analysis.R Codebook

Jeffrey Strickland

12/20/2021

### Markdown

* This Codebook is an R Markdown document (Knitted). Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents.
* Included embedded R code chunks within the document and output

## Preliminaries

* Library loading and installation if required
* Data retrieval, decomressing and naming assignments

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

getwd()

## [1] "C:/Users/jeff/Documents/R/Getting\_and\_Cleaning\_Data"

filename <- "Coursera\_DS3\_Final.zip"  
  
# Checking if archive already exists.  
if (!file.exists(filename)){  
 fileURL <- "https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip"  
 download.file(fileURL, filename, method="curl")  
}   
  
# Checking if folder exists  
if (!file.exists("UCI HAR Dataset")) {   
 unzip(filename)   
}

## STEP 1. Merging Training and Test Set into one Data Set

* Assign each data to variables
* Merges feature sets
* merges activities code sets
* merges subject sets
* merges all prior sets

#### The features selected for this database come from the accelerometer and gyroscope 3-axial raw signals tAcc-XYZ and tGyro-XYZ.

features <- read.table("UCI HAR Dataset/features.txt", col.names = c("n","functions"))

#### Activities List

List of activities performed when the corresponding measurements were taken and its codes (labels)

activities <- read.table("UCI HAR Dataset/activity\_labels.txt", col.names = c("code", "activity"))

#### Volunteer Test Subject

Contains test data of 9/30 volunteer test subjects being observed

subject\_test <- read.table("UCI HAR Dataset/test/subject\_test.txt", col.names = "subject")

#### Recorder Feature Data

Contains recorded features testing data

x\_test <- read.table("UCI HAR Dataset/test/X\_test.txt", col.names = features$functions)

#### Activities Codes

Contains testing data of activities’code labels

y\_test <- read.table("UCI HAR Dataset/test/y\_test.txt", col.names = "code")

#### Subject Training Set

Contains train data of 21/30 volunteer subjects being observed

subject\_train <- read.table("UCI HAR Dataset/train/subject\_train.txt", col.names = "subject")

#### Feature Training Set

contains recorded features train data

x\_train <- read.table("UCI HAR Dataset/train/X\_train.txt", col.names = features$functions)

#### Acitivities Code Lables

Contains train data of activities’code labels

y\_train <- read.table("UCI HAR Dataset/train/y\_train.txt", col.names = "code")

### Set Merging

* Merges the training and the test sets (x) to create one data set

X <- rbind(x\_train, x\_test)

* Merges the train data of activities’code (y) labels

Y <- rbind(y\_train, y\_test)

* The set Subject is created by merging subject\_train and subject\_test using rbind() function

Subject <- rbind(subject\_train, subject\_test)

* The set Merged\_Data is created by merging Subject, Y and X using cbind() function

Merged\_Data <- cbind(Subject, Y, X)

## STEP2. Measure Extraction

Extracts only the measurements on the mean and standard deviation for each measurement

TidyData <- Merged\_Data %>% select(subject, code, contains("mean"), contains("std"))

## STEP 3. Activity Naming

Uses descriptive activity names to name the activities in the data set

TidyData$code <- activities[TidyData$code, 2]

## STEP 4. Code Labeling Process

Appropriately labels the data set with descriptive variable names \* The code column in TidyData renamed into activities

names(TidyData)[2] = "activity"

* All Acc in column’s name replaced by Accelerometer

names(TidyData)<-gsub("Acc", "Accelerometer", names(TidyData))

* All Gyro in column’s name replaced by Gyroscope

names(TidyData)<-gsub("Gyro", "Gyroscope", names(TidyData))

* All BodyBody in column’s name replaced by Body

names(TidyData)<-gsub("BodyBody", "Body", names(TidyData))

* All Mag in column’s name replaced by Magnitude

names(TidyData)<-gsub("Mag", "Magnitude", names(TidyData))

* All start with character t in column’s name replaced by Time

names(TidyData)<-gsub("^t", "Time", names(TidyData))

* All start with character f in column’s name replaced by `Frequency

names(TidyData)<-gsub("^f", "Frequency", names(TidyData))

* All tbody in column’s name replaced by TimeBody

names(TidyData)<-gsub("tBody", "TimeBody", names(TidyData))

* All -mean() in column’s name replaced by Mean

names(TidyData)<-gsub("-mean()", "Mean", names(TidyData), ignore.case = TRUE)

* All -std() in column’s name replaced by Stand\_Dev

names(TidyData)<-gsub("-std()", "STD", names(TidyData), ignore.case = TRUE)

* All -freq() in column’s name replaced by Frequency

names(TidyData)<-gsub("-freq()", "Frequency", names(TidyData), ignore.case = TRUE)

## STEP 5. Tidy Data Creation & Final Dataset

* 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
* The set FinalData (180 rows, 88 columns) is created by sumarizing TidyData taking the means of each variable for each activity and each subject, after groupped by subject and activity.

