CodeBook: Project for Getting and Cleaning Data Course

Contents

Variables in the data set

ActivityNames

This variable represents the name of an experiment activity. The type of the variable is Character. The possible names are:

- WALKING
- WALKING_UPSTAIRS
- WALKING_DOWNSTAIRS
- SITTING
- STANDING
- LAYING

DataSetName

Types of the data set, it is a character variable. The two possible categories come from the experiment definition:

- Train
- Test

SubjectId

Identification numbers of the subjects involved in the experiment. They are anonymous participants and this SubjectId field is an integer number from 1 to 30.

TBodyAccMeanX

Mean of the mean for TBodyAcc in the X axis. It is a numeric field.

TBodyAccMeanY

Mean of the mean for TBodyAcc in the Y axis. It is a numeric field.

TBodyAccMeanZ

Mean of the mean for TBodyAcc in the Z axis. It is a numeric field.

TBodyAccStdX

Mean of the Standard Deviation for TBodyAcc in the X axis. It is a numeric field.

TBodyAccStdY

Mean of the Standard Deviation for TBodyAcc in the Y axis. It is a numeric field.

TBodyAccStdZ

Mean of the Standard Deviation for TBodyAcc in the Z axis. It is a numeric field.

TGravityAccMeanX

Mean of the Mean for TGravityAcc in the X axis. It is a numeric field.

TGravityAccMeanY

Mean of the Mean for TGravityAcc in the Y axis. It is a numeric field.

TGravityAccMeanZ

Mean of the Mean for TGravityAcc in the Z axis. It is a numeric field.

TGravityAccStdX

Mean of the Standard deviation for TGravityAcc in the X axis. It is a numeric field.

TGravityAccStdY

Mean of the Standard deviation for TGravityAcc in the Y axis. It is a numeric field.

TGravityAccStdZ

Mean of the Standard deviation for TGravityAcc in the Z axis. It is a numeric field.

TBodyAccJerkMeanX

Mean of the Mean for TBodyAccJerk in the X axis. It is a numeric field.

TBodyAccJerkMeanY

Mean of the Mean for TBodyAccJerk in the Y axis. It is a numeric field.

TBodyAccJerkMeanZ

Mean of the Mean for TBodyAccJerk in the Z axis. It is a numeric field.

TBodyAccJerkStdX

Mean of the Standard Deviation for TBodyAccJerk in the X axis. It is a numeric field.

TBodyAccJerkStdY

Mean of the Standard Deviation for TBodyAccJerk in the Y axis. It is a numeric field.

TBodyAccJerkStdZ

Mean of the Standard Deviation for TBodyAccJerk in the Z axis. It is a numeric field.

TBodyGyroMeanX

Mean of the Mean for TBodyGyro in the X axis. It is a numeric field.

TBodyGyroMeanY

Mean of the Mean for TBodyGyro in the Y axis. It is a numeric field.

TBodyGyroMeanZ

Mean of the Mean for TBodyGyro in the Z axis. It is a numeric field.

TBodyGyroStdX

Mean of the Standard Deviation for TBodyGyro in the X axis. It is a numeric field.

TBodyGyroStdY

Mean of the Standard Deviation for TBodyGyro in the Y axis. It is a numeric field.

TBodyGyroStdZ

Mean of the Standard Deviation for TBodyGyro in the Z axis. It is a numeric field.

TBodyGyroJerkMeanX

Mean of the Mean for TBodyGyroJerk in the X axis. It is a numeric field.

TBodyGyroJerkMeanY

Mean of the Mean for TBodyGyroJerk in the Y axis. It is a numeric field.

TBodyGyroJerkMeanZ

Mean of the Mean for TBodyGyroJerk in the Z axis. It is a numeric field.

TBodyGyroJerkStdX

Mean of the Standard Deviation for TBodyGyroJerk in the X axis. It is a numeric field.

