**DATA CODE BOOK**

**“X\_data”**

subjects

Integral, from 1 to 30, each number represents a volunteer ID. We have total 30 volunteers. They provided total 10299 recodes in data set.

activity\_labels

There are six activities for each person to perform. There are:

* WALKING
* WALKING\_UPSTAIRS
* WALKING\_DOWNSTAIRS
* SITTING
* STANDING
* LAYING

Recording Variables

There are total 561 variables recorded.

These signals were used to estimate variables of the feature vector for each pattern:

'-XYZ' is used to denote 3-axial signals in the X, Y and Z directions.

* tBodyAcc-XYZ
* tGravityAcc-XYZ
* tBodyAccJerk-XYZ
* tBodyGyro-XYZ
* tBodyGyroJerk-XYZ
* tBodyAccMag
* tGravityAccMag
* tBodyAccJerkMag
* tBodyGyroMag
* tBodyGyroJerkMag
* fBodyAcc-XYZ
* fBodyAccJerk-XYZ
* fBodyGyro-XYZ
* fBodyAccMag
* fBodyAccJerkMag
* fBodyGyroMag
* fBodyGyroJerkMag

The set of variables that were estimated from these signals are:

* mean(): Mean value
* std(): Standard deviation
* mad(): Median absolute deviation
* max(): Largest value in array
* min(): Smallest value in array
* sma(): Signal magnitude area
* energy(): Energy measure. Sum of the squares divided by the number of values.
* iqr(): Interquartile range
* entropy(): Signal entropy
* arCoeff(): Autorregresion coefficients with Burg order equal to 4
* correlation(): correlation coefficient between two signals
* maxInds(): index of the frequency component with largest magnitude
* meanFreq(): Weighted average of the frequency components to obtain a mean frequency
* skewness(): skewness of the frequency domain signal
* kurtosis(): kurtosis of the frequency domain signal
* bandsEnergy(): Energy of a frequency interval within the 64 bins of the FFT of each window.
* angle(): Angle between to vectors.

Additional vectors obtained by averaging the signals in a signal window sample. These are used on the angle() variable:

* gravityMean
* tBodyAccMean
* tBodyAccJerkMean
* tBodyGyroMean
* tBodyGyroJerkMean

**“X\_mean\_std”**

The “subject and activity\_labels” variables are the same as in “X\_data” or “Q5\_1\_3\_4.csv” explained above.

Other 79 variables are subset from “X\_data” by extracting only the measurements on the mean and standard deviation for each measurement.

**“X\_mean\_sub\_activity”**

The “subject and activity\_labels” variables are the same as in “X\_data” or “Q5\_1\_3\_4.csv” explained above.

mean\_average

Average of each variable mean for each activity and each subject. Since we have total 6 activities and 30 subjects, we have total 180 groups.

n

How many records used in above group averaging