**CodeBook**

This is a code book that describes the variables, data, and any transformations or work performed to clean up the data.

**The data source**

* Original data are found at: <https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip>
* Original description of the dataset: <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

**Data Set Information**

The experiments have been carried out with a group of 30 volunteers within an age bracket of 19-48 years. Each person performed six activities (WALKING, WALKING\_UPSTAIRS, WALKING\_DOWNSTAIRS, SITTING, STANDING, LAYING) wearing a smartphone (Samsung Galaxy S II) on the waist. Using its embedded accelerometer and gyroscope, we captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The experiments have been video-recorded to label the data manually. The obtained dataset has been randomly partitioned into two sets, where 70% of the volunteers was selected for generating the training data and 30% the test data.

The sensor signals (accelerometer and gyroscope) were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 readings/window). The sensor acceleration signal, which has gravitational and body motion components, was separated using a Butterworth low-pass filter into body acceleration and gravity. The gravitational force is assumed to have only low frequency components, therefore a filter with 0.3 Hz cutoff frequency was used. From each window, a vector of features was obtained by calculating variables from the time and frequency domain.

**The data**

The dataset includes the following files:

* 'README.txt'
* 'features\_info.txt': Shows information about variables used on feature vector.
* 'features.txt': List of all features.
* 'activity\_labels.txt': Links class labels with their activity name.
* 'train/X\_train.txt': Training set.
* 'train/y\_train.txt': Training labels.
* 'test/X\_test.txt': Test set.
* 'test/y\_test.txt': Test labels.

The following files are available for the train and test data. Their descriptions are equivalent.

* 'train/subject\_train.txt': Each row identifies the subject who performed the activity for each window sample. Its range is from 1 to 30.
* 'train/Inertial Signals/total\_acc\_x\_train.txt': The acceleration signal from the smartphone accelerometer X axis in standard gravity units 'g'. Every row shows a 128 element vector. The same description applies for the 'total\_acc\_x\_train.txt' and 'total\_acc\_z\_train.txt' files for the Y and Z axis.
* 'train/Inertial Signals/body\_acc\_x\_train.txt': The body acceleration signal obtained by subtracting the gravity from the total acceleration.
* 'train/Inertial Signals/body\_gyro\_x\_train.txt': The angular velocity vector measured by the gyroscope for each window sample. The units are radians/second.

**Transformation details**

There are 5 parts:

1. Merges the training and the test sets to create one data set.
2. Extracts only the measurements on the mean and standard deviation for each measurement.
3. Uses descriptive activity names to name the activities in the data set
4. Appropriately labels the data set with descriptive activity names.
5. Creates a second, independent tidy data set with the average of each variable for each activity and each subject.

**How run\_analysis.R implements the above steps:**

* Download data from url and unzip data
* Require reshapre2 and data.table librareis.
* Load both test and train data
* Load the features and activity labels.
* Extract the mean and standard deviation column names and data.
* Process the data. There are two parts processing test and train data respectively.
* Merge data set.