ...

$ Activity: Factor w/ 6 levels "WALKING","WALKING\_UPSTAIRS",..: 1 1 1 1 1 1 1 1 1 1 ...

$ Variable: Factor w/ 86 levels "tBodyAcc-mean-X",..: 1 1 1 1 1 1 1 1 1 1 ...

$ Value : num 0.277 0.276 0.276 0.279 0.278 ...

<SUBJECT> 30 Subjects: Ranging from 1:30 (No Names, only a number)

<ACTIVITY>

6 Activities

1 WALKING

2 WALKING\_UPSTAIRS

3 WALKING\_DOWNSTAIRS

4 SITTING

5 STANDING

6 LAYING

<VARIABLE> 86 Features are listed in the variable column, its measure is in the Value column

- Triaxial acceleration from the accelerometer (total acceleration) and the estimated body acceleration.   
- Triaxial Angular velocity from the gyroscope.

tBodyAcc-mean-X tBodyAcc-mean-Y

tBodyAcc-mean-Z tBodyAcc-std-X

tBodyAcc-std-Y tBodyAcc-std-Z

tGravityAcc-mean-X tGravityAcc-mean-Y

tGravityAcc-mean-Z tGravityAcc-std-X

tGravityAcc-std-Y tGravityAcc-std-Z

tBodyAccJerk-mean-X tBodyAccJerk-mean-Y

tBodyAccJerk-mean-Z tBodyAccJerk-std-X

tBodyAccJerk-std-Y tBodyAccJerk-std-Z

tBodyGyro-mean-X tBodyGyro-mean-Y

tBodyGyro-mean-Z tBodyGyro-std-X

tBodyGyro-std-Y tBodyGyro-std-Z

tBodyGyroJerk-mean-X tBodyGyroJerk-mean-Y

tBodyGyroJerk-mean-Z tBodyGyroJerk-std-X

tBodyGyroJerk-std-Y tBodyGyroJerk-std-Z

tBodyAccMag-mean tBodyAccMag-std

tGravityAccMag-mean tGravityAccMag-std

tBodyAccJerkMag-mean tBodyAccJerkMag-std

tBodyGyroMag-mean tBodyGyroMag-std

tBodyGyroJerkMag-mean tBodyGyroJerkMag-std

fBodyAcc-mean-X fBodyAcc-mean-Y

fBodyAcc-mean-Z fBodyAcc-std-X

fBodyAcc-std-Y fBodyAcc-std-Z

fBodyAcc-meanFreq-X fBodyAcc-meanFreq-Y

fBodyAcc-meanFreq-Z fBodyAccJerk-mean-X

fBodyAccJerk-mean-Y fBodyAccJerk-mean-Z

fBodyAccJerk-std-X fBodyAccJerk-std-Y

fBodyAccJerk-std-Z fBodyAccJerk-meanFreq-X

fBodyAccJerk-meanFreq-Y fBodyAccJerk-meanFreq-Z

fBodyGyro-mean-X fBodyGyro-mean-Y

fBodyGyro-mean-Z fBodyGyro-std-X

fBodyGyro-std-Y fBodyGyro-std-Z

fBodyGyro-meanFreq-X fBodyGyro-meanFreq-Y

fBodyGyro-meanFreq-Z fBodyAccMag-mean

fBodyAccMag-std fBodyAccMag-meanFreq

fBodyBodyAccJerkMag-mean fBodyBodyAccJerkMag-std

fBodyBodyAccJerkMag-meanFreq fBodyBodyGyroMag-mean

fBodyBodyGyroMag-std fBodyBodyGyroMag-meanFreq

fBodyBodyGyroJerkMag-mean fBodyBodyGyroJerkMag-std

fBodyBodyGyroJerkMag-meanFreq angletBodyAccMean,gravity

angletBodyAccJerkMean,gravityMean) angletBodyGyroMean,gravityMean

angletBodyGyroJerkMean,gravityMean angleX,gravityMean

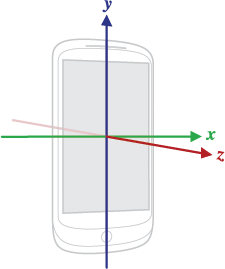
angleY,gravityMean angleZ,gravityMean

<http://www.livescience.com/40103-accelerometer-vs-gyroscope.html>

Gyrocscope – measures rotation around an axis

An accelerometer is a compact device designed to measure non-gravitational acceleration. When the object it’s integrated into goes from a standstill to any velocity, the [accelerometer](http://www.livescience.com/40102-accelerometers.html) is designed to respond to the vibrations associated with such movement. It uses microscopic crystals that go under stress when vibrations occur, and from that stress a voltage is generated to create a reading on any acceleration.

<http://blog.contus.com/how-to-measure-acceleration-in-smartphones-using-accelerometer/>

[](http://blog.contus.com/wp-content/uploads/2013/05/axis_device.png)The accelerometer in the mobile device provides the XYZ coordinate values, which is used to measure the position and the acceleration of the device

<VALUES> Mean value in IMU – Inertial Measurement Unit is assumed unless otherwise documented differently in the source documentation

##Sources and Acknowledgements

For more information about this dataset please contact: activityrecognition '@' smartlab.ws

Acknowledging R for Everyone, Jared Lander

and numerous posts attempting to compare or explain the subtle differences

in syntax dplyr, plyr, tidyr, reshape

the above steps may contain some extra steps, but it helped me verify

based on my interpretation of the requirement

Wikipedia (accelerometers, smartphones, gyroscopes… )

Livescience <http://www.livescience.com/40103-accelerometer-vs-gyroscope.html>

<http://blog.contus.com/how-to-measure-acceleration-in-smartphones-using-accelerometer/>

License:

========

Use of this dataset in publications must be acknowledged by referencing the following publication [1]

[1] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges, Belgium 24-26 April 2013.