==========

Human Activity Recognition Using Smartphones Dataset

Author: Gabriel Ernesto Amodeo

==========

Files:

-'Dataset': UCI HAR dataset

-'Script': Contains the R script and a subfolder (data) with the required data in .txt form

'Tableau workbook': Contains the required workbook file.

‘readme’: Contains information on the script, dataset and workbook.

==========

Requirements:

-R . Version 4.0.5

-RStudio (highly recommended)

-Tableau desktop Version 2011.1.1

*When loading the R script remember to set the working directory to the script location through Session > Set working directory > To source file location. (in RStudio).

==========

Introduction:

The provided dataset was originally intended for machine learning, and because of this it is divided into two sets, a training set, which is used to "teach" the machine-learning algorithm how to recognize a pattern (in this case, the prediction would be to determine what activity the person is doing (WALKING, STANDING, LYING, WALKING UPSTAIRS, WALKING DOWNSTAIRS SITTING)), and a test set, which is used to provide an unbiased evaluation of a *final* model fit on the training dataset.

To analyze the dataset, it was required that both individual sets of data were combined in a single set. The resulting set of data needed to be cleaned and merged with the respective labels for each activity. Finally, to plot the data it was required a time measurement for reference, but the dataset did not contain this information. The R script will load the data into data frames, merge them, clean them and finally add the time column required to advance with the analysis. The resulting data frame is loaded into a csv file (UCIdataset.csv), which is used in Tableau to create the dashboard provided.

==========

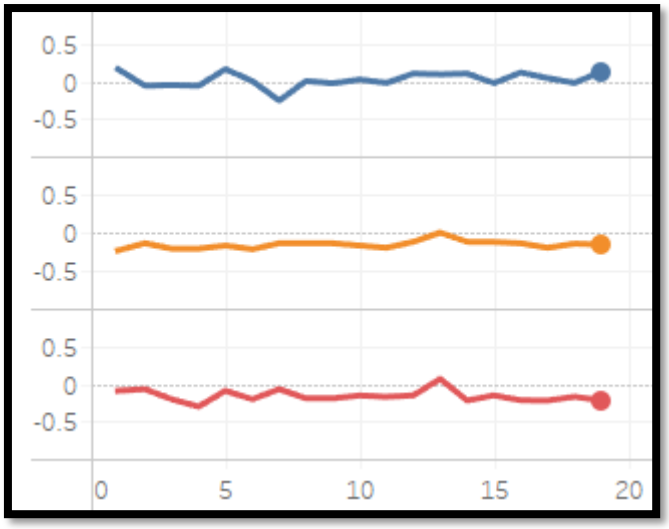
Instructions for the dashboard

*Note : For optimal usage of the dashboard it is recommended that “presentation” mode is enabled and that no more than 3 measurements be selected at a time.

Activity selection bar

The icons in the bar are clicklable and will filter all the other plots and tables with the corresponding activity.

Main plot



Represents the selected measure(s) in regards to the corresponding Subject(s).

Axis labels are as follows:

X-axis: Time [sec]

Y-axis:

Accelerometer [m/s^2]

Gyroscope[rad/sec]

The main line plot is useful to compare the values of the different measurements obtained from the data set. Multiple observations can be done, for instance if we look at subject's N°1 and N°2 BodyAcceleration on the Z axis, we can clearly see that subject N°2 had a greater de-acceleration while climbing stairs. This might tell us that the participant N° 2 might be older, and thus requires to stop on each step taken.

