linear_regression.R

kora

2020-06-12

```r
library(devtools)
```

```
## Warning: package 'devtools' was built under R version 3.6.3

## Loading required package: usethis

## Warning: package 'usethis' was built under R version 3.6.3
```

```r
library(Biobase)
```

```
## Loading required package: BiocGenerics

## Loading required package: parallel

##
## Attaching package: 'BiocGenerics'

## The following objects are masked from 'package:parallel':
##
##      clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##      clusterExport, clusterMap, parApply, parCapply, parLapply,
##      parLapplyLB, parRapply, parSapply, parSapplyLB

## The following objects are masked from 'package:stats':
##
##      IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':
##
##      anyDuplicated, append, as.data.frame, basename, cbind, colnames,
##      dirname, do.call, duplicated, eval, evalq, Filter, Find, get, grep,
##      grepl, intersect, is.unsorted, lapply, Map, mapply, match, mget,
##      order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
##      rbind, Reduce, rownames, sapply, setdiff, sort, table, tapply,
##      union, unique, unsplit, which, which.max, which.min

## Welcome to Bioconductor
##
##      Vignettes contain introductory material; view with
##      'browseVignettes()'. To cite Bioconductor, see
##      'citation("Biobase")', and for packages 'citation("pkgname")'.
```

```r
library(broom)
```

```
## Warning: package 'broom' was built under R version 3.6.3
```

```r
con =url("http://bowtie-bio.sourceforge.net/recount/ExpressionSets/bodymap_eset.RData")
load(file=con)
close(con)
bm = bodymap.eset
pdata=pData(bm)
edata=as.data.frame(exprs(bm))
fdata = fData(bm)

edata = as.matrix(edata)

#Fit the first gene to the age to yield a linear model
lm1 = lm(edata[1,] ~ pdata$age)
tidy(lm1)
```

```
## # A tibble: 2 x 5
##   term        estimate std.error statistic  p.value
##   <chr>          <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)   2188.      403.       5.43 0.0000888
## 2 pdata$age      -23.2       6.39     -3.64 0.00269
```

```r
plot(pdata$age, edata[1,], col = 2 )
#adding the linear model fit to this dataset
#lwd = line width.
abline(lm1, col =4 , lwd = 4)
```