POL 138 Quantitative Reasoning in Political Science
KEY Practice Final Exam Part 1

Directions

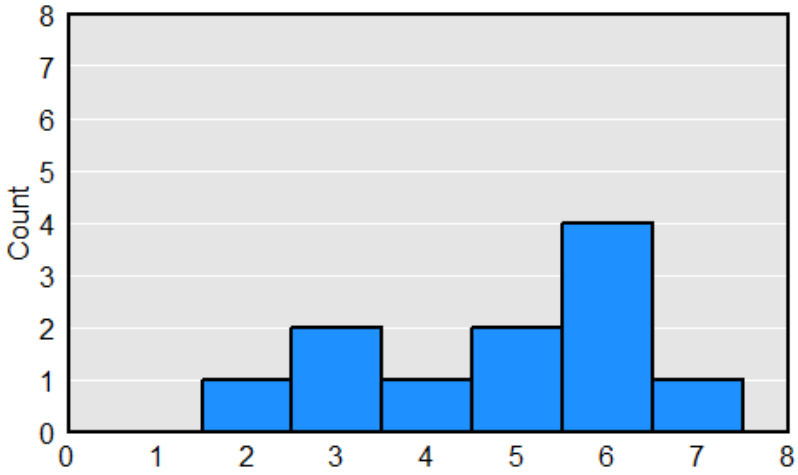
This pdf file contains the multiple-choice items for the Final Exam. Do not upload this pdf file. Instead, place your multiple-choice responses IN CAPITAL LETTERS on the accompanying "POL 138 Final Exam Part 1 submission" Microsoft Word file: either submit the Microsoft Word file with your responses in the table, or submit a photo/scan of a hard copy of the file with your responses written in the table. Also, on the "POL 138 Final Exam Part 1 submission" Microsoft Word file, select the academic honesty box if that applies to you.

1	B	13	A	25	C	37	C	49	A	61	C	73	A	85	D
2	C	14	D	26	D	38	A	50	A	62	B	74	C	86	C
3	C	15	A	27	A	39	A	51	A	63	D	75	C	87	C
4	C	16	C	28	C	40	A	52	B	64	B	76	B	88	A
5	C	17	B	29	D	41	A	53	B	65	C	77	A	89	C
6	A	18	B	30	B	42	B	54	C	66	D	78	A	90	C
7	C	19	A	31	B	43	B	55	B	67	C	79	A	91	B
8	A	20	C	32	C	44	B	56	D	68	D	80	A	92	C
9	A	21	D	33	B	45	A	57	A	69	B	81	A	93	B
10	E	22	A	34	E	46	B	58	A	70	C	82	B	94	D
11	A	23	A/B	35	A	47	D	59	C	71	B	83	D	95	D
12	C	24	A	36	B	48	E	60	C	72	C	84	B	96	D

Reminder about academic honesty

For exams, **students are not permitted to work with others, students are not permitted to use artificial intelligence such as ChatGPT that can generate a response to prompts, and students are not permitted to use materials other than the materials indicated below**. Students are permitted to use a calculator or statistical software, course notes and course videos, the student's handwritten notes, and other hard copy or electronic materials or internet sites other than artificial intelligence such as ChatGPT that can generate a response to prompts. Students are also permitted to ask the instructor to clarify exam items.

1. In the histogram below, which is true?



   A. There are 3 observations of 2.
   B. There are 2 observations of 3.

2. Suppose that a sample has 10 Democrats, 5 Independents, and 7 Republicans. What proportion of the sample is Republican, to two decimal places?
   A. 0.11
   B. 0.19
   C. 0.32
   D. 0.45
   E. 0.92

3. If you score at the 90th percentile on a test, that means that you ___.
   A. got 90 percent of items correct
   B. scored lower than 90 percent of scores
   C. scored higher than 90 percent of scores

4. Political scientists weight survey data for which of the following reasons?
   A. Because the sample is too small
   B. Because the population is much larger than the sample
   C. Because the sample characteristics do not match the population characteristics

5. What is the conventional p-value threshold in political science?
   A. 0
   B. 0.01
   C. 0.05
   D. 0.50
   E. 0.95
   F. 0.99
   G. 1

6. Data below are from seven employees of a company.

Employee	Gender	Pay	Hours	Employee	Gender	Pay	Hours
1	Male	$50	40	5	Female	$50	40
2	Male	$50	40	6	Female	$30	20
3	Male	$50	40	7	Female	$30	20
4	Male	$30	20				

Based on these data, the mean pay is $45.00 among the four male employees and is $36.67 among the three female employees. Which of the following is the gender gap in mean pay among these employees when controlling for the number of hours worked, to the nearest penny?

