Week 1 Quiz

10 questions

1.

Suppose I conduct a study and publish my findings. Which of the following is an example of a replication of my study?



I take my own data, analyze it again, and publish new findings.



An investigator at another institution conducts a study addressing the same question, collects her own data, analyzes it separately from me, and publishes her own findings.



I give my data to an independent investigator at another institution, she analyzes the data and gets the same results as I originally obtained.



An investigator at another institution conducts a study addressing a different scientific question and publishes her findings.

2.

Which of the following is a requirement for a published data analysis to be reproducible?



The investigator's final publication is made available free of charge.



The data analysis is conducted using R.



The analysis is conducted on a variant of the Unix operating system.



The full computer code for doing the data analysis is made publicly available.

3.

Which of the following is an example of a reproducible study?



The study's analytic data and computer code are not publicly available, but the study was simple enough to be repeated by an independent investigator.



The study's analytic data are publicly available, but the computer code is not.



The study's analytic data and computer code for the data analysis are publicly available. When the code is run on the analytic data, the findings are identical to the published results.



The study's original authors re-run their computer code on their analytic data and confirm publicly that the findings match those of the published results.

4.

Which of the following is a reason that a study might NOT be fully **replicated**?



The original investigator does not want to make the analytic data available.



The original study had null findings.



The original study was very expensive and there is no money to repeat it in a different setting.



The original study was conducted by a well-known investigator.

5.

Which of the following is a reason why publishing **reproducible research** is increasingly important?



The statistical methods for most studies can be accurately described using plain language.



Most studies today are small-scale and easily replicated.



Computing power is limited today, making it difficult to apply sophisticated statistical methods.



New technologies are increasing the rate of data collection, creating datasets that are more complex and extremely high dimensional.

6.

What is the role of *processing code* in the research pipeline?



It transforms the analytic data into computational results.



It transforms the computational results into figures and tables.



It transforms the measured data into analytic data.



It conducts the statistical analysis of the primary outcome.

7.

Which is a goal of literate statistical programming?



Combine explanatory text and data analysis code in a single document.



Ensure that data analysis documents are always exported in PDF format.



Separate figures and tables from other data analytic summaries.



Require that data analysis summaries are always written in LaTeX.

8.

What does it mean to *weave* a literate statistical program?



Transform a literate program from R to python.



Transform the literate program into a human readable document.



Transform the literate program into a machine readable code file.



Compress the literate program so that it takes up less space.

9.

Which of the following is required to implement a literate programming system?



A Unix-based computer system.



A program that views PDF files.



A web server for publishing documents.



A programming language like R.

10.

What is one way in which the knitr system differs from Sweave?



knitr lacks features like caching of code chunks.



knitr is written in python instead of R.



knitr was developed by Friedrich Leisch.



knitr allows for the use of markdown instead of LaTeX.

10 questions unanswered

Upgrade to submit