

Code book

The two datasets produced in this project are based on the data available at: <https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip>

Dataset 1

Name: HumanActivity_1_tidy.csv

Description: This dataset contains data from the train and test datasets included in the downloaded dataset.

Transformations:

1. Integrate subject (`subject_[dataset].txt`), observations (`X_[dataset].txt`) and label (`y_[dataset].txt`), for the train and test datasets.
2. Merge train and test datasets.
3. Add columns names obtained from `features.txt`.
4. Select the columns for average and standard deviation. Additionally, the columns subject and activity were included.
5. Clean column names by removing leading numbers, parentheses, and spaces.

Columns:

1. **subject:** Numeric ID between 1 and 30 to identify the volunteer.
2. **Variables:** 79 variables for each observation.
3. **activity:** Activity names as defined in the original dataset. The included values are: WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, LAYING.

Output of `str(dataset)`

```
file_path <- file.path("data", "HumanActivity_1_tidy.txt")
data <- read.csv(file_path, sep = "")
str(data)

## 'data.frame':    10299 obs. of  81 variables:
## $ subject      : int  1 1 1 1 1 1 1 1 1 1 ...
## $ tBodyAcc.mean.X : num  0.289 0.278 0.28 0.279 0.277 ...
## $ tBodyAcc.mean.Y : num  -0.0203 -0.0164 -0.0195 -0.0262 -0.0166 ...
## $ tBodyAcc.mean.Z : num  -0.133 -0.124 -0.113 -0.123 -0.115 ...
## $ tBodyAcc.std.X  : num  -0.995 -0.998 -0.995 -0.996 -0.998 ...
## $ tBodyAcc.std.Y  : num  -0.983 -0.975 -0.967 -0.983 -0.981 ...
## $ tBodyAcc.std.Z  : num  -0.914 -0.96 -0.979 -0.991 -0.99 ...
## $ tGravityAcc.mean.X : num  0.963 0.967 0.967 0.968 0.968 ...
## $ tGravityAcc.mean.Y : num  -0.141 -0.142 -0.142 -0.144 -0.149 ...
## $ tGravityAcc.mean.Z : num  0.1154 0.1094 0.1019 0.0999 0.0945 ...
## $ tGravityAcc.std.X : num  -0.985 -0.997 -1 -0.997 -0.998 ...
## $ tGravityAcc.std.Y : num  -0.982 -0.989 -0.993 -0.981 -0.988 ...
## $ tGravityAcc.std.Z : num  -0.878 -0.932 -0.993 -0.978 -0.979 ...
## $ tBodyAccJerk.mean.X : num  0.078 0.074 0.0736 0.0773 0.0734 ...
## $ tBodyAccJerk.mean.Y : num  0.005 0.00577 0.0031 0.02006 0.01912 ...
## $ tBodyAccJerk.mean.Z : num  -0.06783 0.02938 -0.00905 -0.00986 0.01678 ...
## $ tBodyAccJerk.std.X : num  -0.994 -0.996 -0.991 -0.993 -0.996 ...
## $ tBodyAccJerk.std.Y : num  -0.988 -0.981 -0.981 -0.988 -0.988 ...
## $ tBodyAccJerk.std.Z : num  -0.994 -0.992 -0.99 -0.993 -0.992 ...
## $ tBodyGyro.mean.X : num  -0.0061 -0.0161 -0.0317 -0.0434 -0.034 ...
## $ tBodyGyro.mean.Y : num  -0.0314 -0.0839 -0.1023 -0.0914 -0.0747 ...
## $ tBodyGyro.mean.Z : num  0.1077 0.1006 0.0961 0.0855 0.0774 ...
```

```

## $ tBodyGyro.std.X           : num -0.985 -0.983 -0.976 -0.991 -0.985 ...
## $ tBodyGyro.std.Y           : num -0.977 -0.989 -0.994 -0.992 -0.992 ...
## $ tBodyGyro.std.Z           : num -0.992 -0.989 -0.986 -0.988 -0.987 ...
## $ tBodyGyroJerk.mean.X      : num -0.0992 -0.1105 -0.1085 -0.0912 -0.0908 ...
## $ tBodyGyroJerk.mean.Y      : num -0.0555 -0.0448 -0.0424 -0.0363 -0.0376 ...
## $ tBodyGyroJerk.mean.Z      : num -0.062 -0.0592 -0.0558 -0.0605 -0.0583 ...
## $ tBodyGyroJerk.std.X       : num -0.992 -0.99 -0.988 -0.991 -0.991 ...
## $ tBodyGyroJerk.std.Y       : num -0.993 -0.997 -0.996 -0.997 -0.996 ...
## $ tBodyGyroJerk.std.Z       : num -0.992 -0.994 -0.992 -0.993 -0.995 ...
