Human Activity Recognition Using Smartphones Dataset

Gambera Matteo

05-05-2020

The purpose of this project is to demonstrate your ability to collect, work with, and clean a data set. The goal is to prepare tidy data that can be used for later analysis.

1. Merges the training and the test sets to create one data set.
2. Extracts only the measurements on the mean and standard deviation for each measurement.
3. Uses descriptive activity names to name the activities in the data set
4. Appropriately labels the data set with descriptive variable names.
5. From the data set in step 4, creates a second, independent tidy data set with the average of each variable for each activity and each subject.

Load Data

```r
library(reshape2)
packages <- c("data.table", "reshape2")
sapply(packages, require, character.only=TRUE, quietly=TRUE)
```

```
##
## Attaching package: 'data.table'

## The following objects are masked from 'package:reshape2':
##
##     dcast, melt

## data.table    reshape2
##       TRUE        TRUE
```

```r
url <- "https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip"
# i can use also getwd() for directory path
path <- getwd()
#path <- "/home/matteo/Scrivania/datasciencecoursera/Course_2/Data"
# download the zip file
download.file(url, file.path(path, "dataFiles.zip"))
# unzip
unzip(zipfile = "dataFiles.zip")
```

fread Similar to read.table but faster and more convenient. All controls such as sep, colClasses and nrows are automatically detected. bit64::integer64 types are also detected and read directly without needing to read as character before converting.

Clean Data

```r
#check the name of what type of activity the accelerometer registers
activityLabels <- fread(file.path(path, "UCI HAR Dataset/activity_labels.txt"), col.names = c("Labels",
#view
activityLabels
```

```
##    Labels          Activity
## 1:      1           WALKING
## 2:      2  WALKING_UPSTAIRS
## 3:      3 WALKING_DOWNSTAIRS
## 4:      4           SITTING
## 5:      5          STANDING
## 6:      6            LAYING
```

```r
str(activityLabels)
```

```
## Classes 'data.table' and 'data.frame':   6 obs. of  2 variables:
##  $ Labels  : int  1 2 3 4 5 6
##  $ Activity: chr  "WALKING" "WALKING_UPSTAIRS" "WALKING_DOWNSTAIRS" "SITTING" ...
##  - attr(*, ".internal.selfref")=<externalptr>
```

There are 6 type of activity

```r
features <- fread(file.path(path, "UCI HAR Dataset/features.txt"), col.names = c("id", "featureNames"))
head(features)
```

```
##    id       featureNames
## 1:  1 tBodyAcc-mean()-X
## 2:  2 tBodyAcc-mean()-Y
## 3:  3 tBodyAcc-mean()-Z
## 4:  4  tBodyAcc-std()-X
## 5:  5  tBodyAcc-std()-Y
## 6:  6  tBodyAcc-std()-Z
```

```r
str(features)
```

```
## Classes 'data.table' and 'data.frame':   561 obs. of  2 variables:
##  $ id          : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ featureNames: chr  "tBodyAcc-mean()-X" "tBodyAcc-mean()-Y" "tBodyAcc-mean()-Z" "tBodyAcc-std()-X"
##  - attr(*, ".internal.selfref")=<externalptr>
```

I don't want all features(exp ENTROPY),i select and use the only ones i'm interested (words with mean and std)

```r
featuresid <- grep("(mean|std)\\(\\)", features[, featureNames])
featuresid
```

```
## [1]    1    2    3    4    5    6   41   42   43   44   45   46   81   82   83   84   85   86  121
## [20]  122  123  124  125  126  161  162  163  164  165  166  201  202  214  215  227  228  240  241
## [39]  253  254  266  267  268  269  270  271  345  346  347  348  349  350  424  425  426  427  428
## [58]  429  503  504  516  517  529  530  542  543
```

