**Preparer:** Fazleali Pirbhai

**Assignment:** Getting and cleaning data – Coursera – John Hopkins University

**Project Name:** Human Activity Recognition Using Smartphones Data Set

**Data**

The files uploaded in github for this assignment are: -

1. README.md
2. Codebook.md
3. Run\_analysis.R

For this assignment, the data was downloaded from <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>.

**Steps used for this assignments**

1. Downloaded the original dataset from <https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip>
2. Load the following datasets in R: -

|  |  |
| --- | --- |
| **Data Frame** | **Original File Name** |
| activity\_labels\_raw | activity\_labels.txt |
| features\_raw | features.txt |
| x\_test\_raw | X\_test.txt |
| y\_test\_raw | y\_test.txt |
| Subject\_test\_raw | subject\_test.txt |
| x\_train\_raw | X\_train.txt |
| y\_train\_raw | y\_train.txt |
| Subject\_train\_raw | subject\_train.txt |

1. Merge the training and test data tables into one, namely ‘x\_test\_train’. This table has 10,299 observations and 561 variables.
2. Extracted only measurements with either ‘mean’ or ‘std’ and saved to a new table ‘mean\_std’
3. Inserted the following variables in the above data frame: ‘subject’, ‘type\_act’ and ‘activity\_description’.
4. Then we labels the remaining variables.
5. Finally, we summarize the table by grouping the table by subject,Type\_Activity and activity description. For each of this group, we provide the mean for each variable.