**R1: Data science for the scientific life cycle**

**Quiz Type:** Graded Quiz

**Shuffle Answers:** Yes

**Time Limit:** No Time Limit

**Multiple Attempts:** Yes

**Score to Keep:** Latest

**Attempts:** 3

**View Responses:** Always

**Show Correct Answers:**

From {1 MINUTE AFTER DUE} at 12:02am to {ONE WEEK AFTER} at 12:01am

**One Question at a Time:** No

**Require Respondus LockDown Browser:** No

**Required to View Quiz Results:** No

**Webcam Required:** No

Complete the following reading(s) and do your best to think critically while doing so:

*Data science for the scientific life cycle,*by Ezer & Whitaker, *eLife*2019. (Direct PDF download here.)

https://elifesciences.org/articles/43979

https://github.com/NeuralDataScience/Readings/blob/main/Ezer%20and%20Whitaker%20-%202019%20-%20Data%20science%20for%20the%20scientific%20life%20cycle.pdf

After you've completed the reading, complete the reading quiz. These quizzes will help hold you accountable for reading the assigned papers and articles. Then, the discussions in section are meant to really help to ensure you're able to discuss these papers and their nuances with classmates, professors, and any data scientists you work with in your future careers.

**Question 1 (1 pt)**

What distinguishes "data science" in science from just a "rebranding of 'science'"?

***Correct***

The use of many larger datasets for training predictive models.

***Incorrects***

The application of more advanced, non-linear statistical approaches.

Data science is no different from traditional science.

The broad successes of deep learning in applied contexts.

**Question 2 (1 pt)**

What is one way that datasets should be shared to best support developing new methods, generating new hypotheses, and aiding reproducibility?

***Correct***

By making data FAIR.

***Incorrects***

By collecting as much "big data" as possible.

By leveraging cloud storage and compute resources.

By hiring a staff data scientist.

**Question 3 (1 pt)**

What is a major factor that limits the traditional process of hypothesis generation?

***Correct***

How much knowledge an individual can accumulate and retain.

***Incorrects***

The difficulty for NLP algorithms in detecting sarcasm.

The number of FAIR datasets available.

The lack of open sophisticated machine learning / deep learning methods.

**Question 4 (1 pt)**

How can data science be leveraged for improving/optimizing experimental design?

***Correct***

By discovering data collection parameters for optimizing future experiments.

***Incorrects***

By making experiments shareable on GitHub.

By reducing how long experiments need to be run.

By increasing the amount of data that can be collected.

**Question 5 (1 pt)**

How can societal and experimental biases lead to feedback loops that influence data science?

***Correct***

Existing datasets are collected according to scientific fads and trends.

***Incorrects***

By using larger datasets, smaller effects can be uncovered.

Larger, more established labs can analyze data faster, and thus get more grants.

Societal biases have no effect on data science.