

CODEBOOK FOR SAMSUNG DATA

Aditya

September 26, 2015

CODE BOOK FOR COURSE PROJECT ON "GETTING AND CLEANING DATA"

VARIABLE NAMES

SUBJECT_NUMBER 2

This variable identifies the ID of the subject who performed the designated activity for each window sample (Integer: No units of measure)

1..30 .Unique identifier assigned to each subject

ACTIVITY_NUMBER 1

This variable identifies the type of activity performed by the subject. Each number maps to a particular activity type (Integer: No units of measure)

1..6

1. WALKING
2. WALKING_UPSTAIRS
3. WALKING_DOWNSTAIRS
4. SITTING
5. STANDING
6. LAYING

ACTIVITY_NAME 17

This variable is the value mapped by the key which is the activity_number. (Character vector/string: No units of measure)

WALKING..LAYING

TBODYACC-MEAN() -X 20

Average Mean time domain body accelerometer signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACC-MEAN()-Y 20

Average Mean time domain body accelerometer signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACC-MEAN()-Z 20

Average Mean time domain body accelerometer signal in Z direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACC-STD()-X 20

Average standard deviation of time domain body accelerometer signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACC-STD()-Y 20

Average standard deviation of time domain body accelerometer signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACC-STD()-Z 20

Average standard deviation of time domain body accelerometer signal in Z direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TGRAVITYACC-MEAN()-X 20

Average mean of time domain gravity accelerometer signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TGRAVITYACC-MEAN()-Y 20

Average mean of time domain gravity accelerometer signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TGRAVITYACC-MEAN()-Z 20

Average mean of time domain gravity accelerometer signal in Z direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TGRAVITYACC-STD()-X 20

Average standard deviation of time domain gravity accelerometer signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TGRAVITYACC-STD()-Y 20

Average standard deviation of time domain gravity accelerometer signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TGRAVITYACC-STD()-Z 20

Average standard deviation of time domain gravity accelerometer signal in Z direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCJERK-MEAN()-X 20

Average mean of time domain body acceleration jerk signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCJERK-MEAN()-Y 20

Average mean of time domain body acceleration jerk signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCJERK-MEAN()-Z 20

Average mean of time domain body acceleration jerk signal in Z direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCJERK-STD()-X 20

Average standard deviation of time domain body acceleration jerk signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCJERK-STD()-Y 20

Average standard deviation of time domain body acceleration jerk signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCJERK-STD()-Z 20

Average standard deviation of time domain body acceleration jerk signal in Z direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYRO-MEAN()-X 20

Average mean of time domain body acceleration gyroscope signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYRO-MEAN()-Y 20

Average mean of time domain body acceleration gyroscope signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYRO-MEAN()-Z 20

Average mean of time domain body acceleration gyroscope signal in Z direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYRO-STD()-X 20

Average standard deviation of time domain body acceleration gyroscope signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYRO-STD()-Y 20

Average standard deviation of time domain body acceleration gyroscope signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYRO-STD()-Z 20

Average standard deviation of time domain body acceleration gyroscope signal in Z direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROJERK-MEAN()-X 20

Average mean of time domain body acceleration gyroscope jerk signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROJERK-MEAN()-Y 20

Average mean of time domain body gyroscope jerk signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROJERK-MEAN()-Z 20

Average standard deviation of time domain body gyroscope jerk signal in Z direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROJERK-STD()-X 20

Average standard deviation of time domain body gyroscope jerk signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROJERK-STD()-Y 20

Average standard deviation of time domain body gyroscope jerk signal in Y direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROJERK-STD()-Z 20

Average standard deviation of time domain body gyroscope jerk signal in X direction calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCMAG-MEAN() 20

Average triaxial mean of time domain body acceleration magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCMAG-STD() 20

Average triaxial standard deviation of time domain body acceleration magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TGRAVITYACCMAG-MEAN() 20

Average triaxial mean of time domain gravity acceleration magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TGRAVITYACCMAG-STD() 20

Average triaxial standard deviation of time domain gravity acceleration magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCJERKMAG-MEAN() 20

Average triaxial mean of time domain body acceleration jerk magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYACCJERKMAG-STD() 20

Average triaxial standard deviation of time domain body acceleration jerk magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROMAG-MEAN() 20

Average triaxial mean of time domain body gyrometric magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROMAG-STD() 20

Average triaxial standard deviation of time domain body gyrometric magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROJERKMAG-MEAN() 20

Average triaxial mean of time domain body gyrometric jerk magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

TBODYGYROJERKMAG-STD() 20

Average triaxial standard deviation of time domain body gyrometric jerk magnitude calculated for all measurements within 1 activity for a subject (Numeric: seconds)

FBODYACC-MEAN()-X 20

Average mean of frequency domain body acceleration in X direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACC-MEAN()-Y 20

Average mean of frequency domain body acceleration in Y direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACC-MEAN()-Z 20

Average mean of frequency domain body acceleration in Z direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACC-STD()-X 20

Average standard deviation of frequency domain body acceleration in X direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACC-STD()-Y 20

