Your answers go on Canvas Test_3_2023.

One answer and only one answer per question. Leaving a question blank or filling in 2+ answers will be incorrect no matter what. (Canvas should not allow you to choose more than 1.)

A = True, B = False unless indicated otherwise. If any part of an answer is incorrect, treat all of it as incorrect. If different parts of an option are inconsistent with each other, consider it incorrect.

Data Quality: Errors and fixes

(RPA = rounding, precision, accuracy; H&T = human & technical; standards = knowns)

1-6. (1 pt each) Identify the type of error indicated. If there is no indication of error in the data, choose (F).

A) Sampling B) Bias C) RPA D) Human and technical (E) Flawed analysis F) None

1. (A)(B)(C)(D)(E)(F). Your bathroom scale consistently reads your weight at 140.7 pounds, but your true weight is 139.3.

2. (A)(B)(C)(D)(E)(F). The average cancer rate calculated from the observed numbers of cancers in a population does not yield the exact underlying cancer rate

3. (A)(B)(C)(D)(E)(F). A person wins the lottery with a single ticket despite the odds being against winning.

4. (A)(B)(C)(D)(E)(F). Use of a blind design yields one set of data, absence of blind gives a very different set of data from the same population. (The data are collected the same way both times except for blind.)

5. (A)(B)(C)(D)(E)(F). Male and female responses to a questionnaire consistently show a large difference across multiple surveys, where the goal is to compare male and female responses.

6. (A)(B)(C)(D)(E)(F). A technician accidentally mixes up samples in a DNA test

7-11 (1 pt each). Identify the type of error indicated <u>as requested or indicated by the underlined text</u>?

A) Sampling B) Bias C) RPA D) Human and technical (E) Flawed analysis F) None

7 (A)(B)(C)(D)(E)(F). A firm uses two methods to assess customer satisfaction of its products. One method uses a postcard included with each product that the customer voluntarily returns anonymously. The other method uses a phone survey of a random sample of customers. The two methods yield the same level of customer satisfaction. <u>What type of error is indicated by the agreement between both sampling methods</u>?

8 (A)(B)(C)(D)(E)(F). A sheet of paper weighs 4 gm to the nearest gram. For the pre-stamped envelope you have, you can mail 40 gm of contents for a fixed rate, so you put in 10 sheets. However, the total weight of contents is actually 43 gm. <u>What error likely explains the discrepancy between the actual total weight and the predicted total weight</u>?

9 (A)(B)(C)(D)(E)(F). Two lab technicians manually code fingerprint images obtained in an investigation for entry into a database, using the same protocol. The characteristics used for fingerprint scoring are discrete (all or none) so there is no ambiguity about how something should be scored. Even though they both work off the very same images as each other, each is unaware of the other's work, and <u>the two technicians enter different numbers for some of the characteristics</u>.

10-12 (3 pts). Read the following paragraph. For each question that follows, indicate whether the ideal data feature is correctly identified in the <u>underlined text</u> that is quoted from the paragraph.

You are trying to determine whether chemical fertilizers yield more vegetables in your garden than does manure (which is organic). You divide your garden into 4 plots. Each of the 4 plots is planted with the same strains of tomato and beans. You flip a coin to choose two of the plots for the manure and use the chemical fertilizer on the other two plots. As produce is harvested, you weigh the amounts from each plot. Periodically, you confirm that the scale is accurate by weighing a known 2-pound piece of steel. At the end of the year, you tally the total produce yield from each of the 4 plots.

(A) The feature is correctly indicated by the underlined text. **(B)** not correctly indicated or absent

10. (A)(B) Replication: <u>choose two of the plots for the manure and use the chemical fertilizer on the other two plots.</u>

11. (A)(B) Standards : <u>You flip a coin to choose</u>

12. (A)(B) Random : <u>You divide your garden into 4 plots</u>

(**13-24). (4pts each set)** For each of the following paragraphs, mark the appropriate letters that describe the data design features present. Mark a data feature only if it is explicitly present at some level in the problem description.

(A) = True/present (B) = False/not described

13-16. Phase III clinical trials of drug efficacy typically use over a thousand patients. Half are randomly assigned to receive the treatment, half to receive a placebo. Blood tests are done by a lab that does not know the status of the patients but ensures data quality by subjecting its staff to blind proficiency tests.