## $ tBodyAccMag.mean          : num -0.959 -0.979 -0.984 -0.987 -0.993 ...
## $ tBodyAccMag.std           : num -0.951 -0.976 -0.988 -0.986 -0.991 ...
## $ tGravityAccMag.mean       : num -0.959 -0.979 -0.984 -0.987 -0.993 ...
## $ tGravityAccMag.std        : num -0.951 -0.976 -0.988 -0.986 -0.991 ...
## $ tBodyAccJerkMag.mean      : num -0.993 -0.991 -0.989 -0.993 -0.993 ...
## $ tBodyAccJerkMag.std       : num -0.994 -0.992 -0.99 -0.993 -0.996 ...
## $ tBodyGyroMag.mean         : num -0.969 -0.981 -0.976 -0.982 -0.985 ...
## $ tBodyGyroMag.std          : num -0.964 -0.984 -0.986 -0.987 -0.989 ...
## $ tBodyGyroJerkMag.mean     : num -0.994 -0.995 -0.993 -0.996 -0.996 ...
## $ tBodyGyroJerkMag.std      : num -0.991 -0.996 -0.995 -0.995 -0.995 ...
## $ fBodyAcc.mean.X           : num -0.995 -0.997 -0.994 -0.995 -0.997 ...
## $ fBodyAcc.mean.Y           : num -0.983 -0.977 -0.973 -0.984 -0.982 ...
## $ fBodyAcc.mean.Z           : num -0.939 -0.974 -0.983 -0.991 -0.988 ...
## $ fBodyAcc.std.X            : num -0.995 -0.999 -0.996 -0.996 -0.999 ...
## $ fBodyAcc.std.Y            : num -0.983 -0.975 -0.966 -0.983 -0.98 ...
## $ fBodyAcc.std.Z            : num -0.906 -0.955 -0.977 -0.99 -0.992 ...
## $ fBodyAcc.meanFreq.X       : num 0.252 0.271 0.125 0.029 0.181 ...
## $ fBodyAcc.meanFreq.Y       : num 0.1318 0.0429 -0.0646 0.0803 0.058 ...
## $ fBodyAcc.meanFreq.Z       : num -0.0521 -0.0143 0.0827 0.1857 0.5598 ...
## $ fBodyAccJerk.mean.X       : num -0.992 -0.995 -0.991 -0.994 -0.996 ...
## $ fBodyAccJerk.mean.Y       : num -0.987 -0.981 -0.982 -0.989 -0.989 ...
## $ fBodyAccJerk.mean.Z       : num -0.99 -0.99 -0.988 -0.991 -0.991 ...
## $ fBodyAccJerk.std.X        : num -0.996 -0.997 -0.991 -0.991 -0.997 ...
## $ fBodyAccJerk.std.Y        : num -0.991 -0.982 -0.981 -0.987 -0.989 ...
## $ fBodyAccJerk.std.Z        : num -0.997 -0.993 -0.99 -0.994 -0.993 ...
## $ fBodyAccJerk.meanFreq.X   : num 0.8704 0.6085 0.1154 0.0358 0.2734 ...
## $ fBodyAccJerk.meanFreq.Y   : num 0.2107 -0.0537 -0.1934 -0.093 0.0791 ...
## $ fBodyAccJerk.meanFreq.Z   : num 0.2637 0.0631 0.0383 0.1681 0.2924 ...
## $ fBodyGyro.mean.X          : num -0.987 -0.977 -0.975 -0.987 -0.982 ...
## $ fBodyGyro.mean.Y          : num -0.982 -0.993 -0.994 -0.994 -0.993 ...
## $ fBodyGyro.mean.Z          : num -0.99 -0.99 -0.987 -0.987 -0.989 ...
## $ fBodyGyro.std.X           : num -0.985 -0.985 -0.977 -0.993 -0.986 ...
## $ fBodyGyro.std.Y           : num -0.974 -0.987 -0.993 -0.992 -0.992 ...
## $ fBodyGyro.std.Z           : num -0.994 -0.99 -0.987 -0.989 -0.988 ...
## $ fBodyGyro.meanFreq.X      : num -0.2575 -0.0482 -0.2167 0.2169 -0.1533 ...
## $ fBodyGyro.meanFreq.Y      : num 0.0979 -0.4016 -0.0173 -0.1352 -0.0884 ...
## $ fBodyGyro.meanFreq.Z      : num 0.5472 -0.0682 -0.1107 -0.0497 -0.1622 ...
## $ fBodyAccMag.mean          : num -0.952 -0.981 -0.988 -0.988 -0.994 ...
## $ fBodyAccMag.std           : num -0.956 -0.976 -0.989 -0.987 -0.99 ...
## $ fBodyAccMag.meanFreq      : num -0.0884 -0.0441 0.2579 0.0736 0.3943 ...
## $ fBodyBodyAccJerkMag.mean   : num -0.994 -0.99 -0.989 -0.993 -0.996 ...
## $ fBodyBodyAccJerkMag.std    : num -0.994 -0.992 -0.991 -0.992 -0.994 ...
## $ fBodyBodyAccJerkMag.meanFreq : num 0.347 0.532 0.661 0.679 0.559 ...
## $ fBodyBodyGyroMag.mean     : num -0.98 -0.988 -0.989 -0.989 -0.991 ...
## $ fBodyBodyGyroMag.std      : num -0.961 -0.983 -0.986 -0.988 -0.989 ...

