You must turn in this hard copy (with your name on it) and your scantron to receive credit for this exam.

One answer and only one answer per question. Leaving a question blank or filling in 2+ answers will be incorrect no matter what.

A = True, B = False   unless indicated otherwise

Demos and videos

1-7 (8 pts) Which topic most spefically indicates the purpose of a video or demo? If a demo or video was not given in this segment of the class, choose 'none (I).'

(A) Sampling error   (B) Bias   (C) Standards   (D) Data fabrication   (E) Data conversion   (F) Bad forensics

      (G) eye witness fallability   (H) good DNA methods   (I) None

  1. (A)(B)(C)(D)(E)(F)(G)(H)(I)   "it goes to eleven" (Rock Band video)
  2. (A)(B)(C)(D)(E)(F)(G)(H)(I)   Monty Python penguin segment
  3. (A)(B)(C)(D)(E)(F)(G)(H)(I)   Coin flip (virtual or real)
  4. (A)(B)(C)(D)(E)(F)(G)(H)(I)   Choose a random odd number
  5. (A)(B)(C)(D)(E)(F)(G)(H)(I)   Bullet lead analysis ('60 Minutes')
  6. (A)(B)(C)(D)(E)(F)(G)(H)(I)   Terrorist on roof seen through window
  7. (A)(B)(C)(D)(E)(F)(G)(H)(I)   Louie Louie (by the Kingsmen)

Data Quality: Errors and fixes

8-10. (4 pts) There are two phases in most instances of data collection. The first (phase A) involves acquiring the subjects or objects that will be measured. Phase B involves taking the measurements themselves. Which of the following would usually be associated with phase A only, which with phase B only, both, or neither?

  Answer as follows: (A) phase A only (acquisition)
                                   (B) phase B only (making the measurements)
                                   (C) <u>both</u> phase A and B
                                   (D) <u>none</u> of the above

  8. (A) (B) (C) (D) Sampling error

  9. (A) (B) (C) (D) Data obtained with a thermometer

  10. (A) (B) (C) (D) Bias fixed by randomization

11-18 (3.5pts each). In the following questions, indicate which type of error is indicated.

11. The manager of Fiesta wants to know whether fewer customers shop on Fridays falling on the 13th of the month than on other Fridays, hence whether fewer cashiers are needed on Friday the 13th. From records obtained over the last decade, the number of customers on 13th Fridays is 97% that of other Fridays, but the difference is not statistically significant. To what type of error in data do we attribute the observed reduction in customers on Fridays that happen to fall on the 13th?

     A) Sampling   B) Bias   C) RPA   D) Human and technical   E) None

12. A professor has one large class but has two time slots to give an exam. She lets students choose which time to take the exam, so some students volunteer to take the exam during the first time slot, others volunteer for the second. She expects the average score to be the same between both time slots because all students are from the same class. Contrary to her expectation, the average score from the first period is statistically higher than the average for the entire class. The same exam was given at both times and the testing conditions/environment were the same, so the only reasonable interpretation is that the students taking the early exam were better prepared than those taking it later. What type of error is the most plausible explanation for the fact that the average exam score from early time slot was greater than the overall class average?

     A) Sampling   B) Bias   C) RPA   D) Human and technical   E) None

13,14. Joe James bubbled the wrong key code for his exam 1 and thus got a much lower score than if he had bubbled in the correct key code. Use this context to answer the next two questions.

13. What type of error <u>in the machine grading</u> underlies his score being lower than if he had bubbled in the correct key code?

        A) Sampling   B) Bias   C) RPA   D) Human and technical   E) None

14. What type of error underlies Joe's <u>providing the incorrect data (incorrect key code)</u>?

        A) Sampling   B) Bias   C) RPA   D) Human and technical   E) None

15. A grade schooler wishes to measure the average weight of an apple seed, but her scale weighs only to the nearest 0.1 gram. She first weighs each of 100 seeds separately and arrives at an average weight of 0.700 grams. Not being satisfied, she weighs all 100 seeds together and then divides by 100 and obtains 0.763. What type of error is reduced by her second method of calculation?

        A) Sampling   B) Bias   C) RPA   D) Human and technical   E) None

16. A newspaper hires a firm to see if the responses of Republicans differ from those of Democrats to editorials published in the newsaper. The firm takes 3 surveys. Each survey uses 100 randomly chosen Democrats and 100 randomly chosen Republicans. Results of the 3 surveys are that 73%, 70% and 74% of Republicans approve of the editorials, but the corresponding numbers for Democrats is 46%, 45% and 50%. Thus there is a consistently large difference in the response between the two parties. What type of data error is indicated by the consistent difference between the Republican and Democrat responses?

        A) Sampling   B) Bias   C) RPA   D) Human and technical   E) None

17. A newspaper hires a firm to measure the responses of the average reader to editorials published in the newsaper. The firm takes 3 surveys, and for each survey, it sends the same two employees to gather 100 respones each. Unknown to the firm, one employee always polls a random set of Democrats for each survey, the other always polls a random set of Republicans for each survey. When the data are compared between the two employees, one reports 73%, 70% and 74% approval of the editorials, but the other reports 46%, 45% and 50% approval. Thus there is a consistently large difference in the response between the surveys taken by the two employees, even though they should have been representing the same population. What type of data error is indicated by the consistent difference between the responses obtained by the two employees?