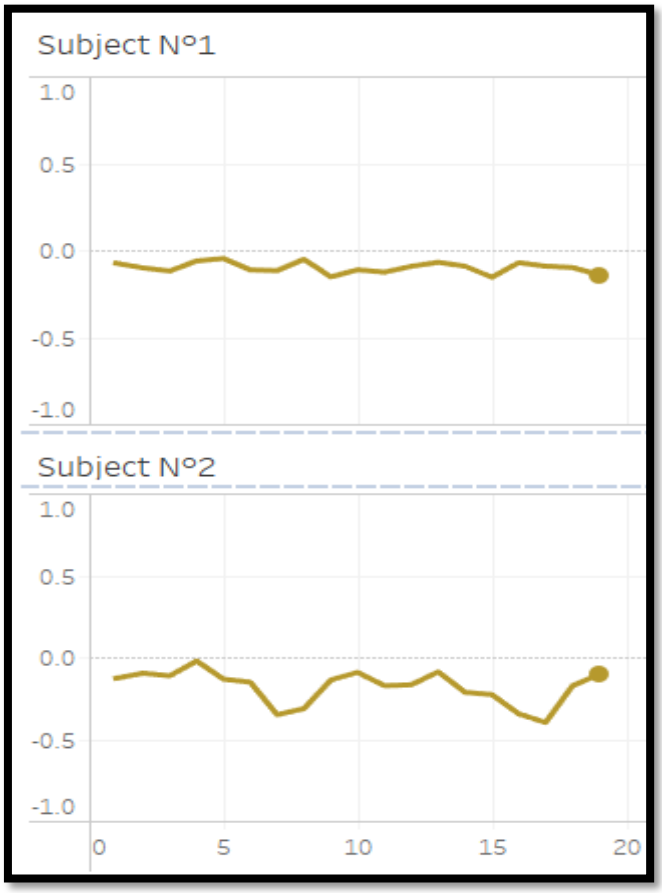


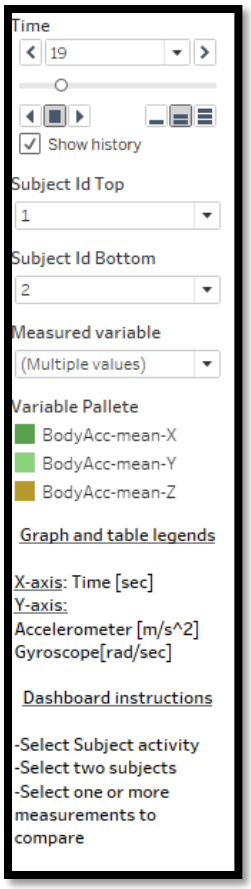
Fig 1. Subject 1 vs Subject 2

Min, Avg and Max table

	Max	Avg	Min
BodyAcc-mean-X	0.470	0.266	-0.361
BodyAcc-mean-Y	0.189	-0.018	-0.684
BodyAcc-mean-Z	0.347	-0.108	-1.000

This table shows the maximum, average and minimum values for the selected measurements for each of participant.

Navigation bar



The navigation bar contains:

-The time controls for the graphs.

-Subject single selection dropdown menu

-Measurement multiple selection dropdown menu

-Graph and table legends.

Considerations: In order to analyze and later plot the required information it was necessary to remove or modify the following items:

signals with Fast Fourier Transformations, prefix 'f' (we are more interested in the time domain signals, prefix 't') are removed

fBodyAcc-XYZ

fBodyAccJerk-XYZ

fBodyGyro-XYZ

fBodyAccMag

fBodyAccJerkMag

fBodyGyroMag

fBodyGyroJerkMag

Magnitude of signals calculated with Euclidean form are removed:

tBodyAccMag

tGravityAccMag

tBodyAccJerkMag

tBodyGyroMag

tBodyGyroJerkMag

For the simplification of this analysis, only mean, std, max and min variables will be kept, discarding the following:

mad(): Median absolute deviation

sma(): Signal magnitude area

energy(): Energy measure. Sum of the squares divided by the number of values.

iqr(): Interquartile range

entropy(): Signal entropy

arCoeff(): Autorregresion coefficients with Burg order equal to 4

correlation(): correlation coefficient between two signals

maxInds(): index of the frequency component with largest magnitude

meanFreq(): Weighted average of the frequency components to obtain a mean frequency

skewness(): skewness of the frequency domain signal

kurtosis(): kurtosis of the frequency domain signal

bandsEnergy(): Energy of a frequency interval within the 64 bins of the FFT of each window.

angle(): Angle between to vectors.

This leaves the following signals and variables up for analysis:

#signals

tBodyAcc-XYZ

tGravityAcc-XYZ

tBodyAccJerk-XYZ

tBodyGyro-XYZ

tBodyGyroJerk-XYZ

#variations

mean()

std()

max()

min()

This results in a dataset of 10299 data points with 62 columns (the first two being the subject ID and the activity performed), while the rest is a combination of the #signal (with the respective XYZ axis) and the #variations.

Introduction on the data set:

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. See 'features_info.txt' for more details.

For each record it is provided:

==========

- Triaxial acceleration from the accelerometer (total acceleration) and the estimated body acceleration.

- Triaxial Angular velocity from the gyroscope.

- A 561-feature vector with time and frequency domain variables.

- Its activity label.

- An identifier of the subject who carried out the experiment.

The dataset includes the following files:

==========

- 'README.txt'

- 'features_info.txt': Shows information about the variables used on the feature vector.

- 'features.txt': List of all features.

- 'activity_labels.txt': Links the 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.

Notes:

======

- Features are normalized and bounded within [-1,1].

- Each feature vector is a row on the text file.

- The units used for the accelerations (total and body) are 'g's (gravity of earth -> 9.80665 m/seg2).

- The gyroscope units are rad/seg.

- A video of the experiment including an example of the 6 recorded activities with one of the participants can be seen in the following link: http://www.youtube.com/watch?v=XOEN9W05_4A