

run_analysis_codebook

This codebook builds upon details from within <https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip>

The following files are referenced:

features_info.txt, activity_labels.txt, README.txt

=====
subject: subject numbers associated with Samsung data. There are 30 subjects who participated in the study

Storage mode: integer

=====
activity: These are the 6 activities described in the study, and labeled in activity_labels.txt

Storage mode: integer Factor with 6 levels

Values and labels	N	Percent
1 'LAYING'	1980	16.7
2 'SITTING'	1980	16.7
3 'STANDING'	1980	16.7
4 'WALKING'	1980	16.7
5 'WALKING_DOWNSTAIRS'	1980	16.7
6 'WALKING_UPSTAIRS'	1980	16.7

=====
domain: Within the features, variables had been concatenated to include the domain (frequency or time). This variable splits out the domain in order to better understand time data vs frequency data

Storage mode: character

Length: 11880
Class: character
Mode: character

=====
signal: within the features list, signal is the specific set of variables looked at in features_info.txt, they are described as follows. This column is the result of pulling out domain (t|f), and axis (-XYZ) so that they may be looked at independently

Storage mode: character

Length: 11880
Class: character
Mode: character

=====

calctype: in the original dataset, a number of calculations had been performed and reported. This dataset includes only mean() and std() calculations

Storage mode: character

Length: 11880
Class: character
Mode: character

=====

axis: within the features list, signal is the specific set of variables. A number of signals include axis information (-XYZ). Data in this column includes on which axis data was collected. For some signals, axis is not applicable, and is reported as such

Storage mode: character

Length: 11880
Class: character
Mode: character

=====

average: The mean for all observations/calculations for this group of data