Average standard deviation of frequency domain body acceleration in Y direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACC-STD()-Z 20

Average standard deviation of frequency domain body acceleration in Z direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACC-MEANFREQ()-X 20

Average mean frequency of the frequency domain body acceleration in X direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACC-MEANFREQ()-Y 20

Average mean frequency of the frequency domain body acceleration in Y direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACC-MEANFREQ()-Z 20

Average mean frequency of the frequency domain body acceleration in Z direction calculated for all

measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCJERK-MEAN()-X 20

Average mean of the frequency domain body acceleration jerk in X direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCJERK-MEAN()-Y 20

Average mean of the frequency domain body acceleration jerk in Y direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCJERK-MEAN()-Z 20

Average mean of the frequency domain body acceleration jerk in Z direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCJERK-STD()-X 20

Average standard deviation of the frequency domain body acceleration jerk in X direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCJERK-STD()-Y 20

Average standard deviation of the frequency domain body acceleration jerk in Y direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCJERK-STD()-Z 20

Average standard deviation of the frequency domain body acceleration jerk in Z direction calculated for all

measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCJERK-MEANFREQ()-X 20

Average mean frequency of the frequency domain body acceleration jerk in X direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCJERK-MEANFREQ()-Y 20

Average mean frequency of the frequency domain body acceleration jerk in Y direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCJERK-MEANFREQ()-Z 20

Average mean frequency of the frequency domain body acceleration jerk in Z direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYGYRO-MEAN()-X 20

Average mean of the frequency domain body gyrometric in X direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYGYRO-MEAN()-Y 20

Average mean of the frequency domain body gyrometric in Y direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYGYRO-MEAN()-Z 20

Average mean of the frequency domain body gyrometric in Z direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

`FBODYGYRO-STD()-X 20`

Average standard deviation of the frequency domain body gyrometric in X direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

`FBODYGYRO-STD()-Y 20`

Average standard deviation of the frequency domain body gyrometric in Y direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

`FBODYGYRO-STD()-Z 20`

Average standard deviation of the frequency domain body gyrometric in Z direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

`FBODYGYRO-MEANFREQ()-X 20`

Average mean frequency of the frequency domain body gyrometric in X direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

`FBODYGYRO-MEANFREQ()-Y 20`

Average mean frequency of the frequency domain body gyrometric in Y direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

`FBODYGYRO-MEANFREQ()-Z 20`

Average mean frequency of the frequency domain body gyrometric in Z direction calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCMAG-MEAN() 20

Average mean of the frequency domain body accelerometer magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCMAG-STD() 20

Average standard deviation of the frequency domain body accelerometer magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYACCMAG-MEANFREQ() 20

Average mean frequency of the frequency domain body accelerometer magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYBODYACCJERKMAG-MEAN() 20

Average mean of the frequency domain body accelerometer jerk magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYBODYACCJERKMAG-STD() 20

Average standard deviation of the frequency domain body accelerometer jerk magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYBODYACCJERKMAG-MEANFREQ() 20

Average mean frequency of the frequency domain body accelerometer jerk magnitude calculated for all

measurements within 1 activity for a subject (Numeric: Hz)

FBODYBODYGYROMAG-MEAN() 20

Average mean of the frequency domain body gyroemtric magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYBODYGYROMAG-STD() 20

Average standard deviation of the frequency domain body gyrometric magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYBODYGYROMAG-MEANFREQ() 20

Average mean frequency of the frequency domain body gyrometric magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYBODYGYROJERKMAG-MEAN() 20

Average mean of the frequency domain body gyrometric jerk magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

FBODYBODYGYROJERKMAG-STD() 20

Average standard deviation of the frequency domain body gyrometric jerk magnitude calculated for all measurements within 1 activity for a subject (Numeric: Hz)

SUMMARIES CALCULATED

In order to group and compute the average of the measurements which belonged to the same activity for a particular subject, the **apply** function was used, as shown in the code. Here, *i* is a counter variable looped

from 1 through 30 for each of the subjects. The first line is used for grouping the data based on the subject number while the second line finds the mean for all activity numbers within that subject number.

MERGING AND SUBSETTING THE DATASETS

In order to merge the two datasets, the function **rbind** was used. The function accepts the two datasets: *training* and *testing* datasets. The column names were added to the *testing* dataset before the merging. The column names assigned were meaningful and intuitive as suggested by the question.

The *testing* dataset has 2947 observations and 561 variables, while *training* dataset has 7352 observations and 561 variables. After merging, the resulting dataset has 10299 observations and 561 variables.

Three more columns were added to the datasets: subject number, activity number and activity name. The dataset was then subsetted using the **subset** function and using the **select** function within it to weed out all columns that do not perform either the *mean* and *standard deviation*. After the subsetting, the resulting dataset has 10299 observations but 82 + 3 variables. While 82 variables correspond to the originally present, the last 3 (later moved to the front/head of the dataset) correspond to the subject number, activity number and activity name respectively.

The above detailed summarizing was applied to this dataset in order to obtain a tidy dataset which satisfies the basic principles of tidy dataset. This dataset has 180 observations and 85 variables.