FinalData <- TidyData %>%  
 group\_by(subject, activity) %>%  
 summarise\_all(funs(mean))

## Warning: `funs()` was deprecated in dplyr 0.8.0.  
## Please use a list of either functions or lambdas:   
##   
## # Simple named list:   
## list(mean = mean, median = median)  
##   
## # Auto named with `tibble::lst()`:   
## tibble::lst(mean, median)  
##   
## # Using lambdas  
## list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_warnings()` to see where this warning was generated.

### Final File Checking

## str(FinalData) not printed

### File Export FinalData into FinalData.txt file.

write.table(FinalData, "FinalData.txt", row.name=FALSE)

### Print FianlData

FinalData

## # A tibble: 180 x 88  
## # Groups: subject [30]  
## subject activity TimeBodyAccelerome~ TimeBodyAccelerome~ TimeBodyAccelerom~  
## <int> <chr> <dbl> <dbl> <dbl>  
## 1 1 LAYING 0.222 -0.0405 -0.113   
## 2 1 SITTING 0.261 -0.00131 -0.105   
## 3 1 STANDING 0.279 -0.0161 -0.111   
## 4 1 WALKING 0.277 -0.0174 -0.111   
## 5 1 WALKING\_D~ 0.289 -0.00992 -0.108   
## 6 1 WALKING\_U~ 0.255 -0.0240 -0.0973  
## 7 2 LAYING 0.281 -0.0182 -0.107   
## 8 2 SITTING 0.277 -0.0157 -0.109   
## 9 2 STANDING 0.278 -0.0184 -0.106   
## 10 2 WALKING 0.276 -0.0186 -0.106   
## # ... with 170 more rows, and 83 more variables:  
## # TimeGravityAccelerometer.mean...X <dbl>,  
## # TimeGravityAccelerometer.mean...Y <dbl>,  
## # TimeGravityAccelerometer.mean...Z <dbl>,  
## # TimeBodyAccelerometerJerk.mean...X <dbl>,  
## # TimeBodyAccelerometerJerk.mean...Y <dbl>,  
## # TimeBodyAccelerometerJerk.mean...Z <dbl>, TimeBodyGyroscope.mean...X <dbl>,  
## # TimeBodyGyroscope.mean...Y <dbl>, TimeBodyGyroscope.mean...Z <dbl>,  
## # TimeBodyGyroscopeJerk.mean...X <dbl>, TimeBodyGyroscopeJerk.mean...Y <dbl>,  
## # TimeBodyGyroscopeJerk.mean...Z <dbl>,  
## # TimeBodyAccelerometerMagnitude.mean.. <dbl>,  
## # TimeGravityAccelerometerMagnitude.mean.. <dbl>,  
## # TimeBodyAccelerometerJerkMagnitude.mean.. <dbl>,  
## # TimeBodyGyroscopeMagnitude.mean.. <dbl>,  
## # TimeBodyGyroscopeJerkMagnitude.mean.. <dbl>,  
## # FrequencyBodyAccelerometer.mean...X <dbl>,  
## # FrequencyBodyAccelerometer.mean...Y <dbl>,  
## # FrequencyBodyAccelerometer.mean...Z <dbl>,  
## # FrequencyBodyAccelerometer.meanFreq...X <dbl>,  
## # FrequencyBodyAccelerometer.meanFreq...Y <dbl>,  
## # FrequencyBodyAccelerometer.meanFreq...Z <dbl>,  
## # FrequencyBodyAccelerometerJerk.mean...X <dbl>,  
## # FrequencyBodyAccelerometerJerk.mean...Y <dbl>,  
## # FrequencyBodyAccelerometerJerk.mean...Z <dbl>,  
## # FrequencyBodyAccelerometerJerk.meanFreq...X <dbl>,  
## # FrequencyBodyAccelerometerJerk.meanFreq...Y <dbl>,  
