Feedback — Quiz 4: Data Science Technology **Please Note: No Grace Period**

Help Center

Authentication is not required for this quiz.

You submitted this quiz on **Wed 5 Aug 2015 9:01 PM PDT**. You got a score of **8.00** out of **10.00**.

Question 1

Which of the following are required for sharing a data set?

Your Answer		Score	Explanation
○ A tidy data set			
○ An explicit and exact recipe to go from the raw to the tidy data			
◉ All of these options	✔	1.00	
○ A code book describing each variable and its values			
Total		1.00 / 1.00	

Question 2

Which of the following should be included in data tidying recipes?

Your Answer		Score	Explanation
○ Power calculations			
◉ Parameter values for all functions	✔	1.00	

○ Preprocessed data		
○ Units of variables		
Total		1.00 / 1.00

Question 3

What is the central dogma of statistics?

Your Answer		Score	Explanation
○ Using measurements on a population to infer knowledge about a sample			
◉ Using measurements on a probabilistically selected sample to infer knowledge about a population	✔	1.00	
○ Using Bayes rule to calculate probabilities we care about			
○ Estimating parameters using frequencies of observed events			
Total		1.00 / 1.00	

Question 4

Which of the following are types of variability in all genomic data?

Your Answer		Score	Explanation
○ Missing data variability			
○ Geographic variability			
◉ Natural biological variability	✔	1.00	
○ Genetic drift			

Total	1.00 / 1.00

Question 5

Which of the following will increase power in a statistical analysis?

Your Answer		Score	Explanation
○ Increasing measurement variation			
○ Using a new technology			
◉ Decreasing variance	✔	1.00	
○ Adjusting for confounders			
Total		1.00 / 1.00	

Question 6

If 100 p-values are calculated on a data set with no signal, how many p-values would we expect to be less than 0.05 on average?

Your Answer		Score	Explanation
○ 0.05			
○ 50			
○ 20			
◉ 5	✔	1.00	
Total		1.00 / 1.00	

Question 7

If we report 500 results as significant out of 10,000 tests while controlling the family-wise error rate at 5%, about how many false positives do we expect?

Your Answer		Score	Explanation
○ 200			
○ 10			
◉ 500	✖	0.00	
○ 0			
Total		0.00 / 1.00	

Question 8

What is the most common confounder in genomics?

Your Answer		Score	Explanation
◉ Batch effects	✔	1.00	
○ Population stratification			
○ Age			
○ Sex			
Total		1.00 / 1.00	

Question 9

Which of the following can be used to address potential confounders at the experimental design stage?

Your Answer		Score	Explanation
Increasing sample size			
Multiple testing correction	✖	0.00	
Blocking			
Using linear models			
Total		0.00 / 1.00	

Question 10

Which of the following are benefits of making big data as small as possible as soon as possible?

Your Answer		Score	Explanation
Reducing the data will reduce the number of hypothesis tests			
Reduced data will increase the power of statistical tests			
Smaller data sets are easier to share	✔	1.00	
Smaller data sets will decrease false discovery rates			
Total		1.00 / 1.00	