CodeBook

December 26, 2015

- Code Book for "tidy.txt"

The tidy.txt containsts of 180 records each with 81 variables. Each record is the results of multi-step process that is explained below. These 180 observations are related to 30 subjects performing 6 main activities (30 * 6 = 180). For each subject, there are 79 columns.

Database Structure

```r
str(tidy_data)
```

```
## 'data.frame':    180 obs. of  81 variables:
##  $ Activity                 : int  1 1 1 1 1 1 2 2 2 2 ...
##  $ Subject                  : Factor w/ 6 levels "LAYING","SITTING",..: 1 2 3 4 5 6 1 2 3 4
##  $ tBodyAcc.mean.X_Avg      : num  0.222 0.261 0.279 0.277 0.289 ...
##  $ tBodyAcc.mean.Y_Avg      : num  -0.04051 -0.00131 -0.01614 -0.01738 -0.00992 ...
##  $ tBodyAcc.mean.Z_Avg      : num  -0.113 -0.105 -0.111 -0.111 -0.108 ...
##  $ tBodyAcc.std.X_Avg       : num  -0.928 -0.977 -0.996 -0.284 0.03 ...
##  $ tBodyAcc.std.Y_Avg       : num  -0.00933 -0.20119 -0.00543 -0.19123 -0.22554 ...
##  $ tBodyAcc.std.Z_Avg       : num  -0.0234 0.1183 0.2883 0.3802 0.0757 ...
##  $ tGravityAcc.mean.X_Avg   : num  -0.249 0.832 0.943 0.935 0.932 ...
##  $ tGravityAcc.mean.Y_Avg   : num  0.706 0.204 -0.273 -0.282 -0.267 ...
##  $ tGravityAcc.mean.Z_Avg   : num  0.4458 0.332 0.0135 -0.0681 -0.0621 ...
##  $ tGravityAcc.std.X_Avg    : num  -0.897 -0.968 -0.994 -0.977 -0.951 ...
##  $ tGravityAcc.std.Y_Avg    : num  -0.0781 -0.4975 0.1378 -0.0401 -0.2974 ...
##  $ tGravityAcc.std.Z_Avg    : num  -0.209 0.301 0.47 0.281 0.101 ...
##  $ tBodyAccJerk.mean.X_Avg  : num  0.0811 0.0775 0.0754 0.074 0.0542 ...
##  $ tBodyAccJerk.mean.Y_Avg  : num  0.003838 -0.000619 0.007976 0.028272 0.02965 ...
##  $ tBodyAccJerk.mean.Z_Avg  : num  0.01083 -0.00337 -0.00369 -0.00417 -0.01097 ...
##  $ tBodyAccJerk.std.X_Avg   : num  -0.9585 -0.9864 -0.9946 -0.1136 -0.0123 ...
##  $ tBodyAccJerk.std.Y_Avg   : num  0.146 0.16 0.156 -0.362 -0.358 ...
##  $ tBodyAccJerk.std.Z_Avg   : num  -0.022 0.0166 0.0744 0.4906 0.116 ...
##  $ tBodyGyro.mean.X_Avg     : num  -0.0166 -0.0454 -0.024 -0.0418 -0.0351 ...
##  $ tBodyGyro.mean.Y_Avg     : num  -0.0645 -0.0919 -0.0594 -0.0695 -0.0909 ...
##  $ tBodyGyro.mean.Z_Avg     : num  0.1487 0.0629 0.0748 0.0849 0.0901 ...
##  $ tBodyGyro.std.X_Avg      : num  -0.874 -0.977 -0.987 -0.474 -0.458 ...
##  $ tBodyGyro.std.Y_Avg      : num  0.25254 0.15251 -0.00513 -0.06303 -0.19103 ...