13. (A)(B) Replication **14. (A)(B)** Random **15. (A)(B)** Blind **16. (A)(B)** Standards

17-20. Before subjecting your employees to drug tests, you decide to test the accuracy of the testing lab. You take a sample from yourself, split it into 3 tubes, each with completely different identifying information, and send all 3 tubes for testing by the same lab. You repeat the test of the lab the same way a week later.

17. (A)(B) Replication **18. (A)(B)** Random **19. (A)(B)** Blind **20. (A)(B)** Standards

21-24. A professor wants to know whether ending a lecture with an unsolved problem increases attendance for the following lecture. On odd-numbered class days throughout one month, she ends lecture with an unsolved question, even-numbered days does not; attendance is recorded for each of the following lectures. Her class is not told the purpose of this study, or even that the study is being conducted.

21. (A)(B) Replication **22. (A)(B)** Random **23. (A)(B)** Blind **24. (A)(B)** Standards

25-29. (**5 pts**) Which options identify a valid "fix" for the type of error (or problem) indicated; a "fix" may either reduce that error or allow you to detect that error some of the time. If the error is bias, then the fix must clearly remove the possibility of bias. The (possible) error or problem you need to solve is indicated with underline.

A = the fix is valid; **B** = the fix is not valid or has obvious flaws

The 'Error'	Fix (choose **A** if the fix is valid)
25. The manager of a crime lab is worried that her analytical team is analyzing the data in ways to intentionally support their prejudices about who they think is guilty.	**25. (A)(B)** Ask the analytical team to re-analyze the same data, telling them you want them to be sure of their results.
26. You weigh a single envelope at 4 gm. Even though your scale is known to be accurate, your scale reads only to the nearest gram, so an envelope could weigh between 3.5 and 4.5 grams. You need to know the weight of an envelope to the nearest tenth of a gram.	**26. (A)(B)** Using the same scale, weigh 10 of the same envelopes together and divide the total weight by 10.
27. A drug testing company often accidentally mixes samples and thus gives the wrong result.	**27**. (**A)(B)** Send the lab the same samples twice (for separate analyses) but label them differently so they don't know the samples are the same.
28. You are worried that your breathalyzer is not calibrated correctly and thus gives readings that are consistently low for people who have alcohol in their breath. For example, the machine reads 0.04 when the person is really 0.06. The machine, however, cannot give a value less than 0.0.	**28. (A)(B)** Compare a reading from someone who has imbibed an unknown amount of alcohol to a reading from someone who has imbibed no alcohol.
29. You compare the average caffeine content of Mountain Dew in a sample of 100 bottles produced in Seattle with that of 200 bottles produced in Los Angeles. What can you do to be more confident that sampling error does not account for any difference you find between the Seattle bottles and the Los Angeles bottles?	**29. (A)(B)** Increase the size of the both samples by 100 bottles.

Forensics (Criminal Justice)

30-32. (3 pts) Which options identify a "fix" for the type of error indicated in a protocol; a "fix" may either reduce that error or at least allow you to detect/measure that error.

(A) = The "fix" will correct the problem or identify it **(B**) the fix will not help

30. (A)(B) error: lab falsifies results to give the prosecution its desired results. **Fix**: code samples, leaving names off them; submit samples to the lab without providing any information related to the case.

31. (A)(B) error: unintentional sample mix-up during testing. **Fix**: subject the lab to blind proficiency tests to detect how often mistakes are made

32. (A)(B) error: lab occasionally and accidentally declares matches that are not real, and they often go undetected. **Fix**: code the samples, leaving names off them

33-38. (6 pts) In each of the following, a property of a forensic method, underlined, is described as absent or what it might consist of. The text following the underline gives a possible example of what the underlined text describes. Which examples correctly describe the underlined text?

(A) The non-underlined text correctly describes the underlined.

(B) The underlined and non-underlined do not agree

33. (A)(B) A proficiency test is indicated: over two years, a DNA-typing lab is sent 100 samples of unknown types and is paid $750 per sample because they never fail to deliver results on time.

34. (A)(B) The means for 'independent verification' is present: when at least two experts from the same lab testify about the evidence and report the same results.

