Codebook

Abhimanyu Banerjee

January 1, 2017

# Getting and Cleaning Data Project

## File Description

This is a codebook for the the tidy dataset produced as a result of running the 'run\_analysis.R' script. It contains information about the variables in the data and the transformations performed on the dataset to get the final tidy dataset.

## Data Description

A full description of the data used in this project can be found at [The UCI Machine Learning Repository](http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones)

## Variable Information

The features selected for the UCI HAR Dataset come from the accelerometer and gyroscope 3-axial raw signals **tAcc-XYZ** and **tGyro-XYZ** (prefix *t* to denote time). The acceleration and gyroscope signals were then separated into body and gravity acceleration signals (**tBodyAcc-XYZ**, **tGravityAcc-XYZ**, **tBodyGyro-XYZ**, **tGravityGyro-XYZ**).

Subsequently, the body linear acceleration and angular velocity were derived in time to obtain Jerk signals (**tBodyAccJerk-XYZ** and **tBodyGyroJerk-XYZ**). Also the magnitude of these three-dimensional signals were calculated using the Euclidean norm (**tBodyAccMag**, **tGravityAccMag**, **tBodyAccJerkMag**, **tBodyGyroMag**, **tBodyGyroJerkMag**).

Finally a **Fast Fourier Transform** (FFT) was applied to some of these signals producing **fBodyAcc-XYZ**, **fBodyAccJerk-XYZ**, **fBodyGyro-XYZ**, **fBodyAccJerkMag**, **fBodyGyroMag**, **fBodyGyroJerkMag** (prefix *f* to indicate frequency domain signals).

These signals were used to estimate variables of the feature vector for each pattern:  
(*-XYZ* is used to denote 3-axial signals in the X, Y and Z directions)

1. tBodyAcc-XYZ
2. tGravityAcc-XYZ
3. tBodyAccJerk-XYZ
4. tBodyGyro-XYZ
5. tBodyGyroJerk-XYZ
6. tBodyAccMag
7. tGravityAccMag
8. tBodyAccJerkMag
9. tBodyGyroMag
10. tBodyGyroJerkMag
11. fBodyAcc-XYZ
12. fBodyAccJerk-XYZ
13. fBodyGyro-XYZ
14. fBodyAccMag
15. fBodyAccJerkMag
16. fBodyGyroMag
17. fBodyGyroJerkMag

The following set of variables were estimated from the signals above:

1. mean(): Mean value
2. std(): Standard deviation
3. mad(): Median absolute deviation
4. max(): Largest value in array
5. min(): Smallest value in array
6. sma(): Signal magnitude area
7. energy(): Energy measure. Sum of the squares divided by the number of values.
8. iqr(): Interquartile range
9. entropy(): Signal entropy
10. arCoeff(): Autorregresion coefficients with Burg order equal to 4
11. correlation(): correlation coefficient between two signals
12. maxInds(): index of the frequency component with largest magnitude
13. meanFreq(): Weighted average of the frequency components to obtain a mean frequency
14. skewness(): skewness of the frequency domain signal
15. kurtosis(): kurtosis of the frequency domain signal
16. bandsEnergy(): Energy of a frequency interval within the 64 bins of the FFT of each window.
17. angle(): Angle between to vectors.

Additional vectors obtained by averaging the signals in a signal window sample: (They were then used on the angle() variable)

1. gravityMean
2. tBodyAccMean
3. tBodyAccJerkMean
4. tBodyGyroMean
5. tBodyGyroJerkMean