        A) Sampling   B) Bias   C) RPA   D) Human and technical   E) None

18. The random choice of 10 exam scores from all 520 Bio301d scores to estimate the average could lead to what type of error?

        A) Sampling   B) Bias   C) RPA   D) Human and technical   E) None

19-22. (6pts) Which options identify a valid "fix" for the type of error indicated; a "fix" should either reliably reduce that error or allow you to measure that error. **A = the fix is valid; B = the fix is not valid**

19. (A) (B) <u>error</u>: a phone survey gets a biased sample of opinions because people with phones have different attitudes than those who lack phones. **Fix**: choose a random subset of opinions obtained in the phone survey

20. (A) (B) <u>error</u>: dog sniffing tests are biased by the dog handler's prior knowledge of the right answer **Fix**: set up a choice test with the dog and handler out of the room so that the right answer is unknown to them

21. (A) (B) <u>error</u>: the scantron misread your form because of a poor erasure **Fix**: have the machine read the forms twice instead of once, and look for a difference in the two readings of each form.

22. (A) (B) <u>error</u>: a random sample of 10 exam scores from class does not give the average score to within one point. **Fix**: replace the first random sample with a second random sample of 10 different exam scores

23-25 (5pts). The following paragraph is a description of a design. In the questions following the paragraph, mark whether the quoted text correctly indicates the data feature is present.
A = the quote indicates the feature, B = the quote does not

An instructor gives two exams during the first part of the semester. Both exams are on Wednesdays, as has been the policy in the past. The class complains that Monday exams would lead to higher scores, on the grounds that students are more rested and have more time to study immediately before Mondays. The instructor bows to class pressure and gives the next two exams on Mondays. Everyone is surprised that the average scores on Monday exams are not higher than the averages for the Wednesday exams.

23 (A) (B) Replication: "gives the next two exams on Mondays"
24 (A) (B) Random: "Both exams are on Wednesdays"
25 (A) (B) Blind: "Everyone is surprised that"

26-30. (**7pts**). Indicate which ideal data features are clearly described:

You decide to test whether sober people can routinely pass the Standardized Field Sobriety Test (SFST), and whether age affects performance. You recruit 200 people of different ages and inform them only that they will be given the SFST, that they must be sober at the time (verified with a breathalyzer test that is calibrated against a blank), and that you are interested in whether men are better than women at passing the test; they are not told about your interest in the effect of age. They are asked to show up in alphabetical order on the same day. The test is administered by officers in uniform that are certified to administer the test and who follow formal test procedures, the actual trials are video taped and verified by others who are also certified.

> **(A)** The feature is indicated. **(B)** not

> **26.** (A)(B) Replication
> **27.** (A)(B) Standards
> **28.** (A)(B) Random
> **29.** (A)(B) Blind at least one way
> **30.** (A)(B) Double blind

31-38. Do-it-yourself protocol. You are conducting an external review/test of a genotyping lab. Your job is to send **two tubes** to the lab, with labels. Several options are given for the content of and label on a tube. You must decide which contents to send and how to label the tubes so that the features of ideal data requested in the question are present. If a tube has a person's name on it, the lab can assume that the tube contents belong to the name of the person on the label and can infer gender. If a tube is labeled with a number, the contents are unknown to the lab but known to you. **?? indicates that you do not know the individual's status for that characteristic.** Your options for tube contents and tube labels are:

tube	tube label	Contents are from	Blood type	Gender	Marker status
(1)	Laura Baker	Laura Baker	AB	Female	negative
(2)	Darin Rokyta	Darin Rokyta	B	Male	??
(3)	Rachael Springman	Rachael Springman	AB	Female	+
(4)	#101	Harry Wichman	A	Male	+
(5)	#218	Patsy Cline	A	Female	negative
(6)	#10	Pam Hines	O	Female	negative
(7)	Jerry Allison	Jerry Allison	A	Male	+
(8)	Brent Iverson	Brent Iverson	B	Male	+

In the following questions, indicate which pairs of tubes (if any) will satisfy the specified criteria.

> **(A)** = criteria are satisfied **(B)** – not satisfied

31-34 (4pts). Choose two tubes to achieve <u>replication of blood type</u> but for which <u>gender is not replicated</u>. You should be able to know that blood type is replicated and that gender is not; marker status is not a factor. Furthermore, the <u>replication of blood type should be blind to the lab</u>. It does not matter if the absence of gender replication is known.

> **31.** (A)(B) tubes 1 & 3
> **32.** (A)(B) tubes 1 & 4
> **33.** (A)(B) tubes 5 & 7
> **34.** (A)(B) tubes 2 & 7

35-38 (4pts). Choose two tubes so that you are certain that <u>none of blood type, gender, marker status is replicated.</u> <u>Absence of replication should be blind to the lab</u>. Assume that the lab will know gender from a name on the tube.