A. $0.00
B. $8.33
C. $10.00
D. $20.00

7. Of the following, which best describes what a p-value measures?
   A. the precision of an estimate
   B. the size of an association controlling for other model factors
   C. the strength of evidence against the null hypothesis

8. Suppose that, in 2023, 40% of students at a college are women, but that, in 2024, 48% of students at the college are women. That change can be correctly expressed as an increase of ___.
   A. 20 percent
   B. 20 percentage points

9. Men are on average taller than women are. If Amy is at the 50th percentile height among women, then Amy is likely ___ among men.
   A. less than the 50th percentile
   B. at the 50th percentile
   C. greater than the 50th percentile

10. Suppose that a course has two exams: Exam 1 is worth 30% of the overall grade for the course, Exam 2 is worth 70% of the overall grade for the course. If a student scored 50% on Exam 1 and 100% on Exam 2, what would be that student's overall percentage for the course?
    A. 37%
    B. 68%
    C. 72%
    D. 78%
    E. 85%

11. Amy randomly samples 100 students, and Bob randomly samples 200 students. Is the mean height among Bob's students likely to be closer to the population mean, compared to the mean height among Amy's students?
    A. Yes
    B. No

12. Of the following, which term is most appropriate to describe a measure of the extent to which the values of one variable associate with the values of another variable?
    A. a percentile
    B. an inference
    C. a correlation
    D. a standard deviation

13. The random selection from a population ___.
    A. better permits the researcher to make an inference about the population
    B. better permits the researcher to make an inference about whether the treatment had an effect

14. The "Madden Curse" is the name given to the supposed phenomenon in which the NFL player who appears on the cover of the Madden NFL video game in a given year has a bad next season. For example, in the season before appearing on the cover of Madden NFL, Seattle Seahawks running back Shaun Alexander rushed for 1,880 years and had 28 touchdowns, but, the next season, Alexander rushed for only 896 yards and had only 7 touchdowns. Of the phenomena below, which phenomenon below is the most likely reason for any "Madden curse"?
    A. Kelley's paradox
    B. Simpson's paradox
    C. an ecological fallacy
    D. regression toward the mean

15. Suppose that we are estimating the mean gender difference in height between male students and female students. For this study, we randomly select 200 male students and randomly select 300 female students. The fact that the sample of female students is larger than the sample of male students ___.
    A. will not bias our estimate of the mean gender difference
    B. will bias our estimate of the mean gender difference to be smaller than it truly is
    C. will bias our estimate of the mean gender difference to be larger than it truly is

16. In the population of U.S. adults, about 45% are married. If, in our sample of U.S. adults, only 40% is married, which of the following would be the best weight to apply to observations from married persons?
    A. 45 + 40
    B. 45 * 40
    C. 45 / 40
    D. 40 / 45

17. Suppose that, for two participants, a researcher has data from two surveys, one in April and another in June of the same year. For each participant and for both months, the dataset has an indication of the participant's political party (D or R) and an indication of whether the participant supports or opposes affirmative action. Data are below, with each participant identified with an ID:

ID	April	June
1	D + Support	R + Support
2	R + Oppose	D + Oppose

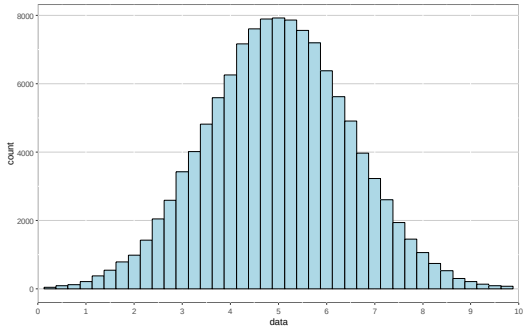
Based on these data only, which of the following is more supported?

