**CodeBook for the tidy dataset**

**Data source**

This dataset is derived from the "Human Activity Recognition Using Smartphones Data Set" which was originally made avaiable here: <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

**Feature Selection**

I refer you to the README and features.txt files in the original dataset to learn more about the feature selection for this dataset. And there you will find the follow description:

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).

The reasoning behind my selection of features is that the assignment explicitly states "Extracts only the measurements on the mean and standard deviation for each measurement." To be complete, I included all variables having to do with mean or standard deviation.

In short, for this derived dataset, 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 (and kept for this assignment) from these signals are:

* mean(): Mean value
* std(): Standard deviation

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

* gravityMean
* tBodyAccMean
* tBodyAccJerkMean
* tBodyGyroMean
* tBodyGyroJerkMean

Other estimates have been removed for the purpose of this excercise.

Note: features are normalized and bounded within [-1,1].

The resulting variable names are of the following form: tbodyaccmeanx, which means the mean value of tBodyAcc-XYZ.