# Codebook

## Codebook - run_analysis.r script

The `run_analysis.r` script collects and cleans a data set as described int the Getting and Cleaning Data project

### Preparation

- Downloading required packages: `dplyr`, `data.table`
- Downloading and unzipping dataset
- Assining each data to variables
  - `features` <- 'features.txt' The features selected for this database come from accelerometer and gyroscope 3-axial raw signals tAcc-XYZ and tGyro-XYZ
  - `activity` <- `activity_labels.txt` Links the class labels (code) with their activity name.
  - `subject_test` <- `subject_test.txt` contains test data of volunteer test subjects beeing observed
  - `X_test` <- `X_test.txt` contains recorded feautres data set
  - `Y_test` <- `Y_test.txt` contains test data of activities code labels
  - `subject_train` <- `subject_train.txt` contains train data of volunteer subject being observed
  - `X_train` <- `X_train.txt` contains recorded features train data
  - `Y_train` <- `Y_train.txt` contains train data of activities code labes -Looking at the properties of the assigned data frames

### 1. Merges the training and the test sets to create one data set.

- `merged_X` <- binding the rows of `X_test` and `X_train`
- `merged_Y` <- binding the rows of `Y_test` and `Y_train`
- `merged_subject` <- binding the rows of `subject_test` and `subject_train`
- `data` <- binding all data into one dataset `merged_X`, `merged_Y`, `merged_subject`

### 2. Extracts only the measurements on the mean and standard deviation for each measurement.

`tidyData` <- selecting only the measurements with mean and standard deviation from `data`

### 3. Uses descriptive activity names to name the activities in the data set

Combines the names from the activities from the `activity` data set with the class labels from `tidyData$code`

### 4. Appropriately labels the data set with descriptive variable names.

- `code` column in the `tidyData` renamed to `activity`
- All `Acc` in column names replaced by `Accelerometer`
- All `Gyro` in column names replaced by `Gyroscope`
- All `t` in the the begining of column names replaced by `time`

- All `BodyBody in column names replaced by`Body‘
- All `Mag` in column names replaced by `Frequency`
- All `f` in the the begining of column names replaced by `time`
- All `mean` in column names replaced by `Mean`
- All `STD` in column names replaced by `STD`
- All `freq` in column names replaced by `Frequency`
- All `angle` in column names replaced by `Angle`

**5. From the data set in step 4, creates a second, independent tidy data set with the average of each variable for each activity and each subject.**

- `FinalData` <- final data set, created by grouping `tidyData` by `subject` and `activity` and calculatin the `mean`
- Viewing info about `FinalData`