TBodyGyroJerkStdY

Mean of the Standard Deviation for TBodyGyroJerk in the Y axis. It is a numeric field.

TBodyGyroJerkStdZ

Mean of the Standard Deviation for TBodyGyroJerk in the Z axis. It is a numeric field.

TBodyAccMagMean

Mean of the Mean for TBodyAccMag. It is a numeric field.

TBodyAccMagStd

Mean of the Standard Deviation for TBodyAccMag. It is a numeric field.

TGravityAccMagMean

Mean of the Mean for TGravityAccMag. It is a numeric field.

TGravityAccMagStd

Mean of the Standard Deviation for TGravityAccMag. It is a numeric field.

TBodyAccJerkMagMean

Mean of the Mean for TBodyAccJerkMag. It is a numeric field.

TBodyAccJerkMagStd

Mean of the Standard Deviation for TBodyAccJerkMag. It is a numeric field.

TBodyGyroMagMean

Mean of the Mean for TBodyGyroMag. It is a numeric field.

TBodyGyroMagStd

Mean of the Standard Deviation for TBodyGyroMag. It is a numeric field.

TBodyGyroJerkMagMean

Mean of the Mean for TBodyGyroJerkMag. It is a numeric field.

TBodyGyroJerkMagStd

Mean of the Standard Deviation for TBodyGyroJerkMag. It is a numeric field.

FBodyAccMeanX

Mean of the Mean for FBodyAcc for the X axis. It is a numeric field.

FBodyAccMeanY

Mean of the Mean for FBodyAcc for the Y axis. It is a numeric field.

FBodyAccMeanZ

Mean of the Mean for FBodyAcc for the Z axis. It is a numeric field.

FBodyAccStdX

Mean of the Standard Deviation for FBodyAcc for the X axis. It is a numeric field.

FBodyAccStdY

Mean of the Standard Deviation for FBodyAcc for the Y axis. It is a numeric field.

FBodyAccStdZ

Mean of the Standard Deviation for FBodyAcc for the Z axis. It is a numeric field.

FBodyAccMeanFreqX

Mean of the mean of FBodyAcc frequency for X Axis. It is a numeric field.

FBodyAccMeanFreqY

Mean of the mean of FBodyAcc frequency for Y Axis. It is a numeric field.

FBodyAccMeanFreqZ

Mean of the mean of FBodyAcc frequency for Z Axis. It is a numeric field.

FBodyAccJerkMeanX

Mean of the mean of FBodyAccJerk for X Axis. It is a numeric field.

FBodyAccJerkMeanY

Mean of the mean of FBodyAccJerk for Y Axis. It is a numeric field.

FBodyAccJerkMeanZ

Mean of the mean of FBodyAccJerk for Z Axis. It is a numeric field.

FBodyAccJerkStdX

Mean of the Standard Deviation of FBodyAccJerk for X Axis. It is a numeric field.

FBodyAccJerkStdY

Mean of the Standard Deviation of FBodyAccJerk for Y Axis. It is a numeric field.

FBodyAccJerkStdZ

Mean of the Standard Deviation of FBodyAccJerk for Z Axis. It is a numeric field.

FBodyAccJerkMeanFreqX

Mean of the mean of FBodyAccJerk Frequency for X Axis. It is a numeric field.

FBodyAccJerkMeanFreqY

Mean of the mean of FBodyAccJerk Frequency for Y Axis. It is a numeric field.

FBodyAccJerkMeanFreqZ

Mean of the mean of FBodyAccJerk Frequency for Z Axis. It is a numeric field.

FBodyGyroMeanX

Mean of the mean of FBodyGyro for X Axis. It is a numeric field.

FBodyGyroMeanY

Mean of the mean of FBodyGyro for Y Axis. It is a numeric field.

FBodyGyroMeanZ

Mean of the mean of FBodyGyro for Z Axis. It is a numeric field.