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | * Min. | * 1st Qu. | * Median | * Mean 3 | * rd Qu. | * Max. |
| * subjectID | * 1.00000 | * 8.0000000 | * 15.500000 | * 15.500000 | * 23.000000 | * 30.000000 |
| * activityID | * 1.00000 | * 2.0000000 | * 3.500000 | * 3.500000 | * 5.000000 | * 6.000000 |
| * activity | * 30.00000 | * 30.0000000 | * 30.000000 | * 30.000000 | * 30.000000 | * 30.000000 |
| * timeBodyAccMeanX | * 0.22160 | * 0.2712000 | * 0.277000 | * 0.274300 | * 0.280000 | * 0.301500 |
| * timeBodyAccMeanY | * -0.04051 | * -0.0200200 | * -0.017260 | * -0.017880 | * -0.014940 | * -0.001308 |
| * timeBodyAccMeanZ | * -0.15250 | * -0.1121000 | * -0.108200 | * -0.109200 | * -0.104400 | * -0.075380 |
| * timeBodyAccStdX | * -0.99610 | * -0.9799000 | * -0.752600 | * -0.557700 | * -0.198400 | * 0.626900 |
| * timeBodyAccStdY | * -0.99020 | * -0.9421000 | * -0.509000 | * -0.460500 | * -0.030770 | * 0.616900 |
| * timeBodyAccStdZ | * -0.98770 | * -0.9498000 | * -0.651800 | * -0.575600 | * -0.230600 | * 0.609000 |
| * timeGravityAccMeanX | * -0.68000 | * 0.8376000 | * 0.920800 | * 0.697500 | * 0.942500 | * 0.974500 |
| * timeGravityAccMeanY | * -0.47990 | * -0.2332000 | * -0.127800 | * -0.016210 | * 0.087730 | * 0.956600 |
| * timeGravityAccMeanZ | * -0.49510 | * -0.1173000 | * 0.023840 | * 0.074130 | * 0.149500 | * 0.957900 |
| * timeGravityAccStdX | * -0.99680 | * -0.9825000 | * -0.969500 | * -0.963800 | * -0.950900 | * -0.829600 |
| * timeGravityAccStdY | * -0.99420 | * -0.9711000 | * -0.959000 | * -0.952400 | * -0.937000 | * -0.643600 |
| * timeGravityAccStdZ | * -0.99100 | * -0.9605000 | * -0.945000 | * -0.936400 | * -0.918000 | * -0.610200 |
| * timeBodyAccJerkMeanX | * 0.04269 | * 0.0739600 | * 0.076400 | * 0.079470 | * 0.083300 | * 0.130200 |
| * timeBodyAccJerkMeanY | * -0.03869 | * 0.0004664 | * 0.009470 | * 0.007565 | * 0.013400 | * 0.056820 |
| * timeBodyAccJerkMeanZ | * -0.06746 | * -0.0106000 | * -0.003861 | * -0.004953 | * 0.001958 | * 0.038050 |
| * timeBodyAccJerkStdX | * -0.99460 | * -0.9832000 | * -0.810400 | * -0.594900 | * -0.223300 | * 0.544300 |
| * timeBodyAccJerkStdY | * -0.98950 | * -0.9724000 | * -0.775600 | * -0.565400 | * -0.148300 | * 0.355300 |
| * timeBodyAccJerkStdZ | * -0.99330 | * -0.9827000 | * -0.883700 | * -0.736000 | * -0.512100 | * 0.031020 |
| * timeBodyGyroMeanX | * -0.20580 | * -0.0471200 | * -0.028710 | * -0.032440 | * -0.016760 | * 0.192700 |
| * timeBodyGyroMeanY | * -0.20420 | * -0.0895500 | * -0.073180 | * -0.074260 | * -0.061130 | * 0.027470 |
| * timeBodyGyroMeanZ | * -0.07245 | * 0.0747500 | * 0.085120 | * 0.087440 | * 0.101800 | * 0.179100 |
| * timeBodyGyroStdX | * -0.99430 | * -0.9735000 | * -0.789000 | * -0.691600 | * -0.441400 | * 0.267700 |
| * timeBodyGyroStdY | * -0.99420 | * -0.9629000 | * -0.801700 | * -0.653300 | * -0.419600 | * 0.476500 |
| * timeBodyGyroStdZ | * -0.98550 | * -0.9609000 | * -0.801000 | * -0.616400 | * -0.310600 | * 0.564900 |
| * timeBodyGyroJerkMeanX | * -0.15720 | * -0.1032000 | * -0.098680 | * -0.096060 | * -0.091100 | * -0.022090 |
| * timeBodyGyroJerkMeanY | * -0.07681 | * -0.0455200 | * -0.041120 | * -0.042690 | * -0.038420 | * -0.013200 |
| * timeBodyGyroJerkMeanZ | * -0.09250 | * -0.0617200 | * -0.053430 | * -0.054800 | * -0.048980 | * -0.006941 |
| * timeBodyGyroJerkStdX | * -0.99650 | * -0.9800000 | * -0.839600 | * -0.703600 | * -0.462900 | * 0.179100 |
| * timeBodyGyroJerkStdY | * -0.99710 | * -0.9832000 | * -0.894200 | * -0.763600 | * -0.586100 | * 0.295900 |
| * timeBodyGyroJerkStdZ | * -0.99540 | * -0.9848000 | * -0.861000 | * -0.709600 | * -0.474100 | * 0.193200 |
| * timeBodyAccMagnitudeMean | * -0.98650 | * -0.9573000 | * -0.482900 | * -0.497300 | * -0.091900 | * 0.644600 |
| * timeBodyAccMagnitudeStd | * -0.98650 | * -0.9430000 | * -0.607400 | * -0.543900 | * -0.209000 | * 0.428400 |