check for correct grep

```
meanstdmeasure <- features[featuresid, featureNames]
meanstdmeasure
```

```
##  [1] "tBodyAcc-mean()-X"         "tBodyAcc-mean()-Y"
##  [3] "tBodyAcc-mean()-Z"         "tBodyAcc-std()-X"
##  [5] "tBodyAcc-std()-Y"          "tBodyAcc-std()-Z"
##  [7] "tGravityAcc-mean()-X"      "tGravityAcc-mean()-Y"
##  [9] "tGravityAcc-mean()-Z"      "tGravityAcc-std()-X"
## [11] "tGravityAcc-std()-Y"       "tGravityAcc-std()-Z"
## [13] "tBodyAccJerk-mean()-X"     "tBodyAccJerk-mean()-Y"
## [15] "tBodyAccJerk-mean()-Z"     "tBodyAccJerk-std()-X"
## [17] "tBodyAccJerk-std()-Y"      "tBodyAccJerk-std()-Z"
## [19] "tBodyGyro-mean()-X"        "tBodyGyro-mean()-Y"
## [21] "tBodyGyro-mean()-Z"        "tBodyGyro-std()-X"
## [23] "tBodyGyro-std()-Y"         "tBodyGyro-std()-Z"
## [25] "tBodyGyroJerk-mean()-X"    "tBodyGyroJerk-mean()-Y"
## [27] "tBodyGyroJerk-mean()-Z"    "tBodyGyroJerk-std()-X"
## [29] "tBodyGyroJerk-std()-Y"     "tBodyGyroJerk-std()-Z"
## [31] "tBodyAccMag-mean()"        "tBodyAccMag-std()"
## [33] "tGravityAccMag-mean()"     "tGravityAccMag-std()"
## [35] "tBodyAccJerkMag-mean()"    "tBodyAccJerkMag-std()"
## [37] "tBodyGyroMag-mean()"       "tBodyGyroMag-std()"
## [39] "tBodyGyroJerkMag-mean()"   "tBodyGyroJerkMag-std()"
## [41] "fBodyAcc-mean()-X"         "fBodyAcc-mean()-Y"
## [43] "fBodyAcc-mean()-Z"         "fBodyAcc-std()-X"
## [45] "fBodyAcc-std()-Y"          "fBodyAcc-std()-Z"
## [47] "fBodyAccJerk-mean()-X"     "fBodyAccJerk-mean()-Y"
## [49] "fBodyAccJerk-mean()-Z"     "fBodyAccJerk-std()-X"
## [51] "fBodyAccJerk-std()-Y"      "fBodyAccJerk-std()-Z"
## [53] "fBodyGyro-mean()-X"        "fBodyGyro-mean()-Y"
## [55] "fBodyGyro-mean()-Z"        "fBodyGyro-std()-X"
## [57] "fBodyGyro-std()-Y"         "fBodyGyro-std()-Z"
## [59] "fBodyAccMag-mean()"        "fBodyAccMag-std()"
## [61] "fBodyBodyAccJerkMag-mean()" "fBodyBodyAccJerkMag-std()"
## [63] "fBodyBodyGyroMag-mean()"    "fBodyBodyGyroMag-std()"
## [65] "fBodyBodyGyroJerkMag-mean()" "fBodyBodyGyroJerkMag-std()"
```