FBodyGyroStdX

Mean of the Standard Deviation of FBodyGyro for X Axis. It is a numeric field.

FBodyGyroStdY

Mean of the Standard Deviation of FBodyGyro for Y Axis. It is a numeric field.

FBodyGyroStdZ

Mean of the Standard Deviation of FBodyGyro for Z Axis. It is a numeric field.

FBodyGyroMeanFreqX

Mean of the Mean of FBodyGyro Frequency for X Axis. It is a numeric field.

FBodyGyroMeanFreqY

Mean of the Mean of FBodyGyro Frequency for Y Axis. It is a numeric field.

FBodyGyroMeanFreqZ

Mean of the Mean of FBodyGyro Frequency for Z Axis. It is a numeric field.

FBodyAccMagMean

Mean of the Mean of FBodyAccMag. It is a numeric field.

FBodyAccMagStd

Mean of the Standard Deviation of FBodyAccMag. It is a numeric field.

FBodyAccMagMeanFreq

Mean of the Mean of FBodyAccMag Frequency. It is a numeric field.

FBodyBodyAccJerkMagMean

Mean of the Mean of FBodyBodyAccJerkMag Frequency. It is a numeric field.

FBodyBodyAccJerkMagStd

Mean of the Standard Deviation of FBodyBodyAccJerkMag Frequency. It is a numeric field.

FBodyBodyAccJerkMagMeanFreq

Mean of the Mean of FBodyBodyAccJerkMag Frequency. It is a numeric field.

FBodyBodyGyroMagMean

Mean of the Mean of FBodyBodyGyroMag. It is a numeric field.

FBodyBodyGyroMagStd

Mean of the Standard Deviation of FBodyBodyGyroMag. It is a numeric field.

FBodyBodyGyroMagMeanFreq

Mean of Mean of FBodyBodyGyroMag Frequency. It is a numeric field.

FBodyBodyGyroJerkMagMean

Mean of Mean of FBodyBodyGyroJerkMag. It is a numeric field.

FBodyBodyGyroJerkMagStd

Mean of Standard Deviation of FBodyBodyGyroJerkMag. It is a numeric field.

FBodyBodyGyroJerkMagMeanFreq

Mean of Mean of FBodyBodyGyroJerkMag Frequency. It is a numeric field.

Angle_TBodyAccMean_Gravity

Mean of the Angle of TBodyAcc vs Gravity. It is a numeric field.

Angle_TBodyAccJerkMean_GravityMean

Mean of the Angle of TBodyAccJerk vs Gravity. It is a numeric field.

Angle_TBodyGyroMean_GravityMean

Mean of the Angle of TBodyAccJerk vs Gravity. It is a numeric field.

Angle_TBodyGyroJerkMean_GravityMean

Mean of the Angle of TBodyGyroJerk vs Gravity. It is a numeric field.

Angle_X_GravityMean

Mean of the Angle of X axis vs Gravity. It is a numeric field.

Angle_Y_GravityMean

Mean of the Angle of Y axis vs Gravity. It is a numeric field.

Angle_Z_GravityMean

Mean of the Angle of Z axis vs Gravity. It is a numeric field.

Summarization Techniques

The data was group by Subject Id, Activity Name and Data Set type. The Data Set Type consists of two categories: Train and Test. Because of the two different purposes of train and test data the distinction was made.

Following the instructions for the project, the data was processed using the dplyr package. From the package, the group_by function was used with the summarize_each function to create means for each measurement.

The tidy data is a wide data set of columns. The measurements were added as columns so that each mean of measurements of observations are group by activity, subject and Data Set Type.

Study Design

The data was collected for a specific study:

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.

The authors created a Human activity recognition database. The data was collected from 30 subjects doing a specific set of activities while they were wearing a Smartphone (Samsung Galaxy S II). In this project, the R scripts are merging and transforming the data to generate a tidy data set.

For information about the original experiment and details please follow the official link for the dataset:

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