A. political party influences attitudes about affirmative action
B. attitudes about affirmative action influences political party

18. If the mean survey weight for a group is 0.9, then that means that the group was ___.

A. undersampled, relative to the group's percentage of the population
B. oversampled, relative to the group's percentage of the population
C. neither undersampled nor oversampled, relative to the group's percentage of the population

19. The image below is an example of a ___.



A. normal distribution
B. uniform distribution

20. Suppose that a test has a mean of 100 and a standard deviation of 15. Scores on the test follow a normal distribution. About 95% of scores should fall within which two scores?

A. 15 and 100
B. 55 and 145
C. 70 and 130
D. 80 and 120
E. 85 and 115
F. 90 and 110

21. Randomly assigning participants to groups helps a randomized experiment identify causes by ___.
    A. eliminating demand effects as much as possible
    B. helping to avoid regression toward the mean
    C. getting the sample to be as representative of the population as possible without weighting
    D. getting the groups to be as similar to each other as possible before the difference in treatment

22. Suppose that scores on a national test follow a normal distribution and have a mean of 100 and a standard deviation of 10. If Student A raises her score from 90 to 100, and Student B raises her score from 120 to 130, which of the following statements is true?
    A. Student A had the higher percentile increase on the test.
    B. Student B had the higher percentile increase on the test.
    C. Student A had the same percentile increase on the test as Student B had.

23. Which of the following is expected to be wider?
    A. the 90% confidence interval for the mean weight of a random sample of 400 U.S. residents
    B. the 99% confidence interval for the mean weight of a random sample of 400 U.S. residents

24. Which of the following is expected to be wider?
    A. the 90% confidence interval for the mean weight of a random sample of 200 U.S. residents
    B. the 90% confidence interval for the mean weight of a random sample of 500 U.S. residents

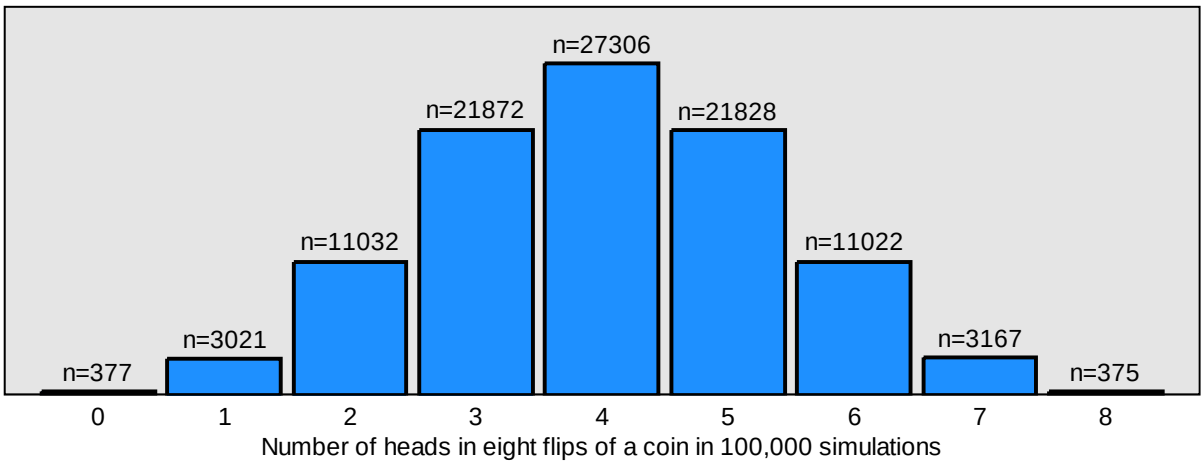
25. Which best indicates what the null hypothesis is?
    A. The hypothesis that the effect is not zero
    B. The hypothesis that is true
    C. The hypothesis being tested
    D. The hypothesis that is most supported by the evidence

26. Which one of these is NOT a necessary step in a randomized experiment involving human participants?
    A. Treat each group differently.
    B. Measure some outcome for each group.
    C. Randomly assign participants to groups.
    D. Use control variables to eliminate alternate explanations.