35. (A)(B) A relevant reference database for bullet lead analysis would be: individual analyses of bullet lead from thousands of bullets manufactured in the same time frame as the crime.

36. (A)(B) A possible reference database for matching bite marks would be: a description of the characteristics and methods used to score the bite properties.

37. (A)(B) The means for 'independent verification' is present: the technician conducting the method does the analysis twice to be confident of the results

38. (A)(B) A proficiency test is indicated: the lab has been able to convince juries of its results in 50 separate trials.

39-43 (5 pts). Read the following paragraph. Then answer the questions about 'ideal forensic' properties. The underlined text in each question refers to an ideal forensic property that is either present or absent. Other parts of the question explain why the property is present or absent. Choose **A** if the property (or its absence) is correctly explained in the non-underlined text.

The only evidence in a murder trial to connect the defendant to the crime is a match between the duct tape used to bind the victim and a roll of tape found in the defendant's car. The lab providing the evidence based the match on both the total number of fibers in the tape and on the fiber dimensions in the tape. The tape in the roll and on the victim both had 513 fibers in it. Also, the microscopic widths of fibers and spacings between fibers were claimed to be too similar between the victim sample and the suspect's car sample to be from different rolls. The lab analyzed 690 tape rolls from stores in 20 states in establishing the typical similarity between different rolls of tape. Although the defense attempted to find an expert to challenge this testimony, it discovered that no other lab in the world does this kind of analysis, and the lab uses a proprietary analytical method to declare a match (that it keeps secret) so that no one else can compete with its business. This court case is in fact the first time the lab has applied its method in testing whether two samples match.

A = explained correctly B= not explained correctly

39. (A) (B) a reference database that can be screened for a RMP is indicated by the fact that the tape from the roll in the car and from the victim both had 513 fibers.

40. (A) (B) the 690 tape rolls are considered discrete characteristics used for matching because they are all-or-none

41. (A) (B) the widths of fibers and spacings between fibers are characteristics that are not discrete because they involve decimals.

42. (A) (B) The lab was indicated as able to pass blind proficiency tests because it analyzed 690 rolls of tape.

43. (A) (B) Independent verification of a declared match is indicated as NOT being possible because this is the first time the method has been used in court.

44-48. (5 pts). Read the following paragraph. Then answer the questions as indicated.

Only two items of evidence in a trial connect the defendant to the crime. One is a lock found in the defendant's car; the other is an oil spot on the pavement that was dripped on the street in front of the crime scene on the night of the crime. A key found at the crime scene fits the lock found in the car; the company that made the lock testifies that their locks and keys are distinct from all other locks and keys made in the world; from their manufacturing records, it is known that the chance one of their keys would open a 'random' lock of theirs is 1 in 10 million. (A key either fits or does not, there is no intermediate.) Thus, the lock and key connect the defendant to the crime scene. The oil drops left at the scene were analyzed by *Chemlytics Inc.* for trace contaminants and found to match those from the oil in the defendant's car. *Chemlytics Inc.* is well known for chemical analyses of unknown substances, but this is the first case in which it has been requested to decide if two samples match.

Which questions correctly identify the features of 'ideal forensics' that the paragraph indicates as being present? For all but 'independent verification', the problem must specifically describe their presence for it to be present. For 'independent verification' the problem must specifically describe it or describe a means by which independent verification could feasibly be performed by different/independent labs. **A = True (B) = False**

44. (A) (B) a reference database that can be screened for a RMP is indicated for at least one of the analyses

45. (A) (B) some of the characteristics used for the analyses of evidence are discrete

46. (A) (B) some of the characteristics used for the analyses of evidence are not discrete

47. (A) (B) *Chemlytics Inc.*l is described as being able to pass blind proficiency tests for their oil analysis

48. (A) (B) For the lock and key, independent verification of the match would be possible

49-51. (3 pts) Miscellaneous questions about the forensic lectures. **A = True (B) = False**

49. (A) (B) In the eye witness 'test' given to the class, only about 20% of the class identified the correct person in the lineup.

50. (A) (B) In the discussion on eyewitness identification, it was suggested that a photo catalog of potential suspects is a valid reference database when it is used to have the eyewitness find the one picture of the individual they saw.

51. (A)(B) From the video shown, bullet lead analysis satisfied the 'reference database' property of an ideal forensic method