```

```
## $ fBodyBodyGyroMag.meanFreq : num -0.129 -0.272 -0.2127 -0.0357 -0.2736 ...
## $ fBodyBodyGyroJerkMag.mean : num -0.992 -0.996 -0.995 -0.995 -0.995 ...
## $ fBodyBodyGyroJerkMag.std : num -0.991 -0.996 -0.995 -0.995 -0.995 ...
## $ fBodyBodyGyroJerkMag.meanFreq: num -0.0743 0.1581 0.4145 0.4046 0.0878 ...
## $ activity : chr "STANDING" "STANDING" "STANDING" "STANDING" ...
```

Dataset 2

Name: HumanActivity_2_tidy.csv

Description: This dataset contains averages of each variable for each activity and each subject.

Columns:

1. **variable:** Variable name corresponding to the variables X to Y from the previous dataset.
2. **activity:** Activity names as defined in the original dataset. The included values are: WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, LAYING.
3. **subject:** Numeric ID for the volunteer which ranges between 1 and 30.
4. **average:** The average for the group defined by variable, activity, and subject.

Output of str(dataset)

```
file_path <- file.path("data", "HumanActivity_2_tidy.txt")
data <- read.csv(file_path, sep = "")
str(data)
```

```
## 'data.frame': 14220 obs. of 4 variables:
## $ variable: chr "fBodyAcc-mean-X" "fBodyAcc-mean-X" "fBodyAcc-mean-X" "fBodyAcc-mean-X" ...
## $ activity: chr "LAYING" "LAYING" "LAYING" "LAYING" ...
## $ subject : int 1 2 3 4 5 6 7 8 9 10 ...
## $ average : num -0.939 -0.977 -0.981 -0.959 -0.969 ...
```