cut "()" and from feautures Names

```
meanstdmeasure <- gsub('[()]', '', meanstdmeasure)
meanstdmeasure
```

```
##  [1] "tBodyAcc-mean-X"         "tBodyAcc-mean-Y"
##  [3] "tBodyAcc-mean-Z"         "tBodyAcc-std-X"
##  [5] "tBodyAcc-std-Y"          "tBodyAcc-std-Z"
##  [7] "tGravityAcc-mean-X"      "tGravityAcc-mean-Y"
```

```
##  [9] "tGravityAcc-mean-Z"        "tGravityAcc-std-X"
## [11] "tGravityAcc-std-Y"         "tGravityAcc-std-Z"
## [13] "tBodyAccJerk-mean-X"       "tBodyAccJerk-mean-Y"
## [15] "tBodyAccJerk-mean-Z"       "tBodyAccJerk-std-X"
## [17] "tBodyAccJerk-std-Y"        "tBodyAccJerk-std-Z"
## [19] "tBodyGyro-mean-X"          "tBodyGyro-mean-Y"
## [21] "tBodyGyro-mean-Z"          "tBodyGyro-std-X"
## [23] "tBodyGyro-std-Y"           "tBodyGyro-std-Z"
## [25] "tBodyGyroJerk-mean-X"      "tBodyGyroJerk-mean-Y"
## [27] "tBodyGyroJerk-mean-Z"      "tBodyGyroJerk-std-X"
## [29] "tBodyGyroJerk-std-Y"       "tBodyGyroJerk-std-Z"
## [31] "tBodyAccMag-mean"          "tBodyAccMag-std"
## [33] "tGravityAccMag-mean"       "tGravityAccMag-std"
## [35] "tBodyAccJerkMag-mean"      "tBodyAccJerkMag-std"
## [37] "tBodyGyroMag-mean"         "tBodyGyroMag-std"
## [39] "tBodyGyroJerkMag-mean"     "tBodyGyroJerkMag-std"
## [41] "fBodyAcc-mean-X"           "fBodyAcc-mean-Y"
## [43] "fBodyAcc-mean-Z"           "fBodyAcc-std-X"
## [45] "fBodyAcc-std-Y"            "fBodyAcc-std-Z"
## [47] "fBodyAccJerk-mean-X"       "fBodyAccJerk-mean-Y"
## [49] "fBodyAccJerk-mean-Z"       "fBodyAccJerk-std-X"
## [51] "fBodyAccJerk-std-Y"        "fBodyAccJerk-std-Z"
## [53] "fBodyGyro-mean-X"          "fBodyGyro-mean-Y"
## [55] "fBodyGyro-mean-Z"          "fBodyGyro-std-X"
## [57] "fBodyGyro-std-Y"           "fBodyGyro-std-Z"
## [59] "fBodyAccMag-mean"          "fBodyAccMag-std"
## [61] "fBodyBodyAccJerkMag-mean"  "fBodyBodyAccJerkMag-std"
## [63] "fBodyBodyGyroMag-mean"     "fBodyBodyGyroMag-std"
## [65] "fBodyBodyGyroJerkMag-mean" "fBodyBodyGyroJerkMag-std"
```