27. If the p-value is p=0.00001 for a single statistical test of a null hypothesis that there is no association, do we have enough evidence to claim that there is statistically significant evidence for the detected association?
    A. Yes
    B. No

28. If we flipped a coin and got 10 heads and 12 tails, what would be the p-value for a statistical test of the null hypothesis that the coin is fair?
    A. 0
    B. 1
    C. something between 0 and 1

29. The histogram below is from a simulation that had 100,000 trials. For each trial, a fair coin was flipped eight times. The horizontal x-axis indicates each number of times that the fair coin could land on heads in the eight flips, and the height of the columns indicates the number of times the fair coin landed on that number of heads in the eight flips. For example, the n=21872 for the x-axis value of 3 indicates that the fair coin landed on heads exactly 3 times in the 8 flips in 21,872 of the 100,000 trials.



    Based on the above simulation data, which of the following is the closest p-value that would occur if we tested the null hypothesis that our coin is fair, if the coin landed on 7 heads in 8 flips?
    A. 3167 ÷ 100,000
    B. (3167 + 375) ÷ 100,000
    C. (3167 + 3021) ÷ 100,000
    D. (3167 + 375 + 3021 + 377) ÷ 100,000

30. Suppose that we conduct 900 well-designed independent tests of a null hypothesis. In reality, the null hypothesis is true. What is the expected percentage of these tests that are expected to have a p-value of $p<0.05$?
    A. 0%
    B. 5%
    C. 50%
    D. 95%
    E. 100%
    F. Cannot be determined without more information

31. Which score below indicates a higher degree of political knowledge for a political knowledge test?
    A. scoring at the 20th percentile on the test
    B. scoring at the 80th percentile on the test

32. Which one of these, all else equal, associates with a smaller p-value?
    A. a smaller sample size
    B. a smaller estimated effect
    C. a smaller standard deviation

33. Suppose that, in 2023, 40% of students at a college are women, but that, in 2024, 48% of students at the college are women. That change can be correctly expressed as an increase of ___.
    A. 8 percent
    B. 8 percentage points

34. If the p-value for a test of a null hypothesis is p=0.65, then we should do which of the following?
    A. accept the null hypothesis and accept the alternative hypothesis
    B. reject the null hypothesis and reject the alternative hypothesis
    C. accept the null hypothesis and reject the alternative hypothesis
    D. reject the null hypothesis and accept the alternative hypothesis
    E. none of the above

35. Suppose that two researchers are estimating the effect of math homework. Researcher A randomly assigns 500 students to a treatment group that is assigned to complete one math homework assignment per week; the other 500 students in the study are assigned to a control group that gets no math homework assignments to complete. Researcher A then compares end-of-the-year scores on the state math assessment between students in the treatment group and students in the control group. Researcher B instead compares end-of-the-year scores on the state math assessment between the 800 students at a school whose teachers have decided to assign them math homework, to 800 students at the same school whose teachers have decided to not assign students math homework. Which research study has a higher internal validity for estimating the effect of math homework?
    A. Researcher A
    B. Researcher B

36. Discrimination in which unknown information for an individual is estimated based on known or perceived data for the individual's group is referred to as ___.
    A. regression toward the mean
    B. statistical discrimination
    C. taste-based discrimination

37. Which of the following is the name for the phenomenon in which patterns within groups do not hold when the groups are combined?
    A. selection bias
    B. Kelley's paradox
    C. Simpson's paradox
    D. an ecological fallacy

38. Suppose that we are testing a null hypothesis. If we want to be extra certain before rejecting the null hypothesis, which p-value threshold below would be more appropriate?
    A. p=0.01
    B. p=0.10

39. Suppose that unemployment in Freedonia was 12% in 2021, 12% in 2022, and 12% in 2023, and then, on January 1, 2024, Freedonia enacted the Unemployment Reduction Act. Unemployment was then 6% in 2024. Researchers are interested in assessing the extent to which the Unemployment Reduction Act caused this decrease in the unemployment rate. Oceania and Pacifica are countries next to Freedonia and are similar to Freedonia in many ways, but neither Oceania nor Pacifica enacted any legislation to reduce unemployment.

