**Codebook**

**By Peter Johnson**

This document is about the Assignment for the JHU Data Cleaning course on Coursera.  
More info on [the Coursera assignment page](https://www.coursera.org/learn/data-cleaning/peer/FIZtT/getting-and-cleaning-data-course-project)

**Other related file**

* run\_analysis.R R script performing the actions required to get a tidy data set from original data project
* Readme.md describing the action performed on the above script
* tidydataset.txt resulting tidy data set file from execution of above script

**Context and initial Project Description**

The original experiments were performed with a group of 30 volunteers within an age bracket of 19-48 years. Each person performed six activities (WALKING, WALKING\_UPSTAIRS, WALKING\_DOWNSTAIRS, SITTING, STANDING, LAYING) wearing a smartphone (Samsung Galaxy S II) on the waist. The smartphone used its embedded accelerometer and gyroscope to capture 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The experiments can be viewed at the following URL (<http://www.youtube.com/watch?v=XOEN9W05_4A>)

**Tidy Data Set Structure**

The Tidy Data Set resulting from the execution of run\_analysis.R contains 68 variables.

The first 2 are identifiers for the activity performed and the subject\_id.  
There were 30 subjects in the experiments so ids are in [1-30] range.  
Activities are the following and were coded as given in parenthesis in the original data set (see Credits below):

* WALKING (1)
* WALKING\_UPSTAIRS (2)
* WALKING\_DOWNSTAIRS (3)
* SITTING (4)
* STANDING (5)
* LAYING (6)

The following 66 columns are measures of mean or standard deviation for each activity and each subjects (ie. they have been computed from orginal measures by aggregating by activity and subject then applying the mean).  
Below is the list of all 66 measures:

1. timeBodyAcceleration-mean-X
2. timeBodyAcceleration-mean-Y
3. timeBodyAcceleration-mean-Z
4. timeBodyAcceleration-stdDev-X
5. timeBodyAcceleration-stdDev-Y
6. timeBodyAcceleration-stdDev-Z
7. timeGravityAcceleration-mean-X
8. timeGravityAcceleration-mean-Y
9. timeGravityAcceleration-mean-Z
10. timeGravityAcceleration-stdDev-X
11. timeGravityAcceleration-stdDev-Y
12. timeGravityAcceleration-stdDev-Z
13. timeBodyAccelerationJerk-mean-X
14. timeBodyAccelerationJerk-mean-Y
15. timeBodyAccelerationJerk-mean-Z
16. timeBodyAccelerationJerk-stdDev-X
17. timeBodyAccelerationJerk-stdDev-Y
18. timeBodyAccelerationJerk-stdDev-Z
19. timeBodyGyroscope-mean-X
20. timeBodyGyroscope-mean-Y
21. timeBodyGyroscope-mean-Z
22. timeBodyGyroscope-stdDev-X
23. timeBodyGyroscope-stdDev-Y
24. timeBodyGyroscope-stdDev-Z
25. timeBodyGyroscopeJerk-mean-X
26. timeBodyGyroscopeJerk-mean-Y
27. timeBodyGyroscopeJerk-mean-Z
28. timeBodyGyroscopeJerk-stdDev-X
29. timeBodyGyroscopeJerk-stdDev-Y
30. timeBodyGyroscopeJerk-stdDev-Z
31. timeBodyAccelerationMagnitude-mean
32. timeBodyAccelerationMagnitude-stdDev
33. timeGravityAccelerationMagnitude-mean
34. timeGravityAccelerationMagnitude-stdDev
35. timeBodyAccelerationJerkMagnitude-mean
36. timeBodyAccelerationJerkMagnitude-stdDev
37. timeBodyGyroscopeMagnitude-mean
38. timeBodyGyroscopeMagnitude-stdDev
39. timeBodyGyroscopeJerkMagnitude-mean
40. timeBodyGyroscopeJerkMagnitude-stdDev
41. freqBodyAcceleration-mean-X
42. freqBodyAcceleration-mean-Y
43. freqBodyAcceleration-mean-Z
44. freqBodyAcceleration-stdDev-X
45. freqBodyAcceleration-stdDev-Y
46. freqBodyAcceleration-stdDev-Z
47. freqBodyAccelerationJerk-mean-X
48. freqBodyAccelerationJerk-mean-Y
49. freqBodyAccelerationJerk-mean-Z
50. freqBodyAccelerationJerk-stdDev-X
51. freqBodyAccelerationJerk-stdDev-Y
52. freqBodyAccelerationJerk-stdDev-Z
53. freqBodyGyroscope-mean-X
54. freqBodyGyroscope-mean-Y
55. freqBodyGyroscope-mean-Z
56. freqBodyGyroscope-stdDev-X
57. freqBodyGyroscope-stdDev-Y
58. freqBodyGyroscope-stdDev-Z
59. freqBodyAccelerationMagnitude-mean
60. freqBodyAccelerationMagnitude-stdDev
61. freqBodyAccelerationJerkMagnitude-mean
62. freqBodyAccelerationJerkMagnitude-stdDev
63. freqBodyGyroscopeMagnitude-mean
64. freqBodyGyroscopeMagnitude-stdDev
65. freqBodyGyroscopeJerkMagnitude-mean
66. freqBodyGyroscopeJerkMagnitude-stdDev

**Notes:**

* features are normalized and bounded within [-1,1]
* Units are in radian/second for time measures, in hertz for frequency measures

**Tidy Data Set Content**

The data set contents 180 observations (6 activities x 30 subjects) of the 66 variables (means of each average and standard deviation) from the original UCI HAR dataset.  
The first line of the file consists of the header with the variables names.