Getting and Cleaning Data Course Project Code Book

Elizabeth Lundeen

3/18/2022

The data set described in this code book is located in the tidy_data.txt file of this repository.

See the README.md file of this repository for background information on this data set.

Data

The original data have been cleaned and prepared as tidy data and are contained in a text file called tidy_data.txt.

The first row in the text file contains the variable names and the subsequent rows contain the data for each of the 30 participants (1 row per participant per activity [30 x 6 = 180 rows]).

Data Cleaning and Transformations to Create the Tidy Dataset

The original dataset included 561 features, including time and frequency domain variables. Signals were measured for the X axis, Y axis, and Z axis.

Steps to create a tidy dataset:

1. The original dataset was randomly partitioned into two sets, where 70% of the volunteers were selected for generating the training data and 30% for the test data. The first step in data cleaning was to append the test and training datasets to create one dataset.
2. From the original 561 features, step 2 was to extract only the measurements on the mean and standard deviation for each measurement. This reduced the dataset down to 79 features.
3. Activities in the data set were labeled using descriptive activity names.
4. The next step was to appropriately label the data set with descriptive variable names.
5. Finally, an independent tidy data set was created which contains the average of each variable for each activity and each subject.

Description of variables in the tidy_data.txt file

- participant
  Unique ID for each study participant (an integer ranging from 1 to 30).
- activity

There are six types of activities in this dataset:
- WALKING
- WALKING UPSTAIRS
- WALKING DOWNSTAIRS
- SITTING
- STANDING
- LAYING

Measurements Taken Using the Accelerometer Sensor Signal

The sensor acceleration signal was separated into body acceleration and gravity. The features are calculated variables from the time and frequency domain. The mean and standard deviation for these measurements are provided.

Time Domain

Mean and standard deviation for accelerometer measurements – body acceleration – time domain
- TimeBodyAccelerometerMeanX
- TimeBodyAccelerometerMeanY
- TimeBodyAccelerometerMeanZ
- TimeBodyAccelerometerStandardDeviationX
- TimeBodyAccelerometerStandardDeviationY
- TimeBodyAccelerometerStandardDeviationZ

Mean and standard deviation for accelerometer measurements – body acceleration jerk – time domain
Jerk is the rate of change of acceleration, or the time derivative of acceleration, and as such the second derivative of velocity.
- TimeBodyAccelerometerJerkMeanX
- TimeBodyAccelerometerJerkMeanY
- TimeBodyAccelerometerJerkMeanZ
- TimeBodyAccelerometerJerkStandardDeviationX
- TimeBodyAccelerometerJerkStandardDeviationY
- TimeBodyAccelerometerJerkStandardDeviationZ

Mean and standard deviation for accelerometer measurements – gravity acceleration – time domain
- TimeGravityAccelerometerMeanX
- TimeGravityAccelerometerMeanY
- TimeGravityAccelerometerMeanZ
- TimeGravityAccelerometerStandardDeviationX
- TimeGravityAccelerometerStandardDeviationY
- TimeGravityAccelerometerStandardDeviationZ

Mean and standard deviation for accelerometer measurements – magnitude of body acceleration – time domain
- TimeBodyAccelerometerMagnitudeMean
- TimeBodyAccelerometerMagnitudeStandardDeviation

Mean and standard deviation for accelerometer measurements – magnitude of gravity acceleration – time domain
- TimeGravityAccelerometerMagnitudeMean
- TimeGravityAccelerometerMagnitudeStandardDeviation

Mean and standard deviation for accelerometer measurements – magnitude of body acceleration jerk – time domain

Jerk is the rate of change of acceleration, or the time derivative of acceleration, and as such the second derivative of velocity.
- TimeBodyAccelerometerJerkMagnitudeMean
- TimeBodyAccelerometerJerkMagnitudeStandardDeviation

Frequency Domain

Mean and standard deviation for accelerometer measurements – body acceleration – frequency domain
- FrequencyAccelerometerMeanX
- FrequencyAccelerometerMeanY
- FrequencyAccelerometerMeanZ
- FrequencyAccelerometerStandardDeviationX
- FrequencyAccelerometerStandardDeviationY
- FrequencyAccelerometerStandardDeviationZ

