Code Book

Each row of total_df is a 563 variables observation, the first two variables being the subject_ID and the act_label:

- subject_ID: Identifies with a number from 1 to 30 the subject/volunteer who performed the activity whose measurements are recorded in the same row.
- act_label: Identifies with a literal label the type of activity that the subject was performing when the measurements were made. Levels: WALKING, WALKING_UPSATAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, LAYING.

These are the only two variables that are common to all four data frames produced by the script.

The names of the rest of the variables begin with one of the following prefixes:

- 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

where:

- "t" indicates that it's a time domain variable and "f" indicates it's a frequency domain variable.
- "body" indicates translational acceleration (in units of g) and "gyro" indicates angular speed (radians/second). Note: there are exceptions of variables with the "body" or "gyro" infixes that have different units from the mentioned before.
- "Jerk" indicates the derivative of acceleration (units of g per second).
- "Gravity" indicates the gravity signal that was separated from the "body" acceleration using a low-pass filter.
- X, Y or Z indicates the axis of the phone along which the measurements were made.

The aforementioned variable names end with one of the following suffixes:

- 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.

NOTE: When running the script, the "()" are systematically replaced by ".." for formatting reasons.

Examples:

- fBodyAccMag-kurtosis() refers to the kurtosis of the FFT of the magnitude of the body acceleration. Since Kurtosis is calculated using standarized forms of the distribution, this variable is unitless. The same thing happens with the skewness() variables.
- tBodyAccMag-energy() has units of the square of body acceleration.

IMPORTANT: In the means_agg and means_agg_by_subject data frames the variables have the same names mentioned above but a slightly different meaning. They are the means of the values of total_df and mean_std by subject and by act_label.