code book of Getting and cleanning data Assignment

Yangtaoyuan

2020/11/17

There is the code book of the Getting and cleanning data project assignment.

1.1 read downloaded file

```
setwd("C:/Users/Administrator/Desktop/coursea/getting_and_cleaning_data/week4/Getting-Cleanning-Project
library(plyr)
```

```
## Warning: package 'plyr' was built under R version 4.0.3
```

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 4.0.3
```

```
subjectTrain = read.table('./train/subject_train.txt',header=FALSE)
xTrain = read.table('./train/x_train.txt',header=FALSE)
yTrain = read.table('./train/y_train.txt',header=FALSE)

subjectTest = read.table('./test/subject_test.txt',header=FALSE)
xTest = read.table('./test/x_test.txt',header=FALSE)
yTest = read.table('./test/y_test.txt',header=FALSE)
```

1.2 Merges the training and testing sets into one dataset called "subjectDataset"

```
xDataSet <- rbind(xTrain, xTest)
yDataSet <- rbind(yTrain, yTest)
subjectDataSet <- rbind(subjectTrain, subjectTest)
dim(xDataSet)
```

```
## [1] 10299   561
```

```
dim(yDataSet)
```

```
## [1] 10299     1
```

```
dim(subjectDataSet)
```

```
## [1] 10299     1
```

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

```
xDataSet_mean_std <- xDataSet[, grep("-(mean|std)\\(\\)", read.table("features.txt")[, 2])]
names(xDataSet_mean_std) <- read.table("features.txt")[grep("-(mean|std)\\(\\)", read.table("features.t
View(xDataSet_mean_std)
dim(xDataSet_mean_std)
```

```
## [1] 10299    66
```

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

```
yDataSet[, 1] <- read.table("activity_labels.txt")[yDataSet[, 1], 2]
names(yDataSet) <- "Activity"
View(yDataSet)
```

4. Appropriately labels the data set with descriptive variable names.

```
names(subjectDataSet) <- "Subject"
summary(subjectDataSet)
```

```
##     Subject
##  Min.   : 1.00
##  1st Qu.: 9.00
##  Median :17.00
##  Mean   :16.15
##  3rd Qu.:24.00
##  Max.   :30.00
```

```
singleDataSet <- cbind(xDataSet_mean_std, yDataSet, subjectDataSet)

# Defining descriptive names for all variables.

names(singleDataSet) <- make.names(names(singleDataSet))
names(singleDataSet) <- gsub('Acc',"Acceleration",names(singleDataSet))
names(singleDataSet) <- gsub('GyroJerk',"AngularAcceleration",names(singleDataSet))
names(singleDataSet) <- gsub('Gyro',"AngularSpeed",names(singleDataSet))
names(singleDataSet) <- gsub('Mag',"Magnitude",names(singleDataSet))
names(singleDataSet) <- gsub('^t',"TimeDomain.",names(singleDataSet))
names(singleDataSet) <- gsub('^f',"FrequencyDomain.",names(singleDataSet))
names(singleDataSet) <- gsub('\\.mean',".Mean",names(singleDataSet))
```

```
names(singleDataSet) <- gsub('\\.std',".StandardDeviation",names(singleDataSet))
names(singleDataSet) <- gsub('Freq\\.',"Frequency.",names(singleDataSet))
names(singleDataSet) <- gsub('Freq$',"Frequency",names(singleDataSet))

View(singleDataSet)
```

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. the data as output as tidydata.txt