## # FrequencyBodyAccelerometerJerk.meanFreq...Z <dbl>,  
## # FrequencyBodyGyroscope.mean...X <dbl>,  
## # FrequencyBodyGyroscope.mean...Y <dbl>,  
## # FrequencyBodyGyroscope.mean...Z <dbl>,  
## # FrequencyBodyGyroscope.meanFreq...X <dbl>,  
## # FrequencyBodyGyroscope.meanFreq...Y <dbl>,  
## # FrequencyBodyGyroscope.meanFreq...Z <dbl>,  
## # FrequencyBodyAccelerometerMagnitude.mean.. <dbl>,  
## # FrequencyBodyAccelerometerMagnitude.meanFreq.. <dbl>,  
## # FrequencyBodyAccelerometerJerkMagnitude.mean.. <dbl>,  
## # FrequencyBodyAccelerometerJerkMagnitude.meanFreq.. <dbl>,  
## # FrequencyBodyGyroscopeMagnitude.mean.. <dbl>,  
## # FrequencyBodyGyroscopeMagnitude.meanFreq.. <dbl>,  
## # FrequencyBodyGyroscopeJerkMagnitude.mean.. <dbl>,  
## # FrequencyBodyGyroscopeJerkMagnitude.meanFreq.. <dbl>,  
## # angle.TimeBodyAccelerometerMean.gravity. <dbl>,  
## # angle.TimeBodyAccelerometerJerkMean..gravityMean. <dbl>,  
## # angle.TimeBodyGyroscopeMean.gravityMean. <dbl>,  
## # angle.TimeBodyGyroscopeJerkMean.gravityMean. <dbl>,  
## # angle.X.gravityMean. <dbl>, angle.Y.gravityMean. <dbl>,  
## # angle.Z.gravityMean. <dbl>, TimeBodyAccelerometer.std...X <dbl>,  
## # TimeBodyAccelerometer.std...Y <dbl>, TimeBodyAccelerometer.std...Z <dbl>,  
## # TimeGravityAccelerometer.std...X <dbl>,  
## # TimeGravityAccelerometer.std...Y <dbl>,  
## # TimeGravityAccelerometer.std...Z <dbl>,  
## # TimeBodyAccelerometerJerk.std...X <dbl>,  
## # TimeBodyAccelerometerJerk.std...Y <dbl>,  
## # TimeBodyAccelerometerJerk.std...Z <dbl>, TimeBodyGyroscope.std...X <dbl>,  
## # TimeBodyGyroscope.std...Y <dbl>, TimeBodyGyroscope.std...Z <dbl>,  
## # TimeBodyGyroscopeJerk.std...X <dbl>, TimeBodyGyroscopeJerk.std...Y <dbl>,  
## # TimeBodyGyroscopeJerk.std...Z <dbl>,  
## # TimeBodyAccelerometerMagnitude.std.. <dbl>,  
## # TimeGravityAccelerometerMagnitude.std.. <dbl>,  
## # TimeBodyAccelerometerJerkMagnitude.std.. <dbl>,  
## # TimeBodyGyroscopeMagnitude.std.. <dbl>,  
## # TimeBodyGyroscopeJerkMagnitude.std.. <dbl>,  
## # FrequencyBodyAccelerometer.std...X <dbl>,  
## # FrequencyBodyAccelerometer.std...Y <dbl>,  
## # FrequencyBodyAccelerometer.std...Z <dbl>,  
## # FrequencyBodyAccelerometerJerk.std...X <dbl>,  
## # FrequencyBodyAccelerometerJerk.std...Y <dbl>,  
## # FrequencyBodyAccelerometerJerk.std...Z <dbl>,  
## # FrequencyBodyGyroscope.std...X <dbl>, FrequencyBodyGyroscope.std...Y <dbl>,  
## # FrequencyBodyGyroscope.std...Z <dbl>,  
## # FrequencyBodyAccelerometerMagnitude.std.. <dbl>,  
## # FrequencyBodyAccelerometerJerkMagnitude.std.. <dbl>,  
## # FrequencyBodyGyroscopeMagnitude.std.. <dbl>,  
## # FrequencyBodyGyroscopeJerkMagnitude.std.. <dbl>