Mean and standard deviation for accelerometer measurements – body acceleration jerk – frequency domain
Jerk is the rate of change of acceleration, or the time derivative of acceleration, and as such the second derivative of velocity.
- FrequencyAccelerometerJerkMeanX
- FrequencyAccelerometerJerkMeanY
- FrequencyAccelerometerJerkMeanZ
- FrequencyAccelerometerJerkStandardDeviationX
- FrequencyAccelerometerJerkStandardDeviationY
- FrequencyAccelerometerJerkStandardDeviationZ

Average, standard deviation, and weighted average for accelerometer measurements – magnitude of body acceleration – frequency domain
- FrequencyAccelerometerMagnitudeMean
- FrequencyAccelerometerMagnitudeStandardDeviation
- FrequencyAccelerometerMagnitudeMeanFrequency

Average, standard deviation, and weighted average – magnitude of body acceleration jerk – frequency domain
- FrequencyBodyAccelerometerJerkMagnitudeMean
- FrequencyBodyAccelerometerJerkMagnitudeStandardDeviation
- FrequencyBodyAccelerometerJerkMagnitudeMeanFrequency

Weighted average of the frequency components – body acceleration – frequency domain
- FrequencyAccelerometerMeanFrequencyX
- FrequencyAccelerometerMeanFrequencyY
- FrequencyAccelerometerMeanFrequencyZ

Weighted average of the frequency components – body acceleration jerk – frequency domain
Jerk is the rate of change of acceleration, or the time derivative of acceleration, and as such the second derivative of velocity.
- FrequencyAccelerometerJerkMeanFrequencyX
- FrequencyAccelerometerJerkMeanFrequencyY
- FrequencyAccelerometerJerkMeanFrequencyZ

Measurements Taken Using the Gyroscope Sensor Signal

Triaxial Angular velocity from the gyroscope.

Time Domain

Mean and standard deviation for gyroscope measurements – body angular velocity– time domain
- TimeBodyGyroscopeMeanX
- TimeBodyGyroscopeMeanY
- TimeBodyGyroscopeMeanZ
- TimeBodyGyroscopeStandardDeviationX
- TimeBodyGyroscopeStandardDeviationY
- TimeBodyGyroscopeStandardDeviationZ

Mean and standard deviation for gyroscope measurements – body angular velocity jerk – time domain
Jerk is the rate of change of acceleration, or the time derivative of acceleration, and as such the second derivative of velocity.
- TimeBodyGyroscopeJerkMeanX
- TimeBodyGyroscopeJerkMeanY
- TimeBodyGyroscopeJerkMeanZ
- TimeBodyGyroscopeJerkStandardDeviationX
- TimeBodyGyroscopeJerkStandardDeviationY
- TimeBodyGyroscopeJerkStandardDeviationZ

Mean and standard deviation for gyroscope measurements – magnitude of body angular velocity – time domain
- TimeBodyGyroscopeMagnitudeMean
- TimeBodyGyroscopeMagnitudeStandardDeviation

Mean and standard deviation for gyroscope measurements – magnitude of body angular velocity jerk – time domain
- TimeBodyGyroscopeJerkMagnitudeMean
- TimeBodyGyroscopeJerkMagnitudeStandardDeviation

Frequency Domain

Mean and standard deviation for gyroscope measurements – body angular velocity – frequency domain
- FrequencyGyroscopeMeanX
- FrequencyGyroscopeMeanY
- FrequencyGyroscopeMeanZ
- FrequencyGyroscopeStandardDeviationX
- FrequencyGyroscopeStandardDeviationY
- FrequencyGyroscopeStandardDeviationZ

Weighted average of the frequency components – body angular velocity – frequency domain
- FrequencyGyroscopeMeanFrequencyX
- FrequencyGyroscopeMeanFrequencyY
- FrequencyGyroscopeMeanFrequencyZ

Mean, standard deviation, and weighted average of the frequency components – magnitude of body angular velocity – frequency domain - FrequencyBodyGyroscopeMagnitudeMean
- FrequencyBodyGyroscopeMagnitudeStandardDeviation
- FrequencyBodyGyroscopeMagnitudeMeanFrequency

Mean, standard deviation, and weighted average of the frequency components – magnitude of body angular velocity jerk – frequency domain
- FrequencyBodyGyroscopeJerkMagnitudeMean
- FrequencyBodyGyroscopeJerkMagnitudeStandardDeviation
- FrequencyBodyGyroscopeJerkMagnitudeMeanFrequency