```
names(singleDataSet)
```

```
##  [1] "TimeDomain.BodyAcceleration.Mean...X"
##  [2] "TimeDomain.BodyAcceleration.Mean...Y"
##  [3] "TimeDomain.BodyAcceleration.Mean...Z"
##  [4] "TimeDomain.BodyAcceleration.StandardDeviation...X"
##  [5] "TimeDomain.BodyAcceleration.StandardDeviation...Y"
##  [6] "TimeDomain.BodyAcceleration.StandardDeviation...Z"
##  [7] "TimeDomain.GravityAcceleration.Mean...X"
##  [8] "TimeDomain.GravityAcceleration.Mean...Y"
##  [9] "TimeDomain.GravityAcceleration.Mean...Z"
## [10] "TimeDomain.GravityAcceleration.StandardDeviation...X"
## [11] "TimeDomain.GravityAcceleration.StandardDeviation...Y"
## [12] "TimeDomain.GravityAcceleration.StandardDeviation...Z"
## [13] "TimeDomain.BodyAccelerationJerk.Mean...X"
## [14] "TimeDomain.BodyAccelerationJerk.Mean...Y"
## [15] "TimeDomain.BodyAccelerationJerk.Mean...Z"
## [16] "TimeDomain.BodyAccelerationJerk.StandardDeviation...X"
## [17] "TimeDomain.BodyAccelerationJerk.StandardDeviation...Y"
## [18] "TimeDomain.BodyAccelerationJerk.StandardDeviation...Z"
## [19] "TimeDomain.BodyAngularSpeed.Mean...X"
## [20] "TimeDomain.BodyAngularSpeed.Mean...Y"
## [21] "TimeDomain.BodyAngularSpeed.Mean...Z"
## [22] "TimeDomain.BodyAngularSpeed.StandardDeviation...X"
## [23] "TimeDomain.BodyAngularSpeed.StandardDeviation...Y"
## [24] "TimeDomain.BodyAngularSpeed.StandardDeviation...Z"
## [25] "TimeDomain.BodyAngularAcceleration.Mean...X"
## [26] "TimeDomain.BodyAngularAcceleration.Mean...Y"
## [27] "TimeDomain.BodyAngularAcceleration.Mean...Z"
## [28] "TimeDomain.BodyAngularAcceleration.StandardDeviation...X"
## [29] "TimeDomain.BodyAngularAcceleration.StandardDeviation...Y"
## [30] "TimeDomain.BodyAngularAcceleration.StandardDeviation...Z"
## [31] "TimeDomain.BodyAccelerationMagnitude.Mean.."
## [32] "TimeDomain.BodyAccelerationMagnitude.StandardDeviation.."
## [33] "TimeDomain.GravityAccelerationMagnitude.Mean.."
## [34] "TimeDomain.GravityAccelerationMagnitude.StandardDeviation.."
## [35] "TimeDomain.BodyAccelerationJerkMagnitude.Mean.."
## [36] "TimeDomain.BodyAccelerationJerkMagnitude.StandardDeviation.."
## [37] "TimeDomain.BodyAngularSpeedMagnitude.Mean.."
## [38] "TimeDomain.BodyAngularSpeedMagnitude.StandardDeviation.."
## [39] "TimeDomain.BodyAngularAccelerationMagnitude.Mean.."
```

```
## [40] "TimeDomain.BodyAngularAccelerationMagnitude.StandardDeviation.."
## [41] "FrequencyDomain.BodyAcceleration.Mean...X"
## [42] "FrequencyDomain.BodyAcceleration.Mean...Y"
## [43] "FrequencyDomain.BodyAcceleration.Mean...Z"
## [44] "FrequencyDomain.BodyAcceleration.StandardDeviation...X"
## [45] "FrequencyDomain.BodyAcceleration.StandardDeviation...Y"
## [46] "FrequencyDomain.BodyAcceleration.StandardDeviation...Z"
## [47] "FrequencyDomain.BodyAccelerationJerk.Mean...X"
## [48] "FrequencyDomain.BodyAccelerationJerk.Mean...Y"
## [49] "FrequencyDomain.BodyAccelerationJerk.Mean...Z"
## [50] "FrequencyDomain.BodyAccelerationJerk.StandardDeviation...X"
## [51] "FrequencyDomain.BodyAccelerationJerk.StandardDeviation...Y"
## [52] "FrequencyDomain.BodyAccelerationJerk.StandardDeviation...Z"
## [53] "FrequencyDomain.BodyAngularSpeed.Mean...X"
## [54] "FrequencyDomain.BodyAngularSpeed.Mean...Y"
## [55] "FrequencyDomain.BodyAngularSpeed.Mean...Z"
## [56] "FrequencyDomain.BodyAngularSpeed.StandardDeviation...X"
## [57] "FrequencyDomain.BodyAngularSpeed.StandardDeviation...Y"
## [58] "FrequencyDomain.BodyAngularSpeed.StandardDeviation...Z"
## [59] "FrequencyDomain.BodyAccelerationMagnitude.Mean.."
## [60] "FrequencyDomain.BodyAccelerationMagnitude.StandardDeviation.."
## [61] "FrequencyDomain.BodyBodyAccelerationJerkMagnitude.Mean.."
## [62] "FrequencyDomain.BodyBodyAccelerationJerkMagnitude.StandardDeviation.."
## [63] "FrequencyDomain.BodyBodyAngularSpeedMagnitude.Mean.."
## [64] "FrequencyDomain.BodyBodyAngularSpeedMagnitude.StandardDeviation.."
## [65] "FrequencyDomain.BodyBodyAngularAccelerationMagnitude.Mean.."
## [66] "FrequencyDomain.BodyBodyAngularAccelerationMagnitude.StandardDeviation.."
## [67] "Activity"
## [68] "Subject"
```

```
Data2<-aggregate(. ~Subject + Activity, singleDataSet, mean)
Data2<-Data2[order(Data2$Subject,Data2$Activity),]
write.table(Data2, file = "tidydata.txt",row.name=FALSE)
```