##  $ tBodyGyro.std.Z_Avg      : num  0.044 -0.233 -0.46 0.108 -0.28 ...
##  $ tBodyGyroJerk.mean.X_Avg : num  -0.1073 -0.0937 -0.0996 -0.09 -0.074 ...
##  $ tBodyGyroJerk.mean.Y_Avg : num  -0.0415 -0.0402 -0.0441 -0.0398 -0.044 ...
##  $ tBodyGyroJerk.mean.Z_Avg : num  -0.0741 -0.0467 -0.049 -0.0461 -0.027 ...
##  $ tBodyGyroJerk.std.X_Avg  : num  -0.919 -0.992 -0.993 -0.207 -0.487 ...
##  $ tBodyGyroJerk.std.Y_Avg  : num  -0.0388 0.1177 -0.0375 -0.3346 -0.2415 ...
##  $ tBodyGyroJerk.std.Z_Avg  : num  -0.09497 0.00285 -0.22817 -0.07155 -0.25142 ...
##  $ tBodyAccMag.mean_Avg     : num  0.0352 0.1532 0.2403 0.0886 0.1391 ...
##  $ tBodyAccMag.std_Avg      : num  -0.0242 -0.1716 -0.2481 0.0818 -0.1533 ...
##  $ tGravityAccMag.mean_Avg  : num  0.0352 0.1532 0.2403 0.0886 0.1391 ...
```

```
##  $ tGravityAccMag.std_Avg            : num  -0.0242 -0.1716 -0.2481 0.0818 -0.1533 ...
##  $ tBodyAccJerkMag.mean_Avg          : num  -0.1116 -0.0308 -0.055 -0.2249 -0.0169 ...
##  $ tBodyAccJerkMag.std_Avg           : num  -0.079 -0.0603 -0.1282 0.2042 0.111 ...
##  $ tBodyGyroMag.mean_Avg             : num  -0.00942 0.16665 0.16383 0.16574 0.42346 ...
##  $ tBodyGyroMag.std_Avg              : num  0.0864 -0.113 -0.0367 -0.1898 -0.2959 ...
##  $ tBodyGyroJerkMag.mean_Avg         : num  -0.03971 -0.00118 0.00219 -0.11631 -0.058 ...
##  $ tBodyGyroJerkMag.std_Avg          : num  -0.11108 -0.00177 -0.28578 -0.18747 0.05727 ...
##  $ fBodyAcc.mean.X_Avg               : num  -0.9391 -0.9796 -0.9952 -0.2028 0.0382 ...
##  $ fBodyAcc.mean.Y_Avg               : num  -0.86707 -0.94408 -0.97707 0.08971 0.00155 ...
##  $ fBodyAcc.mean.Z_Avg               : num  -0.883 -0.959 -0.985 -0.332 -0.226 ...
##  $ fBodyAcc.std.X_Avg                : num  -0.9244 -0.9764 -0.996 -0.3191 0.0243 ...
##  $ fBodyAcc.std.Y_Avg                : num  -0.817 -0.972 -0.909 -0.638 -0.682 ...
##  $ fBodyAcc.std.Z_Avg                : num  -0.854 -0.899 -0.882 -0.92 -0.925 ...
##  $ fBodyAcc.meanFreq.X_Avg           : num  -0.1588 -0.0495 0.0865 -0.2075 -0.3074 ...
##  $ fBodyAcc.meanFreq.Y_Avg           : num  -0.942 -0.988 -0.999 -0.725 -0.718 ...
##  $ fBodyAcc.meanFreq.Z_Avg           : num  -0.996 -0.999 -1 -0.918 -0.81 ...
##  $ fBodyAccJerk.mean.X_Avg           : num  -0.9571 -0.9866 -0.9946 -0.1705 -0.0277 ...
##  $ fBodyAccJerk.mean.Y_Avg           : num  -0.9225 -0.9816 -0.9854 -0.0352 -0.1287 ...