Load Train

load the train dataset

```
train <- fread(file.path(path, "UCI HAR Dataset/train/X_train.txt"))[, featuresid, with = FALSE]
data.table::setnames(train, colnames(train), meanstdmeasure) #rename the columns with meanstdmeasure
head(train, n=1)
```

```
##    tBodyAcc-mean-X tBodyAcc-mean-Y tBodyAcc-mean-Z tBodyAcc-std-X
## 1:      0.2885845     -0.02029417      -0.1329051     -0.9952786
##    tBodyAcc-std-Y tBodyAcc-std-Z tGravityAcc-mean-X tGravityAcc-mean-Y
## 1:     -0.9831106     -0.9135264          0.9633961         -0.1408397
##    tGravityAcc-mean-Z tGravityAcc-std-X tGravityAcc-std-Y tGravityAcc-std-Z
## 1:          0.1153749        -0.9852497        -0.9817084          -0.877625
##    tBodyAccJerk-mean-X tBodyAccJerk-mean-Y tBodyAccJerk-mean-Z
## 1:          0.07799634         0.005000803         -0.06783081
##    tBodyAccJerk-std-X tBodyAccJerk-std-Y tBodyAccJerk-std-Z tBodyGyro-mean-X
## 1:         -0.9935191            -0.98836          -0.993575     -0.006100849
##    tBodyGyro-mean-Y tBodyGyro-mean-Z tBodyGyro-std-X tBodyGyro-std-Y
## 1:      -0.03136479        0.1077254      -0.9853103      -0.9766234
##    tBodyGyro-std-Z tBodyGyroJerk-mean-X tBodyGyroJerk-mean-Y
## 1:      -0.9922053           -0.0991674           -0.05551737
```

```
##     tBodyGyroJerk-mean-Z tBodyGyroJerk-std-X tBodyGyroJerk-std-Y
## 1:           -0.0619858         -0.9921107         -0.9925193
##     tBodyGyroJerk-std-Z tBodyAccMag-mean tBodyAccMag-std tGravityAccMag-mean
## 1:          -0.9920553      -0.9594339      -0.9505515          -0.9594339
##     tGravityAccMag-std tBodyAccJerkMag-mean tBodyAccJerkMag-std
## 1:          -0.9505515           -0.9933059          -0.9943364
##     tBodyGyroMag-mean tBodyGyroMag-std tBodyGyroJerkMag-mean
## 1:         -0.9689591       -0.9643352            -0.9942478
##     tBodyGyroJerkMag-std fBodyAcc-mean-X fBodyAcc-mean-Y fBodyAcc-mean-Z
## 1:           -0.9913676     -0.9947832      -0.9829841      -0.9392687
##     fBodyAcc-std-X fBodyAcc-std-Y fBodyAcc-std-Z fBodyAccJerk-mean-X
## 1:     -0.9954217     -0.983133      -0.906165          -0.9923325
##     fBodyAccJerk-mean-Y fBodyAccJerk-mean-Z fBodyAccJerk-std-X
## 1:          -0.9871699          -0.9896961         -0.9958207
##     fBodyAccJerk-std-Y fBodyAccJerk-std-Z fBodyGyro-mean-X fBodyGyro-mean-Y
## 1:          -0.9909363         -0.9970517      -0.9865744      -0.9817615
##     fBodyGyro-mean-Z fBodyGyro-std-X fBodyGyro-std-Y fBodyGyro-std-Z
## 1:        -0.9895148      -0.9850326      -0.9738861      -0.9940349
##     fBodyAccMag-mean fBodyAccMag-std fBodyBodyAccJerkMag-mean
## 1:       -0.9521547      -0.956134               -0.9937257
##     fBodyBodyAccJerkMag-std fBodyBodyGyroMag-mean fBodyBodyGyroMag-std
## 1:              -0.993755            -0.9801349           -0.9613094
##     fBodyBodyGyroJerkMag-mean fBodyBodyGyroJerkMag-std
## 1:               -0.9919904               -0.9906975
```

```r
str(train)
```

```
## Classes 'data.table' and 'data.frame':   7352 obs. of  66 variables:
##  $ 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 ...
##  $ 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 ...
##  $ 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 ...
##  $ 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 ...
##  $ 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 ...
##  $ 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 ...
##  $ 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 ...
##  - attr(*, ".internal.selfref")=<externalptr>
```

load train activities and subjects

```r
trainActivities <- fread(file.path(path, "UCI HAR Dataset/train/y_train.txt"), col.names = c("Activity")
trainActivities
```

```
##       Activity
##    1:        5
##    2:        5
##    3:        5
##    4:        5
##    5:        5
##   ---
```

```
## 7348:          2
## 7349:          2
## 7350:          2
## 7351:          2
## 7352:          2
```

```r
trainSubjects <- fread(file.path(path, "UCI HAR Dataset/train/subject_train.txt"), col.names = c("Subje
trainSubjects
```

```
##        SubjectNum
##     1:          1
##     2:          1
##     3:          1
##     4:          1
##     5:          1
##    ---
## 7348:         30
## 7349:         30
## 7350:         30
## 7351:         30
## 7352:         30
```