	Unemployment Rate			
	2021	2022	2023	2024
Freedonia	12%	12%	12%	6%
Oceania	14%	14%	14%	15%
Pacifica	10%	10%	10%	12%

    Considering a difference-in-differences method, what do the data in the table above suggest about the decrease in unemployment in Freedonia between 2023 and 2024?
    A. The Unemployment Reduction Act was likely the reason for the decrease in unemployment in Freedonia between 2023 and 2024.
    B. The Unemployment Reduction Act was likely not the reason for the decrease in unemployment in Freedonia between 2023 and 2024.

40. Franco et al. 2014 found that null results were ___ likely to be published than strong results were to be published.
    A. much less
    B. just as
    C. much more

41. Of the p-values below, which p-value is the strongest evidence that an observed difference between the percentage of heads and the percentage of tails from a set of coin flips would have been unlikely to have occurred due to random chance, if the coin is fair?
    A. 0.01
    B. 0.99

42. Recent data indicate that about 12% of U.S. residents receive "food stamp" benefits. Suppose that we conducted a large random survey of U.S. residents that asked participants to respond to the item below:

```
What percentage of U.S. residents receive food stamp benefits?
o  Less than 15%
o  15% to 24%
o  25% to 50%
o  More than 50%
```

If Republicans were less likely than Democrats to select the correct response of "Less than 15%" (with a corresponding p-value of $p<0.05$), would these data be sufficient to conclude that, at least at the time of the survey and at least on average, Republicans were more likely than Democrats were to be misinformed about the percentage of U.S. residents who receive food stamp benefits?

A. Yes, because Republicans were less likely to select the correct response about the percentage of U.S. residents who receive food stamp benefits.
B. No, because the survey item can identify only the misinformation of overestimating the percentage of U.S. residents who receive food stamp benefits, and Democrats might be more likely than Republicans to underestimate the percentage of U.S. residents who receive food stamp benefits.

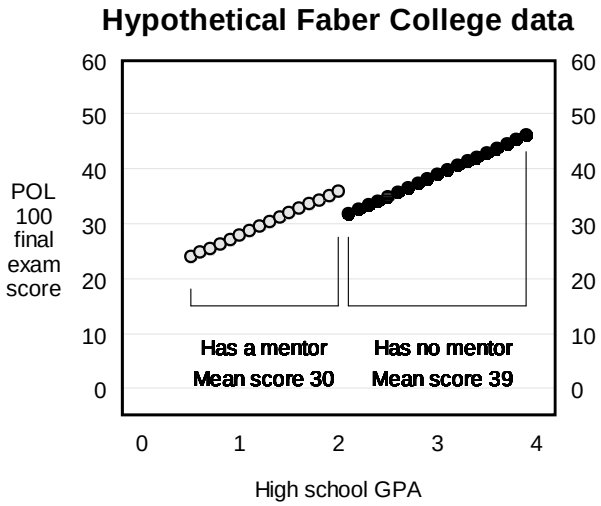
43. Which of the following best indicates what a placebo is?
A. a treatment that has an effect
B. a treatment that has no effect
C. a treatment that has a positive effect
D. a treatment that has a negative effect

44. Which type of validity concerns the ability to make correct claims about the sample?
A. external validity
B. internal validity

45. Suppose that our null hypothesis is that the percentage of U.S. residents who are married is 20 percent. We collect data from a random sample of 1,000 U.S. residents. Of the tests below, which test would be most appropriate for testing that null hypothesis?
A. binomial test
B. Fisher's exact test
C. one-sample t-test
D. two-sample t-test

46. Amy randomly samples 100 students, and Bob randomly samples 200 students. Is the mean height among Bob's students likely to be higher, compared to the mean height among Amy's students?
A. Yes
B. No

47. Suppose that Faber College has a program that assigns freshman students a mentor if and only if the student had a high school GPA that was 2.0 or lower. Each Faber College freshman must take a POL 100 course. The plot below indicates, for each Faber College freshman, the student's high school GPA, the student's final exam score in POL 100, and whether Faber College assigned the student a mentor.



