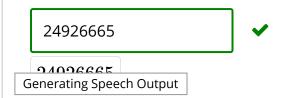edX

Assessment: SummarizedExperiment

Explore a *SummarizedExperiment* object and practice accessing and manipulating data from various components of the object. To complete this assessment, install and load the **airway** package. This package contains a *SummarizedExperiment* object called `airway` containing RNA-seq data from untreated and steroid-treated airway smooth muscle cell lines.
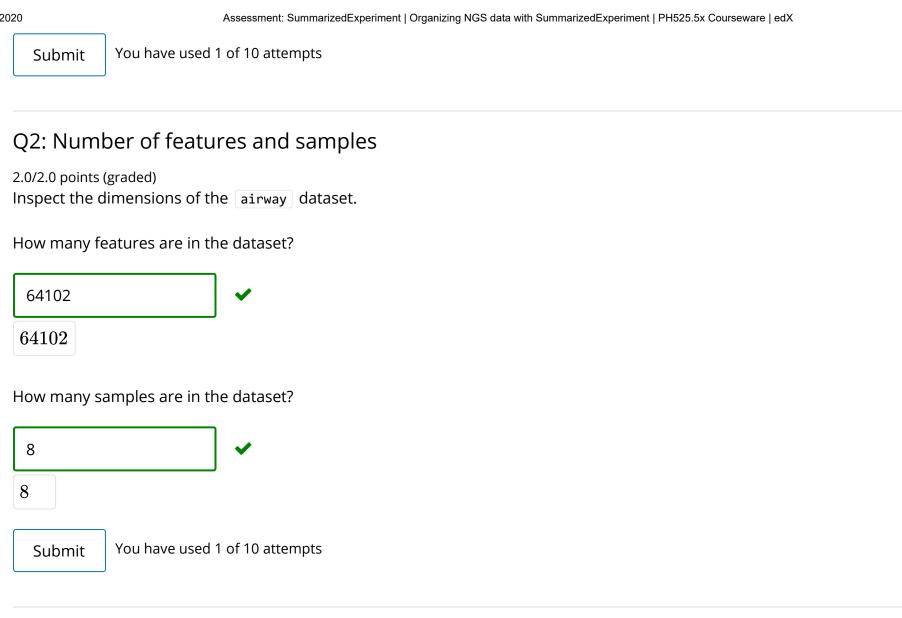
```
library(airway)
data(airway)
airway      # SummarizedExperiment object
```

Q1: Metadata

1.0/1.0 point (graded)
Inspect the metadata for the `airway` dataset.

What is the PubMed ID (PMID) for the original paper supplying these data?

24926665 ✔

24926665

Generating Speech Output

Submit	You have used 1 of 10 attempts

Q2: Number of features and samples

2.0/2.0 points (graded)
Inspect the dimensions of the `airway` dataset.

How many features are in the dataset?

64102	✔

64102

How many samples are in the dataset?

8	✔

8

Submit	You have used 1 of 10 attempts

Q3: Sample data

3.0/3.0 points (graded)
Generating Speech Output

Explore the sample metadata for the `airway` dataset. The columns of the metadata table are given informative abbreviated names, but their exact meaning needs to be inferred.

Which cell line is associated with sample `SRR1039509` ?

N61311 ✔

What is the name of the metadata column specifying whether the sample was treated with dexamethasone (a steroid)?

dex ✔

Which sample (column name) has the shortest average read length?
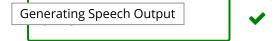
SRR1039513 ✔

Submit   You have used 3 of 10 attempts

Q4: Feature data

2.0/2.0 points (graded)
Explore the feature metadata for the `airway` dataset.

Each feature (row) in the dataset corresponds to a single gene. How many genes are in the dataset?

Generating Speech Output ✔

64102

The row metadata for the `airway` object include *GRanges* specifying the exon boundaries for each gene. How many exons are in the 100th gene?

55 ✔

55

Submit   You have used 1 of 10 attempts

Q5: Exploring a particular feature with GRanges

1.0/3.0 points (graded)
Use *GRanges* operations to explore properties of the 100th gene in the dataset.

On which chromosome is the 100th gene?

7   ✔ Answer: 7

7

Answer code
seqnames(rowRanges(airway)[[100]])

Generating Speech Output ; is the 100th gene (including introns)?

7494   ✖ **Answer:** 78739

7494

Answer code
width(range(rowRanges(airway)[[100]]))

What is the transcription start site (TSS) of the 100th gene?

87031013   ✖ **Answer:** 87109751

87031013

Answer code
resize(range(rowRanges(airway)[[100]]), 1)

Submit   You have used 10 of 10 attempts

ℹ Answers are displayed within the problem

Q6: Assay information for a specific feature

1/4 points (graded)
Consider the gene with ENSEMBL ID "ENSG00000103196".

What is the mean expression level of this gene across all samples?

Generating Speech Output

> 2865.625

✔ **Answer:** 2865.625

\(\)

Answer code

```
mean(assay(airway)["ENSG00000103196", ])
```

What is the mean expression level of this gene in samples treated with dexamethasone?

Use `which()` on the sample information table to identfy which samples have the dexamethasone status "trt".

> 335.22

✖ **Answer:** 4842.75

\(\)

Answer code

```
trt <- mean(assay(airway)["ENSG00000103196", which(colData(airway)$dex == "trt")])
trt
```

What is the mean expression level of this gene in untreated control samples?

Use `which()` on the sample information table to identfy which samples have the dexamethasone status "untrt".

> 349.2917

✖ **Answer:** 888.5

\(\)

Answer code

```
untrt <- mean(assay(airway)["ENSG00000103196", which(colData(airway)$dex == "untrt")])
untrt
```

Generating Speech Output

What is the log (base 2) ratio of mean expression of this gene between treated and untreated samples?

-0.05928798

✖ **Answer:** 2.446383

\(\)

Answer code

```
log2(trt / untrt)
```

Submit   You have used 10 of 10 attempts

ⓘ   Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Section 3 / SummarizedExperiment

Add a Post

Show all posts ▼	by recent activity ▼
❓ **Obtaining Professional Certificate** Do you need to get over a 70% on each assignment for the certificate or an overall average of 70%?	2
💬 **Downloading R 4.0** Is R 4.0 now required? the package airway does not exist on R 3.6	2

Generating Speech Output