

Tidy Data - Code Book

This data set was created from tidying up and summarizing the "Human Activity Recognition Using Smartphones Data Set" so that, for each subject and each activity performed the average value of the "mean()", "std()" and "meanFreq()" variables is calculated. Each variable's name is prefixed with "Avg-" in order to signify this calculation.

Original Dataset:

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

The identity variables:

Subject:	Ordinal	
	1..30	.The subject number
Activity:	Categorical	
	WALKING	.Subject walking
	WALKING_UPSTAIRS	.Subject walking up a stairs
	WALKING_DOWNSTAIRS	.Subject walking down a stairs
	SITTING	.Subject sitting down
	STANDING	.Subject standing up
	LAYING	.Subject laying down on his back

The measurement variables: Continuous

These variables are named using the following format:

Avg-{type of measurement}-{type of statistic}-{measurement axis}

The {type of measurment} signifies the measurement type of the initial data as documented in the "features_info.txt" of the original data set.

The {type of statistic} signifies which statistic was used in the initial data as documented in the "features_info.txt" of the original data set.

The {measurement axis} signifies the axis of measurement.

Summary choices:

The value of each variable is the average value of the corresponding measurement for the particular subject performing the particular activity shown in the first two columns of each row. The average was calculated using the mean() function. The corresponding measurement for each variable is give by the varialbe's name without the prefix "Avg-". For example, "Avg-tBodyAcc-mean()-X" is the average value of the "tBodyAcc-mean()-X" measurement of the initial data.