Suppose that we use a discontinuity method to estimate the effect of having a mentor on a student's POL 100 final exam score, by comparing how much the final exam score among students who were just below the threshold for getting a mentor differs from the final exam score among students who were just above the threshold for getting a mentor. Which of the following best indicates that (better) estimate?

A. The mentor reduced final exam scores by about 9 points, on average.
B. The mentor reduced final exam scores by about 5 points, on average.
C. The mentor did not affect final exam scores, on average.
D. The mentor increased final exam scores by about 5 points, on average.
E. The mentor increased final exam scores by about 9 points, on average.

48. Standard deviation is a measure of ___ a set of numbers.

A. the central tendency of
B. the correctness of
C. the reliability of
D. the validity of
E. the variation in

49. Which of these is closest to what an inference is?

A. a conclusion
B. a flawed idea
C. a prediction
D. a reason for a prediction

50. The sample for a research study is better described as which of the following?
    A. the things that are studied
    B. the things that the research study is interested in

51. Which of the following would provide a more credible estimate of the mean political ideology of a population of 3 million persons?
    A. a random sample of 500 members of the population
    B. a non-random sample of 1,000 members of the population

52. If we flipped a coin and got 10 heads and 10 tails, what would be the p-value for a statistical test of the null hypothesis that the coin is fair?
    A. 0
    B. 1
    C. something between 0 and 1

53. A researcher tested the null hypothesis that an association is zero. The p-value for this test p=0.01. Based on this p-value, which of the following should the researcher do, using the conventional level in political science?
    A. conclude that the association is zero
    B. conclude that the association is not zero
    C. neither of the above

54. A researcher tested the null hypothesis that an association is zero. The p-value for this test p=0.19. Based on this p-value, which of the following should the researcher do, using the conventional level in political science?
    A. conclude that the association is zero
    B. conclude that the association is not zero
    C. neither of the above

55. If the p-value is p=0.00001 for a single statistical test of a null hypothesis that there is no association, do we have enough evidence to claim that there is substantively significant evidence for the detected association?
    A. Yes
    B. No

56. Random assignment error in a randomized experiment ___.
    A. cannot bias an estimate
    B. can bias an estimate of an effect only to be lower than it truly is
    C. can bias an estimate of an effect only to be higher than it truly is
    D. can bias an estimate of an effect to be lower than or higher than it truly is

57. For a test of the null hypothesis that there is no association, "statistically significant evidence" for the association refers to sufficient evidence that a particular association ___.
    A. exists
    B. is large

58. Suppose that a researcher conducted a randomized experiment and then compared the mean response from participants in the control group to the mean response from participants in the treatment group. The p-value was p=0.04 for a test of the null hypothesis that these means equal each other. Based on this p-value, the researcher should conclude that ___.
    A. the treatment had an effect
    B. the treatment did not have an effect
    C. there is not enough evidence to conclude that the treatment had an effect

59. Suppose that our null hypothesis is that the mean ideology of U.S. residents will be 5, on a scale from 0 for very liberal to 10 for very conservative. Of the tests below, which test would be most appropriate for testing that null hypothesis?
    A. binomial test
    B. Fisher's exact test
    C. one-sample t-test
    D. two-sample t-test

60. Suppose that, in a randomized experiment, the mean response from participants in the control group differs from the mean response from participants in the treatment group. One reason for this is that participants in the control group were treated differently than participants in the treatment group. The other possible reason why the mean response from participants in the control group differed from the mean response from participants in the treatment group is ___.
    A. a ceiling effect
    B. Simpson's paradox
    C. random assignment error
    D. regression toward the mean

61. After World War II, Germany was split into two countries: communist East Germany, and non-communist West Germany. In the early 1990s, East Germany and West Germany reunified into a single Germany. Suppose that researchers are interested in the effect of communism on personal income. Researcher A compares the mean personal income among persons in 2024 who live in the areas that used to be East Germany to the mean personal income among persons in 2024 who live in the areas that used to be West Germany. Researcher B compares the mean personal income among persons in 2024 who live in the areas that used to be western East Germany to the mean personal income among persons in 2024 who live in the areas that used to be eastern West Germany. An advantage of Researcher B's research design over Researcher A's research design is that ___.
    A. Researcher B will avoid Simpson's paradox
    B. Researcher B will have a smaller sample size
    C. Researcher B will address an alternate explanation
    D. Researcher B will avoid bias due to regression toward the mean

