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

Italicized phrases are true. Do not assume more than is given in a question.

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.

Correlations

1-4 (4 pts) *A cors236 survey used an old, internal email to UI advisors reporting that, across the US, students taking at least 15hr/semester graduated at higher rates than students taking less than 15hr/semester. (You can still access that survey.) The email recommended that advisors should therefore ensure that all their students take at least 15 hr.* Which of the following points are valid conclusions that may be drawn from this recommendation? You may need to rely on class discussion for some questions.
A = True, B = false

1) (A)(B) The advice to advisors is reasonable because 15 hours per semester ensures that a student is taking their education seriously.

2) (A)(B) The advice to advisors is an example of inferring causation from correlation.

3) (A)(B) Forcing more semester hours on those students signed up for less than 15 hours could actually harm a student's progress toward graduation.

4) (A)(B) The advice is <u>not</u> reasonable because it is very likely that students taking 15 or more hours/semester do not have a higher graduation rate than other students.

5-8. (4 pts) Recall the hypothetical table giving accident rate per car per year per 100,000 miles in each cell (except that the numbers have changed):

		Type of car	
		risky	safe
Car color	red	5	3
	not red	10	6

Assume that no other variables are important. Which of the following are true of this table? **(A) = true (B) = false**

5. (A) (B) It is possible to obtain a correlation in which red cars have higher accident rates than other colored cars.

6. (A) (B) Red cars always have lower accident rates than other colored cars when controlling for type of car.

7. (A) (B) Risky cars always have higher accident rates than safe cars when color is controlled.

8. (A) (B) It is not possible with this table to obtain a correlation in which safe cars have higher accident rates than risky cars.

3rd variables

9-12 (4 pts) Last century, *Cannabis* (marijuana) was considered a 'gateway' drug. This gateway model held that marijuana use (which was illegal across the US) actually led or caused the user to be more likely to try other, illegal 'hard' drugs. The evidence was merely that individuals using hard drugs often had a history of marijuana use (a correlation). Which of the following models explain the correlation without *Cannabis* actually causing individuals to try hard drugs – which models invoke a 3rd variable as the cause?

A = a model in which *Cannabis* is NOT the cause of hard drug use (= a 3rd variable model)

B = *Cannabis* causes hard drug use

Choose (A) if *Cannabis* is not a cause of hard drug use	Causal model
9 (A)(B)	Illegal use of *Cannabis* introduces the user to suppliers of illegal drugs that also sell hard drugs. These introductions allow the suppliers to tempt the user into trying hard drugs. Not using *Cannabis* prevents these introductions and thereby prevents hard drug use.
10. (A)(B)	Use of legal drugs such as tobacco/nicotine makes a person more inclined to try illegal mind-altering chemicals. A person who has experienced tobacco is then motivated to try *Cannabis,* and separately, to try hard drugs.
11. (A)(B)	Some young people are rebellious of government control. Illegal *Cannabis* use is one form of rebellion, hard drug use is merely another. Young people who are the most rebellious are inclined to use both *Cannabis* and hard drugs.
12. (A)(B)	*Cannabis* use primes the brain to desire greater levels of mind-alteration. Hard drugs then become more appealing and are thus tried more often.

Controls and controlled variables

13-15. (2 pts) Epidemiologists in Britain observed higher cancer rates in residents living near nuclear power plants (NPP) than in people living away from NPP. The question is whether residence near NPP or something else is the cause of the elevated cancer. Which of the following options would control for age (a 3rd variable) when re-evaluating that correlation?

A = age controlled B = age not controlled

13. (A)(B) Among residents living near NPP, compare cancer rates of old people with cancer rates of young people.

14. (A)(B) Compare cancer rates of old people living near NPP with cancer rates of old people living away from NPP .

15. (A)(B) Compare cancer rates in a random group of residents living near the power plant with cancer rates in a random group of residents living away from the power plant.

16-20. (4 pts) An associate of yours will market a diet to improve learning for college students. You consider the following experimental design to test whether the diet actually improves learning. You are asked about which variables are controlled or manipulated.

Which are true **(A = TRUE, B = false)**

Design. At one university during one semester, you enlist 2000 students enrolled in the same course; all students are taking more than this one course, and their other courses differ among the 2000 students. All are told that the purpose of the study is to test a method of learning improvement that should help them in all their courses, but instead of telling them that the method involves a diet, they are told it involves mental coaching. Students are randomly assigned to two groups, and the groups are housed in different dorms, where they are fed prepared meals. The group in one dorm is fed a normal diet, the other group fed the learning diet. Both groups of students are given the same supposed mental coaching method and told that they are being given the same mental coaching method; no mention is made of their diet, and they are prevented from knowing the other dorm's food. At semester's end, for the one course they all have in common, you compare performance between the two groups to see if there is an effect of the diet.

Which are true of this Design?

16 (A)(B) It controls for diet

17 (A)(B) It controls for initial body weight of the students

18 (A)(B) It controls for expectation of learning improvement.

19 (A)(B) It controls for dorm.

20 (A)(B) Diet is a treatment variable.

