Codebook for Getting and Cleaning Data Course Project

#09-21-2014

#Susan Lupiani, [susandata2014@gmail.com](mailto:susandata2014@gmail.com)

The data was downloaded from the Internet and unzipped. The following files were loaded into R: Activity\_Labels, Features, Subject\_train, Subject\_test, x\_train, x\_test, y\_train, y\_test. The Features table is going to become the column headings, so at this point the text was cleaned up to remove “()” and “-“ from the names. Once completed, the column headings from the Features table were added to the x\_test and x\_train tables. Column names “ActivityID” and “Activity” were added to y\_test and y\_train and “Subject” was added to Subject\_train and Subject\_test. Activity\_Labels were then merged with the y\_test and y\_train tables. Then one table was created that contained all the data.

Once a table was created with all the data, a second table was created that only contained columns for means and standard deviations.

Finally, a table was created that contained the means of every column, grouped by Subject and Activity.