run_analysis.R

jiahong

2020-07-19

```r
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
```

```r
#load all the ori files#
ActivityTest<-read.table("./test/Y_test.txt")
ActivityTrain<-read.table("./train/y_train.txt")
SubjectTest<-read.table("./test/subject_test.txt")
SubjectTrain<-read.table("./train/subject_train.txt")
FeatureTest<-read.table("./test/X_test.txt")
FeatureTrain<-read.table("./train/X_train.txt")

name<-read.table("./features.txt")
Activityindex<-read.table("./activity_labels.txt")

FeatureData<-rbind(FeatureTest,FeatureTrain)
SubjectData<-rbind(SubjectTest,SubjectTrain)
ActivityData<-rbind(ActivityTest,ActivityTrain)

colnames(SubjectData)<-c("subject")
colnames(ActivityData)<- c("activity_index")
colnames(Activityindex)<-c("activity_index","activity")

#extract the mean and std, save the data in ms_data#
sub<-name$V2[grep("mean()|std()", name$V2)]
sub_data<-FeatureData[,sub]
colnames(sub_data)<-sub

data<-cbind(sub_data,SubjectData,ActivityData)
data<- merge(Activityindex, data , by="activity_index", all.x=TRUE)


names(data)<-gsub("std()", "SD", names(data))
names(data)<-gsub("mean()", "mean", names(data))
names(data)<-gsub("^t", "time", names(data))
names(data)<-gsub("^f", "frequency", names(data))
names(data)<-gsub("Acc", "Accelerometer", names(data))
names(data)<-gsub("Gyro", "Gyroscope", names(data))
names(data)<-gsub("Mag", "Magnitude", names(data))
names(data)<-gsub("BodyBody", "Body", names(data))

data_tidy<-aggregate(. ~subject + activity, data, mean)
write.table(data_tidy, file = "homework_tidydata.txt",row.name=FALSE
)
```