merge this data to train

```r
train <- cbind(trainActivities,trainSubjects,train)
head(train, n=1)
```

```
##     Activity SubjectNum tBodyAcc-mean-X tBodyAcc-mean-Y tBodyAcc-mean-Z
## 1:         5          1       0.2885845     -0.02029417      -0.1329051
##     tBodyAcc-std-X tBodyAcc-std-Y tBodyAcc-std-Z tGravityAcc-mean-X
## 1:      -0.9952786     -0.9831106     -0.9135264          0.9633961
##     tGravityAcc-mean-Y tGravityAcc-mean-Z tGravityAcc-std-X tGravityAcc-std-Y
## 1:         -0.1408397          0.1153749        -0.9852497         -0.9817084
##     tGravityAcc-std-Z tBodyAccJerk-mean-X tBodyAccJerk-mean-Y
## 1:         -0.877625          0.07799634          0.005000803
##     tBodyAccJerk-mean-Z tBodyAccJerk-std-X tBodyAccJerk-std-Y tBodyAccJerk-std-Z
## 1:         -0.06783081         -0.9935191            -0.98836           -0.993575
##     tBodyGyro-mean-X tBodyGyro-mean-Y tBodyGyro-mean-Z tBodyGyro-std-X
## 1:      -0.006100849      -0.03136479        0.1077254      -0.9853103
##     tBodyGyro-std-Y tBodyGyro-std-Z tBodyGyroJerk-mean-X tBodyGyroJerk-mean-Y
## 1:      -0.9766234      -0.9922053          -0.0991674          -0.05551737
##     tBodyGyroJerk-mean-Z tBodyGyroJerk-std-X tBodyGyroJerk-std-Y
## 1:          -0.0619858          -0.9921107          -0.9925193
##     tBodyGyroJerk-std-Z tBodyAccMag-mean tBodyAccMag-std tGravityAccMag-mean
## 1:          -0.9920553      -0.9594339      -0.9505515          -0.9594339
##     tGravityAccMag-std tBodyAccJerkMag-mean tBodyAccJerkMag-std
## 1:         -0.9505515           -0.9933059          -0.9943364
##     tBodyGyroMag-mean tBodyGyroMag-std tBodyGyroJerkMag-mean
## 1:         -0.9689591       -0.9643352            -0.9942478
##     tBodyGyroJerkMag-std fBodyAcc-mean-X fBodyAcc-mean-Y fBodyAcc-mean-Z
## 1:           -0.9913676      -0.9947832      -0.9829841      -0.9392687
##     fBodyAcc-std-X fBodyAcc-std-Y fBodyAcc-std-Z fBodyAccJerk-mean-X
```

```
## 1:       -0.9954217       -0.983133       -0.906165          -0.9923325
##    fBodyAccJerk-mean-Y fBodyAccJerk-mean-Z fBodyAccJerk-std-X
## 1:          -0.9871699          -0.9896961          -0.9958207
##    fBodyAccJerk-std-Y fBodyAccJerk-std-Z fBodyGyro-mean-X fBodyGyro-mean-Y
## 1:          -0.9909363          -0.9970517          -0.9865744          -0.9817615
##    fBodyGyro-mean-Z fBodyGyro-std-X fBodyGyro-std-Y fBodyGyro-std-Z
## 1:          -0.9895148          -0.9850326          -0.9738861          -0.9940349
##    fBodyAccMag-mean fBodyAccMag-std fBodyBodyAccJerkMag-mean
## 1:          -0.9521547          -0.956134          -0.9937257
##    fBodyBodyAccJerkMag-std fBodyBodyGyroMag-mean fBodyBodyGyroMag-std
## 1:          -0.993755          -0.9801349          -0.9613094
##    fBodyBodyGyroJerkMag-mean fBodyBodyGyroJerkMag-std
## 1:          -0.9919904          -0.9906975
```

