**UCI Human Activity Recognition Code Book**

**Data**

This code pulls data from the UCI Machine Learning Repository on Human Activity Recognition software in the Samsung Galaxy S smartphone. The data measures 30 volunteer subjects doing 6 daily activities: walking, walking downstairs, walking upstairs, standing, laying down, and sitting. [[1]](http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones)

**Process**

The code executes the following steps:

1. Uses cbind to merge the “train” and “test” datasets, adding column headers and subject and activity identifiers to the larger dataset from the “features” and “activity labels” files.
2. Extract only the observations on the mean and standard deviation for each variable by selecting only variables with names that contain “std” or “mean”.
3. Name the six activities identified in the dataset, matching the numeric identifier in the larger dataset to the activity names in the “activity labels” files.
4. Labels the data with descriptive names by replacing any non-descriptive variable name components with descriptive ones.
   1. Replaces "acc" with "acceleration”
   2. Replaces "^t" with "time"
   3. Replaces “^f" with "frequency"
   4. Replaces "bodybody" with "body"
   5. Replaces "mag" with "magnitude”
5. Creates an independent tidy dataset with the average of each variable for each activity and subject.
   1. The tidy dataset groups subject identifier and activity, resulting in 180 observations of 89 variables (so there is one mean for each of the 30 subjects and their 6 activities each).