##  $ fBodyAccJerk.mean.Z_Avg           : num  -0.948 -0.986 -0.991 -0.469 -0.288 ...
##  $ fBodyAccJerk.std.X_Avg            : num  -0.9642 -0.9875 -0.9951 -0.1336 -0.0863 ...
##  $ fBodyAccJerk.std.Y_Avg            : num  -0.259 -0.243 -0.32 -0.459 -0.402 ...
##  $ fBodyAccJerk.std.Z_Avg            : num  -0.311 -0.221 -0.121 -0.247 -0.167 ...
##  $ fBodyAccJerk.meanFreq.X_Avg       : num  0.132 0.257 0.314 -0.209 -0.253 ...
##  $ fBodyAccJerk.meanFreq.Y_Avg       : num  -0.992 -0.999 -1 -0.814 -0.727 ...
##  $ fBodyAccJerk.meanFreq.Z_Avg       : num  -0.997 -1 -1 -0.923 -0.822 ...
##  $ fBodyGyro.mean.X_Avg              : num  -0.85 -0.976 -0.986 -0.339 -0.352 ...
##  $ fBodyGyro.mean.Y_Avg              : num  -0.9522 -0.9758 -0.989 -0.1031 -0.0557 ...
##  $ fBodyGyro.mean.Z_Avg              : num  -0.9093 -0.9513 -0.9808 -0.2559 -0.0319 ...
##  $ fBodyGyro.std.X_Avg               : num  -0.882 -0.978 -0.987 -0.517 -0.495 ...
##  $ fBodyGyro.std.Y_Avg               : num  -0.902 -0.915 -0.886 -0.779 -0.856 ...
##  $ fBodyGyro.std.Z_Avg               : num  -0.97 -0.991 -0.88 -0.634 -0.752 ...
##  $ fBodyGyro.meanFreq.X_Avg          : num  -0.00355 0.18915 -0.12029 0.01478 -0.10045 ...
##  $ fBodyGyro.meanFreq.Y_Avg          : num  -0.986 -0.993 -0.999 -0.79 -0.622 ...
##  $ fBodyGyro.meanFreq.Z_Avg          : num  -0.998 -1 -1 -0.919 -0.884 ...
##  $ fBodyAccMag.mean_Avg              : num  -0.644 -0.836 -0.915 0.544 0.653 ...
##  $ fBodyAccMag.std_Avg               : num  -0.266 -0.374 -0.474 -0.549 -0.457 ...
##  $ fBodyAccMag.meanFreq_Avg          : num  -0.508 -0.608 -0.714 -0.835 -0.739 ...
##  $ fBodyBodyAccJerkMag.mean_Avg      : num  -0.827 -0.96 -0.997 0.478 0.525 ...
##  $ fBodyBodyAccJerkMag.std_Avg       : num  -0.389 -0.481 -0.568 -0.104 -0.197 ...
##  $ fBodyBodyAccJerkMag.meanFreq_Avg  : num  -0.674 -0.751 -0.833 -0.478 -0.539 ...
##  $ fBodyBodyGyroMag.mean_Avg         : num  -0.516 -0.709 -0.768 0.683 0.68 ...
##  $ fBodyBodyGyroMag.std_Avg          : num  -0.08434 -0.00776 -0.20117 -0.24686 -0.494 ...
##  $ fBodyBodyGyroMag.meanFreq_Avg     : num  -0.392 -0.284 -0.514 -0.541 -0.772 ...
##  $ fBodyBodyGyroJerkMag.mean_Avg     : num  -0.721 -0.937 -0.957 0.574 0.587 ...
##  $ fBodyBodyGyroJerkMag.std_Avg      : num  -0.5203 -0.0605 0.2702 0.277 0.2672 ...
##  $ fBodyBodyGyroJerkMag.meanFreq_Avg: num  -0.3524 -0.218 0.0123 0.0689 0.065 ...
```