62. The random assignment of participants to groups ___.
    A. better permits the researcher to make an inference about the population
    B. better permits the researcher to make an inference about whether the treatment had an effect

63. How do control variables help improve causal inference in a correlational study?
    A. remove sample bias
    B. randomly assign participants
    C. reduce bias in measurements
    D. help address alternate explanations

64. In 2018, the mean score on the LSAT (the Law School Admission Test) was 154 among political science majors who took the LSAT and was 146 among criminal justice majors who took the LSAT. If the p-value is $p<0.05$ for a test of the null hypothesis that these means equal each other, would that be sufficient evidence at the conventional level in political science to conclude that majoring in political science will, at least on average, cause a higher LSAT score than majoring in criminal justice?
    A. Yes, because the p-value is $p<0.05$.
    B. No, because the analysis should address alternate explanations, such as the possibility that, even before deciding on major, students intending to major in political science already would do better on the LSAT than students intending to major in criminal justice.

65. Data below are from seven employees of a company.

Employee	Gender	Pay	Hours	Employee	Gender	Pay	Hours
1	Male	$60	40	5	Female	$40	40
2	Male	$60	40	6	Female	$20	20
3	Male	$40	20	7	Female	$20	20
4	Male	$40	20				

    Based on these data, the mean pay is $50.00 among the four male employees and is $26.67 among the three female employees. Which of the following is the gender gap in mean pay among these employees when controlling for the number of hours worked, to the nearest penny?
    A. $0.00
    B. $10.00
    C. $20.00
    D. $23.33

66. Suppose that a researcher conducts six independent tests of a true null hypothesis. Which of these is most likely to be the probability that at least one of these tests produces a p-value that is less than $p=0.05$?
    A. 0%
    B. 1%
    C. 5%
    D. 26%
    E. 40%
    F. 100%

67. Of the following, which term best describes the process in which researchers publicly post ahead of time a plan for the research that they will conduct?
    A. Premonition
    B. Replication
    C. Pre-registration
    D. Post-registration

68. Suppose that, in a random sample of 600 Illinois residents, the mean rating about the governor was 60, and the standard deviation of the ratings was 15. Use the formula below to calculate the margin of error for the mean rating about the governor, in which s is the sample standard deviation and n is the sample size.

$$MOE = 1.960 \times \frac{s}{\sqrt{n}}$$

    What is that margin of error, to two decimal places?
    A. 0.58
    B. 0.89
    C. 1.09
    D. 1.20
    E. 1.95

69. Suppose that our null hypothesis is that, in the United States, the percentage of Democrats who support a tax cut equals the percentage of Republicans who support a tax cut. Of the tests below, which test would be most appropriate for testing that null hypothesis?
    A. binomial test
    B. Fisher's exact test
    C. one-sample t-test
    D. two-sample t-test

70. Below are data for two studies:

Study	Estimated effect size	Sample size
A	5	200
B	9	700

    Which of the following is a correct formula for calculating the mean estimated effect size, weighted by sample size?
    A. 5×(200) + 9×(700)
    B. 200×(5/14) + 700×(9/14)
    C. 5×(200/900) + 9×(700/900)
    D. 5×(200/700) + 9×(700/200)

71. An estimated effect from a meta-analysis is typically not calculated as the mean effect size across all studies in the meta-analysis. Instead, larger sample studies tend to receive more weight than smaller samples studies receive. Which of the following better indicates the reason for this extra weight on larger sample studies?
    A. Larger sample studies tend to have larger effect sizes.
    B. Larger sample studies tend to provide more information about the research question.