> **35.** (A)(B) tubes 1 & 4
> **36.** (A)(B) tubes 4 & 5
> **37.** (A)(B) tubes 5 & 7
> **38.** (A)(B) tubes 1 & 8

39-41 (4 pts) How were explicit protocols said to be important in data quality?

> **39.** (A)(B) Following a protocol exactly allows bias and H&T error to be eliminated
> **40.** (A)(B) They enable the data to be gathered consistently from one time to the next and among different people who gather the data.
> **41.** (A)(B) They enable someone who did not gather the data to identify types of error that will likely be present.

Criminal Justice

We mentioned 4 features of an 'ideal' forensic method for matching a suspect with a forensic sample: (i) reference database, (ii) discrete characteristics, (iii) independent verification possible, (iv) labs/experts pass blind proficiency tests.

42-45. (6 pts) Which of the following points correctly identify the main purpose, utility or error reduction principle of the feature? **(A)** = true, **(B)** = false

> **42. (A)(B)** <u>Discrete characters</u>: reduces or eliminates bias in scoring

> **43. (A)(B)** <u>Uniform/universal protocol</u>: enables the results to be replicated by other labs

> **44. (A)(B)** <u>Permanent characteristics</u>: permanence is what causes the characteristic to be 'discrete'

> **45. (A)(B)** <u>Reference database</u>: needed to reference the protocols being used by other labs

46. (5pts) Combining sources of error in wrongful matches. A match has been declared between a forensic sample and a suspect. The RMP (random match probability) is calculated as 1 in ten million (= 0.0000001). It has further been revealed that, through mistakes after the lab receives the sample, the lab wrongfully declares matches 1 in 25 times (0.04). In addition, it has been revealed that the crime scene specialists sometimes mix up samples even before sending them to the lab; this mixup by itself leads to an erroneous match at a rate of 1 in 1000 (0.001). There are thus 3 separate reasons why a suspect may not be the source of a sample when a match is declared. Given these data, what is the approximate chance that the suspect is not the source of the sample?

Choose the answer closest to correct – the exact answer is not listed.

> A) 1 in three – the number of types of errors that may have occurred (1/3 or 0.33)

> B) 1 in 25 (0.04).

> C) 1 in a thousand (0.001)

> D) 1 in ten million (0.000001)

> E) 1 in (25 + 1,000 + ten million) = less than 1 in ten million

> F) You cannot combine error rates from different sources

47-50 (**5 pts**) What points can be drawn from the eyewitness test given to the class (this year and in previous years)? The point must both be correct and be a conclusion that can be drawn from the demonstration.
A = true B = false

> **47.** (A) (B) Instructions given to the eyewitness before observing the lineup can greatly affect the chance of misidentification if the observed individual is absent from the lineup.

> **48.** (A) (B) Eyewitness identification has a high error rate (> 10%) even when instructions are given correctly.

> **49.** (A) (B) The use of discrete characteristics in recalling individual appearances improves correct identification rates.

> **50.** (A)(B)   Double-blind lineups, in which the person administering the lineup does not know 'the answer,' improves the correct identification by the eyewitness.

51-59. Which of the 4 features of 'ideal forensics' are indicated 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 labs.

51-54 (6 pts). The method matches bite marks found on a body with the suspect's teeth. Bite marks on the corpse are measured by the depths of indentations and distances between them, matched to the sizes of cusps of the suspect's teeth and distances between them. Only some types of bites are complete enough for the method. Matching is based on an expert's opinion of the similarity between teeth and bite mark and his/her recollection of whether teeth from other cases would have produced a similar bite impression as observed in the present case.

A = present B= absent, incomplete or not used

> **51.** (A) (B) reference database that can be screened for a RMP

> **52.** (A) (B) discrete characteristics

> **53.** (A) (B) pass blind proficiency tests

> **54.** (A) (B) Independent verification (explicitly present or the means for doing it is described)

55-59 (6 pts) A new hair identification method uses the abundances of trace chemical compounds as a type of chemical 'fingerprint' for matching samples. The match is based on similarity of profiles, not identity, because the chemical signatures are quantitative (measured to many decimal places). The method introduced in a court case claims that hair from the suspect matched forensic samples, whereas neither the suspect's hair nor forensic hair matched over 1000 profiles obtained from the population. The chemical fingerprinting method uses an expensive piece of equipment that is found in many chemistry labs, and matching data are obtained from one of the automated runs programmed into the machine at the factory. The statistical analysis uses a package that is widely available from a company called SAS. The company performing the test and providing the evidence in this case routinely processes standards (unknown to the technicians) and has so far gotten the standards correct every time.

A = present B= absent, incomplete or not used

55. (A) (B) a reference database that can be screened for a RMP

56. (A) (B) some of the characteristics used are discrete

57. (A) (B) some of the characteristics used are not discrete

58. (A) (B) The labs mentioned were able to pass blind proficiency tests

59. (A) (B) Independent verification of a declared match is possible (explicitly present or the means for doing it is described)

60. (4 pts.) Exam Key Code A: **Fill in** bubble **(A)** on question 60 to indicate your exam code; leave the other bubbles blank. Also, fill in the correct <u>bubbles</u> for your name and EID on the scantron form.