Load test

same as train

```
test <- fread(file.path(path, "UCI HAR Dataset/test/X_test.txt"))[, featuresid, with = FALSE]
data.table::setnames(test, colnames(test), meanstdmeasure)
testActivities <- fread(file.path(path, "UCI HAR Dataset/test/y_test.txt")
                        , col.names = c("Activity"))
testSubjects <- fread(file.path(path, "UCI HAR Dataset/test/subject_test.txt")
                      , col.names = c("SubjectNum"))
test <- cbind(testSubjects, testActivities, test)
head(test, n=1)
```

```
##    SubjectNum Activity tBodyAcc-mean-X tBodyAcc-mean-Y tBodyAcc-mean-Z
## 1:          2          5       0.2571778      -0.02328523      -0.01465376
##    tBodyAcc-std-X tBodyAcc-std-Y tBodyAcc-std-Z tGravityAcc-mean-X
## 1:      -0.938404      -0.9200908      -0.6676833          0.9364893
##    tGravityAcc-mean-Y tGravityAcc-mean-Z tGravityAcc-std-X tGravityAcc-std-Y
## 1:          -0.2827192          0.1152882          -0.9254273          -0.9370141
##    tGravityAcc-std-Z tBodyAccJerk-mean-X tBodyAccJerk-mean-Y
## 1:          -0.5642884          0.07204601          0.0457544
##    tBodyAccJerk-mean-Z tBodyAccJerk-std-X tBodyAccJerk-std-Y tBodyAccJerk-std-Z
## 1:          -0.1060427          -0.9066828          -0.9380164          -0.9359358
##    tBodyGyro-mean-X tBodyGyro-mean-Y tBodyGyro-mean-Z tBodyGyro-std-X
## 1:          0.1199762          -0.09179234          0.1896285          -0.8830891
##    tBodyGyro-std-Y tBodyGyro-std-Z tBodyGyroJerk-mean-X tBodyGyroJerk-mean-Y
## 1:          -0.8161636          -0.9408812          -0.2048962          -0.1744877
##    tBodyGyroJerk-mean-Z tBodyGyroJerk-std-X tBodyGyroJerk-std-Y
## 1:          -0.09338934          -0.9012242          -0.9108601
##    tBodyGyroJerk-std-Z tBodyAccMag-mean tBodyAccMag-std tGravityAccMag-mean
## 1:          -0.9392504          -0.8669294          -0.7051911          -0.8669294
##    tGravityAccMag-std tBodyAccJerkMag-mean tBodyAccJerkMag-std
## 1:          -0.7051911          -0.9297665          -0.8959942
##    tBodyGyroMag-mean tBodyGyroMag-std tBodyGyroJerkMag-mean
## 1:          -0.7955439          -0.7620732          -0.9251949
##    tBodyGyroJerkMag-std fBodyAcc-mean-X fBodyAcc-mean-Y fBodyAcc-mean-Z
## 1:          -0.8943436          -0.9185097          -0.9182132          -0.7890915
##    fBodyAcc-std-X fBodyAcc-std-Y fBodyAcc-std-Z fBodyAccJerk-mean-X
```

8

```
## 1:      -0.9482903    -0.9251369    -0.6363167            -0.8996332
##    fBodyAccJerk-mean-Y fBodyAccJerk-mean-Z fBodyAccJerk-std-X
## 1:        -0.937485        -0.9235514        -0.9244291
##    fBodyAccJerk-std-Y fBodyAccJerk-std-Z fBodyGyro-mean-X fBodyGyro-mean-Y
## 1:        -0.9432104        -0.9478915      -0.8235579        -0.807916
##    fBodyGyro-mean-Z fBodyGyro-std-X fBodyGyro-std-Y fBodyGyro-std-Z
## 1:      -0.9179126      -0.9032627      -0.822677      -0.9561651
##    fBodyAccMag-mean fBodyAccMag-std fBodyBodyAccJerkMag-mean
## 1:      -0.7909464      -0.711074              -0.8950612
##    fBodyBodyAccJerkMag-std fBodyBodyGyroMag-mean fBodyBodyGyroMag-std
## 1:            -0.8963596            -0.77061            -0.7971128
##    fBodyBodyGyroJerkMag-mean fBodyBodyGyroJerkMag-std
## 1:              -0.8901655              -0.9073076
```

Merge test and train

```
data <- rbind(train, test)
str(data)
```

```
## Classes 'data.table' and 'data.frame':   10299 obs. of  68 variables:
##  $ Activity             : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ SubjectNum           : 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 ...
```

9

```
## $ 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 ...
## $ 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 ...
## $ 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 ...
## $ 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 ...
## $ 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 ...
## $ 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 ...
## $ 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 ...
## - attr(*, ".internal.selfref")=<externalptr>
```

```r
nrow(data)==sum(nrow(test)+nrow(train))
```

```
## [1] TRUE
```

```r
data[["Activity"]] <- factor(data[, Activity], levels = activityLabels[["classLabels"]], labels = activi
data[["SubjectNum"]] <- as.factor(data[, SubjectNum])
data <- reshape2::melt(data = data, id = c("SubjectNum", "Activity"))
data <- reshape2::dcast(data = data, SubjectNum + Activity ~ variable, fun.aggregate = mean)
```

```r
data.table::fwrite(x = data, file = "tidyData.txt", quote = FALSE)
```