Homework due Apr 29, 2021 23:01 +03

Distance Exercises #1

1/1 point (graded)

If you have not done so already, install the data package **tissuesGeneExpression**.

```
library(devtools)
install_github("genomicsclass/tissuesGeneExpression")
```

The data represents RNA expression levels for seven tissues, each with several *biological replicates*. We call samples that we consider to be from the same population, such as liver tissue from different individuals, *biological replicates*:

```
library(tissuesGeneExpression)
data(tissuesGeneExpression)
head(e)
head(tissue)
```

How many biological replicates are there for hippocampus?

31

✔ **Answer:** 31

31

Explanation

```
sum(tissue=="hippocampus")
##to answer this question for all tissues look at
table(tissue)
```

Submit

You have used 1 of 5 attempts

Answers are displayed within the problem

Distance Exercises #2

1/1 point (graded)
What is the distance between samples 3 and 45?

152.5662 ✔ **Answer:** 152.5662

152.5662

Explanation

```
sqrt( crossprod(e[,3]-e[,45]) )
## or
sqrt( sum((e[,3]-e[,45])^2 ))
```

Submit You have used 1 of 5 attempts

ℹ Answers are displayed within the problem

Distance Exercises #3

1/1 point (graded)
What is the distance between gene `210486_at` and `200805_at`?

41.01153 ✔ **Answer:** 41.01153

41.01153

Explanation

```
sqrt( crossprod(e["210486_at",]-e["200805_at",]) )
## or
sqrt( sum((e["210486_at",]-e["200805_at",])^2 ))
```

Submit You have used 1 of 5 attempts

Answers are displayed within the problem

Distance Exercises #4

1/1 point (graded)
If I run the command (**don't run it!**):

```
d = as.matrix(dist(e))
```

How many cells (number of rows times number of columns) would this matrix have?

493506225

Answer: 493506225

493506225

Explanation

```
##every pair of rows has an entry:
nrow(e)^2
```

Submit

You have used 2 of 5 attempts

Answers are displayed within the problem

Distance Exercises #5

1/1 point (graded)
Compute the distance between all pairs of samples:

```
d = dist(t(e))
```

Read the help file for `dist()`.

How many distances are stored in `d`? (Hint: What is the length of d)?

17766

✔ **Answer:** 17766

17766

Explanation

```
length(d)
```

Submit

You have used 1 of 5 attempts

Answers are displayed within the problem

Distance Exercises #6

1/1 point (graded)

Why is the answer above not `ncol(e)^2`?

- R made a mistake there
- Distances of 0 are left out
- Because R takes advantage of symmetry: only the lower triangular matrix is stored, thus there are only `ncol(e)*(ncol(e)-1)/2` values.
- Because it is equal to `nrow(e)^2`

✔

Explanation

Note that the distance between samples i and j is the same as distance between samples j and i. Also the distance between a sample and itself is 0. The object returned by `dist()` avoids storing all those values.

Submit

You have used 1 of 2 attempts

Answers are displayed within the problem