

output: html\_document

[illegible]

## Data Code Book

# Getting and Cleaning Data Course Final Project

Tarig Elamin      Winter 2015

This code book –as a good habit for a data scientist to have when working on an analytical project by some sort- tries to describe the variables/columns, data, and transformations done to in order to clean up the data and make it ready to next stage of the analytical project.

The data was obtained from:

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

The data used for the project:

<https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip>

What have done is:

The following transformations were done on the origin dataset described above to create the tidy dataset.

1. downloaded the data and a local directory created—see the R program accompanied-and unzipped and then we read the data

1. The test and train data were combined together
2. Columns with `mean()` and `std()` were retained. All other columns were discarded.
3. The `mean()` and `std()` columns were averaged over all participants and activities.
4. This averaging was accomplished using the `melt` and `dcast` functions from the `reshape2` library.

The resulting tidy dataset contains one row for each unique combination of participant and activity.