21-23 (3 pts) A researcher decides to evaluate the effect of exam difficulty on teaching evaluations at UI. One UI instructor is involved in the experiment. She teaches two sections of a lower division course (same course number and content, one section at 8:30 AM and the other at 11:30 AM, with different students). The 8:30 section is given the harder exams. At the end of the semester, student evaluation scores are compared between the 8:30 and 11:30 sections.

21. In this design, 'exam difficulty' is which of the following?

(A) Treatment variable **(B)** Controlled variable **(C)** Insufficient information

22. In this design, 'instructor' is which of the following?

(A) Treatment variable **(B)** Controlled variable **(C)** Insufficient information

23. In this design, 'average student GPA' (measured before the study) is which of the following?

(A) Treatment variable **(B)** Controlled variable **(C)** Insufficient information

24-27. (**4 pts**) Kerry is a chef. He is developing a sauce for roasted goose, but he does not know what combination of spices to use in the sauce – if any. The base of the sauce is a flour-butter roux, which itself is appealing, but he wonders if adding some combination of spices will improve the roux. The spices he is trying are shallots, thyme, sage, and ginger. Each sauce uses a different combination of spices (or no spices), as indicated by rows (A)-(H) in the following table. A + indicates the spice is present in the sauce, a - indicates absence. There is no *a priori* expectation that spices will improve the appeal of the sauce – the spice-less roux may taste best. He uses several friends to score the appeal of the sauce, as indicated in the last column, given only as letters (so you don't know which combination of spices scored highest, lowest, etc.). Which statements in the following questions are true?

A = TRUE, B = false

		spice				
Sauce		shallots	thyme	sage	ginger	**appeal**
	(A)	+	-	+	-	A_A
	(B)	-	+	-	+	A_B
	(C)	-	+	+	+	A_C
	(D)	-	-	-	-	A_D
	(E)	+	+	+	+	A_E
	(F)	+	+	-	+	A_F
	(G)	-	+	+	+	A_G
	(H)	+	+	-	-	A_H

24. **(A) (B)** The appeal of sauce (D) is expected to be lower than the appeal of the others because it has no spices.

25. **(A) (B)** A comparison of sauces (D) and (E) controls for all spices.

26. **(A) (B)** A comparison of sauce (E) with sauce (G) controls for 3 of the 4 spices.

27. **(A) (B)** At least one pair of sauces allows you to assess the effect of ginger when all other spices are controlled.

28-31 (4 pts) We want to control for as many 3rd variables as possible in asking whether wearing a mask reduces the chance of getting covid. The 3rd variables would include age, social activities, number of people living in the same house, and so on. What can we say about the following potential study designs? Questions apply to the design immediately above them. **(A) = True (B) = false**

Design 1: From thousands of people for which you have data on whether or not they wore masks in the past (by their choice) and whether they got covid, you compare covid rates in a randomly chosen subset of mask-wearers to covid rates in a randomly chosen subset of non-wearers.

28 (A)(B) The randomization controls for third variables that may differ between mask wearers and non-wearers.

29 (A)(B) The extent to which the random choice controls for 3rd variables between wearers and non-wearers depends on how big the randomly chosen subsets are.

Design 2: Now imagine that, for the people for which you have histories of mask wearing and covid, you also have information on lifestyles, gender and age. You compare covid rates between wearers and non-wearers when controlling for all recorded characteristics.

30 (A)(B) This approach is correlational.

31 (A)(B) This approach would control for all possible third variables.

Experiments

32-36 (5 pts). Which of the following studies describe(s) experiments, regardless of whether the experiment was designed well or poorly and regardless of ethics. The goal is given (but not underlined). The question is whether the option describes an experiment for obtaining data with respect to the goal. **(A) = is an experiment** **(B) is not**

32. (A)(B) You have always studied for tests only the night before an exam. To consider whether studying 2 nights before an exam is better, you compare your previous exam scores to previous exam scores of friends (in the same classes as you) who studied 2 nights before tests.

33 (A)(B) To consider whether cheap beer contributes to STD rates in 20-25 year-olds, you compare public records of STD rates in 20-25 year-olds with the average price of popular brands of beer across 16 different states.

34. (A)(B) You are a practicing palm reader. (*Palm reading is a supposed psychic method of telling fortunes*.) You have read the books on how to correctly read palms, and it appears to work – your clients tell you that you are 'right on' with your readings. But you are skeptical and want to see if the reading matters. So you decide to read palms the opposite of what the book says and see if client responses are still just as favorable.

35. (A)(B) You buy new tires for your car because your current tires are balding and will no longer pass inspection.

36. (A)(B) A fisherman is having poor success using artificial lures. In attempting to increase his success, he tries fishing with bait.

37. (-1 pt for getting it wrong) (A) **Key code, name, and ID number**. Provide answer (**A**) on question 37 to indicate your key for this version of the exam. Then upload your answers to the field on BbLearn (test 5 upload) in the usual format:

Format: **last name first name** , last 4 digits of your Vandal ID number , 1 answer 2 answer …