Source of the original data:  <https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip>

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

The attached R script (run\_analysis.R) performs the following steps to clean up the data and produce a tidy data set:

* Merges the training and test sets (train/X\_train.txt with test/X\_test.txt) to create one data set, resulting in a 10299x561 data frame, as in the original description ("Number of Instances: 10299" and "Number of Attributes: 561"). Merges the training and test sets train/subject\_train.txt with test/subject\_test.txt, resulting in a 10299x1 data frame with subject IDs, and train/y\_train.txt with test/y\_test.txt, resulting in a 10299x1 data frame with activity IDs.
* Reads the file features.txt and extracts only the measurements on the mean and standard deviation for each measurement. The result is a 10299x66 data frame, because only 66 out of 561 attributes are measurements on the mean and standard deviation. All measurements appear to be floating point numbers in the range (-1, 1).
* Reads the file activity\_labels.txt and applies descriptive activity names to name the activities in the data set: walking, walkingupstairs, walkingdownstairs, sitting, standing, laying.
* The script also appropriately labels the data set with descriptive names: all feature names (attributes) and activity names are converted to lower case, underscores and brackets () are removed. The names of the attributes are similar to the following: tbodyacc-mean-x, tbodyacc-mean-y, tbodyacc-mean-z, tbodyacc-std-x, tbodyacc-std-y, tbodyacc-std-z, tgravityacc-mean-x, tgravityacc-mean-y
* Then it merges the 10299x66 data frame containing features with 10299x1 data frames containing activity labels and subject IDs. The result is a 10299x68 data frame such that the first column contains subject IDs, the second column activity names, and the last 66 columns are measurements. Subject IDs are integers between 1 and 30 inclusive.
* Finally, the script creates a 2nd, independent tidy data set with the average of each measurement for each activity and each subject. The result is saved as uci\_char\_tidy\_data\_set.txt, a 181x68 data frame, where, the first column contains subject IDs, the second column contains activity names (see below), and then the averages for each of the 66 attributes are in columns 3...68. There are 30 subjects and 6 activities, thus 180 rows in this data set with averages. The first row of the text file contains the variable names resulting in a total of 181 rows in the text file.