CODEBOOK

The summary.txt file contains the a dataset summarized from the **Human Activity Recognition Using Smartphones Data Set**:

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

The original dataset represents data collected from the accelerometers from the Samsung Galaxy S smartphone. More information can be found by going to the URL above.

The variables included inside the summary.txt file are:

activity:

TYPE: "CHARACTER"
VALUES: WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, LAYING

Activities are divided into 6 categories describing the actions by individual subjects participating in the study: WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, LAYING

subject_ID:

TYPE: "NUMERIC"
VALUES: 1-30

Subject IDs identifying the person participating in the study. Values range between 1-30.

feature:

TYPE: "CHARACTER"
VALUES:
fBodyAcc-mean()-X
fBodyAcc-mean()-Y
fBodyAcc-mean()-Z
fBodyAcc-meanFreq()-X
fBodyAcc-meanFreq()-Y
fBodyAcc-meanFreq()-Z
fBodyAcc-std()-X
fBodyAcc-std()-Y

fBodyAcc-std()-Z
fBodyAccJerk-mean()-X
fBodyAccJerk-mean()-Y
fBodyAccJerk-mean()-Z
fBodyAccJerk-meanFreq()-X
fBodyAccJerk-meanFreq()-Y
fBodyAccJerk-meanFreq()-Z
fBodyAccJerk-std()-X
fBodyAccJerk-std()-Y
fBodyAccJerk-std()-Z
fBodyAccMag-mean()
fBodyAccMag-meanFreq()
fBodyAccMag-std()
fBodyBodyAccJerkMag-mean()
fBodyBodyAccJerkMag-meanFreq()
fBodyBodyAccJerkMag-std()
fBodyBodyGyroJerkMag-mean()
fBodyBodyGyroJerkMag-meanFreq()
fBodyBodyGyroJerkMag-std()
fBodyBodyGyroMag-mean()
fBodyBodyGyroMag-meanFreq()
fBodyBodyGyroMag-std()
fBodyGyro-mean()-X
fBodyGyro-mean()-Y
fBodyGyro-mean()-Z
fBodyGyro-meanFreq()-X
fBodyGyro-meanFreq()-Y
fBodyGyro-meanFreq()-Z
fBodyGyro-std()-X
fBodyGyro-std()-Y
fBodyGyro-std()-Z
tBodyAcc-mean()-X
tBodyAcc-mean()-Y
tBodyAcc-mean()-Z
tBodyAcc-std()-X
tBodyAcc-std()-Y
tBodyAcc-std()-Z
tBodyAccJerk-mean()-X
tBodyAccJerk-mean()-Y
tBodyAccJerk-mean()-Z
tBodyAccJerk-std()-X
tBodyAccJerk-std()-Y
tBodyAccJerk-std()-Z
tBodyAccJerkMag-mean()
tBodyAccJerkMag-std()
tBodyAccMag-mean()

tBodyAccMag-std()
tBodyGyro-mean()-X
tBodyGyro-mean()-Y
tBodyGyro-mean()-Z
tBodyGyro-std()-X
tBodyGyro-std()-Y
tBodyGyro-std()-Z
tBodyGyroJerk-mean()-X
tBodyGyroJerk-mean()-Y
tBodyGyroJerk-mean()-Z
tBodyGyroJerk-std()-X
tBodyGyroJerk-std()-Y
tBodyGyroJerk-std()-Z
tBodyGyroJerkMag-mean()
tBodyGyroJerkMag-std()
tBodyGyroMag-mean()
tBodyGyroMag-std()
tGravityAcc-mean()-X
tGravityAcc-mean()-Y
tGravityAcc-mean()-Z
tGravityAcc-std()-X
tGravityAcc-std()-Y
tGravityAcc-std()-Z
tGravityAccMag-mean()
tGravityAccMag-std()

The features selected for this database come from the accelerometer and gyroscope 3-axial raw signals tAcc-XYZ and tGyro-XYZ. These time domain signals (prefix 't' to denote time) were captured at a constant rate of 50 Hz. Then they were filtered using a median filter and a 3rd order low pass Butterworth filter with a corner frequency of 20 Hz to remove noise. Similarly, the acceleration signal was then separated into body and gravity acceleration signals (tBodyAcc-XYZ and tGravityAcc-XYZ) using another low pass Butterworth filter with a corner frequency of 0.3 Hz.

Subsequently, the body linear acceleration and angular velocity were derived in time to obtain Jerk signals (tBodyAccJerk-XYZ and tBodyGyroJerk-XYZ). Also the magnitude of these three-dimensional signals were calculated using the Euclidean norm (tBodyAccMag, tGravityAccMag, tBodyAccJerkMag, tBodyGyroMag, tBodyGyroJerkMag).

Finally a Fast Fourier Transform (FFT) was applied to some of these signals producing fBodyAcc-XYZ, fBodyAccJerk-XYZ, fBodyGyro-XYZ, fBodyAccJerkMag, fBodyGyroMag, fBodyGyroJerkMag. (Note the 'f' to indicate frequency domain signals).

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.

The keywords *mean* and *std* denote the mean and standard deviation of the correspondent variables in the original dataset described above.

value_avg:

TYPE: “NUMERIC”

These values represent the average mean and average standard deviation for each activity type and subject ID.