Data Source:

This data is the result of processing raw data collected from *UCI HAR DATASET* dataset, In this dataset, there exist 10299 observations related to datacollections related to each of the 30 subjects performing 6 main activities.

Using a sensory device, sensed data of each subject are collected and recorded in text files. The dataset is delivered as a chunk of test and train dataset (X_test.txt and X_train.txt).

In the following section, we provide further information about dataset and columns of the dataset.

1. *"Subject"*: Is the ID of the 30 subjects participating in the study whose data reflected in this dataset. The input for this colum is read from 'subject_test.txt' and 'subject_train.txt'.

2. *"Activity"*: There are 6 main activities in this study that represents the subject's activity under study.

- 'WALKING'
- 'WALKING_UPSTAIRS'
- 'WALKING_DOWNSTAIRS'
- 'SITTING'
- 'STANDING'
- 'LAYING'.

For each subject and activity, there are 561 features. Out of 561 features, we select only those variables that have 'mean' or 'std' string in their names. The results are 79 variables listed below.

- tBodyAcc-mean-X
- tBodyAcc-mean-Y
- tBodyAcc-mean-Z
- tBodyAcc-std-X
- tBodyAcc-std-Y
- tBodyAcc-std-Z
- tGravityAcc-mean-X
- tGravityAcc-mean-Y
- tGravityAcc-mean-Z
- tGravityAcc-std-X
- tGravityAcc-std-Y
- tGravityAcc-std-Z
- tBodyAccJerk-mean-X
- tBodyAccJerk-mean-Y
- tBodyAccJerk-mean-Z
- tBodyAccJerk-std-X
- tBodyAccJerk-std-Y
- tBodyAccJerk-std-Z
- tBodyGyro-mean-X
- tBodyGyro-mean-Y
- tBodyGyro-mean-Z
- tBodyGyro-std-X
- tBodyGyro-std-Y
- tBodyGyro-std-Z
- tBodyGyroJerk-mean-X
- tBodyGyroJerk-mean-Y
- tBodyGyroJerk-mean-Z
- tBodyGyroJerk-std-X
- tBodyGyroJerk-std-Y
- tBodyGyroJerk-std-Z
- tBodyAccMag-mean

- tBodyAccMag-std
- tGravityAccMag-mean
- tGravityAccMag-std
- tBodyAccJerkMag-mean
- tBodyAccJerkMag-std
- tBodyGyroMag-mean
- tBodyGyroMag-std
- tBodyGyroJerkMag-mean
- tBodyGyroJerkMag-std
- fBodyAcc-mean-X
- fBodyAcc-mean-Y
- fBodyAcc-mean-Z
- fBodyAcc-std-X
- fBodyAcc-std-Y
- fBodyAcc-std-Z
- fBodyAcc-meanFreq-X
- fBodyAcc-meanFreq-Y
- fBodyAcc-meanFreq-Z
- fBodyAccJerk-mean-X
- fBodyAccJerk-mean-Y
- fBodyAccJerk-mean-Z
- fBodyAccJerk-std-X
- fBodyAccJerk-std-Y
- fBodyAccJerk-std-Z
- fBodyAccJerk-meanFreq-X
- fBodyAccJerk-meanFreq-Y
- fBodyAccJerk-meanFreq-Z
- fBodyGyro-mean-X
- fBodyGyro-mean-Y
- fBodyGyro-mean-Z
- fBodyGyro-std-X
- fBodyGyro-std-Y
- fBodyGyro-std-Z
- fBodyGyro-meanFreq-X
- fBodyGyro-meanFreq-Y
- fBodyGyro-meanFreq-Z
- fBodyAccMag-mean
- fBodyAccMag-std
- fBodyAccMag-meanFreq
- fBodyBodyAccJerkMag-mean
- fBodyBodyAccJerkMag-std
- fBodyBodyAccJerkMag-meanFreq
- fBodyBodyGyroMag-mean
- fBodyBodyGyroMag-std
- fBodyBodyGyroMag-meanFreq
- fBodyBodyGyroJerkMag-mean
- fBodyBodyGyroJerkMag-std
- fBodyBodyGyroJerkMag-meanFreq

These variables are further processed in our code. We first melted them using subject and activity. So, the

results is a dataset that has four columns (Subject, Activity, Variable, Value). Then, this dataset is casted. We calculated mean value of all the 79 variables for each subject.

The resulting dataset is the final outcome of this analysis.