| * timeGravityAccMagnitudeMean | * -0.98650 | * -0.9573000 | * -0.482900 | * -0.497300 | * -0.091900 | * 0.644600 |
| * timeGravityAccMagnitudeStd | * -0.98650 | * -0.9430000 | * -0.607400 | * -0.543900 | * -0.209000 | * 0.428400 |
| * timeBodyAccJerkMagnitudeMean | * -0.99280 | * -0.9807000 | * -0.816800 | * -0.607900 | * -0.245600 | * 0.434500 |
| * timeBodyAccJerkMagnitudeStd | * -0.99460 | * -0.9765000 | * -0.801400 | * -0.584200 | * -0.217300 | * 0.450600 |
| * timeBodyGyroMagnitudeMean | * -0.98070 | * -0.9461000 | * -0.655100 | * -0.565200 | * -0.215900 | * 0.418000 |
| * timeBodyGyroMagnitudeStd | * -0.98140 | * -0.9476000 | * -0.742000 | * -0.630400 | * -0.360200 | * 0.300000 |
| * timeBodyGyroJerkMagnitudeMean | * -0.99730 | * -0.9852000 | * -0.864800 | * -0.736400 | * -0.511900 | * 0.087580 |
| * timeBodyGyroJerkMagnitudeStd | * -0.99770 | * -0.9805000 | * -0.880900 | * -0.755000 | * -0.576700 | * 0.250200 |
| * freqBodyAccMeanX | * -0.99520 | * -0.9787000 | * -0.769100 | * -0.575800 | * -0.217400 | * 0.537000 |
| * freqBodyAccMeanY | * -0.98900 | * -0.9536000 | * -0.595000 | * -0.488700 | * -0.063410 | * 0.524200 |
| * freqBodyAccMeanZ | * -0.98950 | * -0.9619000 | * -0.723600 | * -0.629700 | * -0.318300 | * 0.280700 |
| * freqBodyAccStdX | * -0.99660 | * -0.9820000 | * -0.747000 | * -0.552200 | * -0.196600 | * 0.658500 |
| * freqBodyAccStdY | * -0.99070 | * -0.9404000 | * -0.513400 | * -0.481500 | * -0.079130 | * 0.560200 |
| * freqBodyAccStdZ | * -0.98720 | * -0.9459000 | * -0.644100 | * -0.582400 | * -0.265500 | * 0.687100 |
| * freqBodyAccJerkMeanX | * -0.99460 | * -0.9828000 | * -0.812600 | * -0.613900 | * -0.282000 | * 0.474300 |
| * freqBodyAccJerkMeanY | * -0.98940 | * -0.9725000 | * -0.781700 | * -0.588200 | * -0.196300 | * 0.276700 |
| * freqBodyAccJerkMeanZ | * -0.99200 | * -0.9796000 | * -0.870700 | * -0.714400 | * -0.469700 | * 0.157800 |
| * freqBodyAccJerkStdX | * -0.99510 | * -0.9847000 | * -0.825400 | * -0.612100 | * -0.247500 | * 0.476800 |
| * freqBodyAccJerkStdY | * -0.99050 | * -0.9737000 | * -0.785200 | * -0.570700 | * -0.168500 | * 0.349800 |
| * freqBodyAccJerkStdZ | * -0.99310 | * -0.9837000 | * -0.895100 | * -0.756500 | * -0.543800 | * -0.006236 |
| * freqBodyGyroMeanX | * -0.99310 | * -0.9697000 | * -0.730000 | * -0.636700 | * -0.338700 | * 0.475000 |
| * freqBodyGyroMeanY | * -0.99400 | * -0.9700000 | * -0.814100 | * -0.676700 | * -0.445800 | * 0.328800 |
| * freqBodyGyroMeanZ | * -0.98600 | * -0.9624000 | * -0.790900 | * -0.604400 | * -0.263500 | * 0.492400 |
| * freqBodyGyroStdX | * -0.99470 | * -0.9750000 | * -0.808600 | * -0.711000 | * -0.481300 | * 0.196600 |
| * freqBodyGyroStdY | * -0.99440 | * -0.9602000 | * -0.796400 | * -0.645400 | * -0.415400 | * 0.646200 |
| * freqBodyGyroStdZ | * -0.98670 | * -0.9643000 | * -0.822400 | * -0.657700 | * -0.391600 | * 0.522500 |
| * freqBodyAccMagnitudeMean | * -0.98680 | * -0.9560000 | * -0.670300 | * -0.536500 | * -0.162200 | * 0.586600 |
| * freqBodyAccMagnitudeStd | * -0.98760 | * -0.9452000 | * -0.651300 | * -0.621000 | * -0.365400 | * 0.178700 |
| * freqBodyAccJerkMagnitudeMean | * -0.99400 | * -0.9770000 | * -0.794000 | * -0.575600 | * -0.187200 | * 0.538400 |
| * freqBodyAccJerkMagnitudeStd | * -0.99440 | * -0.9752000 | * -0.812600 | * -0.599200 | * -0.266800 | * 0.316300 |
| * freqBodyGyroMagnitudeMean | * -0.98650 | * -0.9616000 | * -0.765700 | * -0.667100 | * -0.408700 | * 0.204000 |
| * freqBodyGyroMagnitudeStd | * -0.98150 | * -0.9488000 | * -0.772700 | * -0.672300 | * -0.427700 | * 0.236700 |
| * freqBodyGyroJerkMagnitudeMean | * -0.99760 | * -0.9813000 | * -0.877900 | * -0.756400 | * -0.583100 | * 0.146600 |
| * freqBodyGyroJerkMagnitudeStd | * -0.99760 | * -0.9802000 | * -0.894100 | * -0.771500 | * -0.608100 | * 0.287800 |