run_analysis.R

Md. Ihtesham

Sun Jun 12 02:40:49 2016

```
setwd("C://MyData/UCI HAR Dataset")

library(reshape2)
```

```
## Warning: package 'reshape2' was built under R version 3.2.4
```

```
filename <- "getdata_dataset.zip"

## Download and unzip the dataset:
if (!file.exists(filename)){
  fileURL <- "https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip "
  download.file(fileURL, filename, method="curl")
}
```

```
## Warning: running command 'curl "https://d396qusza40orc.cloudfront.net/
## getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip " -o
## "getdata_dataset.zip"' had status 127

## Warning in download.file(fileURL, filename, method = "curl"): download had
## nonzero exit status
```

```
if (!file.exists("UCI HAR Dataset")) {
  unzip(filename)
}
```

```
## Warning in unzip(filename): error 1 in extracting from zip file
```

```
# Load activity labels + features
activityLabels <- read.table("activity_labels.txt")
activityLabels[,2] <- as.character(activityLabels[,2])
features <- read.table("features.txt")
features[,2] <- as.character(features[,2])

# Extract only the data on mean and standard deviation
featuresWanted <- grep(".*mean.*|.*std.*", features[,2])
featuresWanted.names <- features[featuresWanted,2]
featuresWanted.names = gsub('-mean', 'Mean', featuresWanted.names)
featuresWanted.names = gsub('-std', 'Std', featuresWanted.names)
featuresWanted.names <- gsub('[-()]', '', featuresWanted.names)


# Load the datasets
train <- read.table("train/X_train.txt")[featuresWanted]
trainActivities <- read.table("train/Y_train.txt")
trainSubjects <- read.table("train/subject_train.txt")
```

```
train <- cbind(trainSubjects, trainActivities, train)

test <- read.table("test/X_test.txt")[featuresWanted]
testActivities <- read.table("test/Y_test.txt")
testSubjects <- read.table("test/subject_test.txt")
test <- cbind(testSubjects, testActivities, test)

# merge datasets and add labels
allData <- rbind(train, test)
colnames(allData) <- c("subject", "activity", featuresWanted.names)

# turn activities & subjects into factors
allData$activity <- factor(allData$activity, levels = activityLabels[,1], labels = activityLabels[,2])
allData$subject <- as.factor(allData$subject)

allData.melted <- melt(allData, id = c("subject", "activity"))
allData.mean <- dcast(allData.melted, subject + activity ~ variable, mean)

write.table(allData.mean, "Clean_Data.txt", row.names = FALSE, quote = FALSE)
```