72. Suppose that we conduct a randomized experiment to estimate the effect of a treatment, but we do not detect sufficient evidence that the treatment differs from zero. One potential reason for this null result is heterogeneous effects, which refers to ___.
    A. socially desirable effects
    B. effects that regress toward the mean
    C. effects that differ between subpopulations

73. Study A has an estimated effect of 5 units and a sample size of 900 participants. Study B has an estimated effect of 10 units and a sample size of 100 participants. Which study should receive more weight in a meta-analysis?
    A. Study A, because Study A has a larger sample size
    B. Study B, because Study B has a larger estimated effect size

74. Sometimes a research study does not permit an inference that the effect of a treatment differs from zero; these inconclusive studies can be referred to as producing a null result. An "informative null" can be used to refer to a study that had a null result but nonetheless provided information that the treatment effect is at most small. Of the measures below, which measure is most useful for assessing whether a result from a study can be appropriately referred to as an informative null?
    A. a p-value
    B. a standard deviation
    C. a 95% confidence interval

75. If we flipped a coin and got 10 heads and 0 tails, what would be the p-value for a statistical test of the null hypothesis that the coin is fair?
    A. 0
    B. 1
    C. something between 0 and 1

76. Which of the following applies to a scenario in which information about a group can reduce errors in predictions about individual members of the group?
    A. selection bias
    B. Kelley's paradox
    C. Simpson's paradox
    D. an ecological fallacy

77. Which of the following is the "blind" element of single-blind peer review of a paper?
    A. Authors are not told the names of the peer reviewers
    B. Peer reviewers are not told the names of the paper's authors

78. Suppose we test whether variation in X causes variation in Y. Which of the following would be worse to add as a predictor to that regression?
A. a variable A that is influenced by X and is influenced by Y
B. a variable B that influences X and influences Y

79. Validity refers to the extent to which a measuring tool ___.
A. measures what the tool is supposed to measure
B. produces consistent results
C. produces statistically significant results

80. Which type of validity concerns the ability of a research result to generalize to the population?
A. external validity
B. internal validity

81. IRBs are organizations designed to protect human subjects in scientific research. Can an IRB, in some circumstances, permit research involving human participants in which the human participants do not provide informed consent?
A. Yes
B. No

82. Suppose that we test whether Drug A has an effect that differs from the effect of a placebo; the p-value is p=0.01 for a test of the null hypothesis that Drug A is just as effective as a placebo. We also test whether Drug B has an effect that differs from the effect of a placebo; the p-value is p=0.40 for a test of the null hypothesis that Drug B is just as effective as a placebo. Is this sufficient evidence to conclude at the conventional level in political science that Drug A has a larger effect than Drug B has?
A. Yes
B. No

83. Suppose that our null hypothesis is that, in the United States, the mean political ideology of women equals the mean political ideology of men. Political ideology is measures on a scale from 0 for very liberal to 10 for very conservative. Of the tests below, which test would be most appropriate for testing that null hypothesis?
A. binomial test
B. Fisher's exact test
C. one-sample t-test
D. two-sample t-test

84. In political science, for peer review of papers that report a statistical analysis, is it typical for the peer reviewers to check the data to see whether the statistical analysis has been correctly conducted?
A. Yes
B. No

85. If the p-value for a test of a null hypothesis is p=0.04, then we should do which of the following?
    A. accept the null hypothesis and accept the alternative hypothesis
    B. reject the null hypothesis and reject the alternative hypothesis
    C. accept the null hypothesis and reject the alternative hypothesis
    D. reject the null hypothesis and accept the alternative hypothesis
    E. none of the above

86. Suppose that a researcher conducted a randomized experiment and then compared the mean response from participants in the control group to the mean response from participants in the treatment group. The p-value was p=0.88 for a test of the null hypothesis that these means equal each other. Based on this p-value, the researcher should conclude that ___.
    A. the treatment had an effect
    B. the treatment did not have an effect
    C. there is not enough evidence to conclude that the treatment had an effect

87. In a non-experimental analysis, omission of a relevant control variable ___.
    A. can bias an estimate of an effect only to be lower than it truly is
    B. can bias an estimate of an effect only to be higher than it truly is
    C. can bias an estimate of an effect to be lower than or higher than it truly is
    D. cannot bias an estimate

88. Suppose that a researcher is interested in the extent to which college causes persons to become more politically liberal. In 2019, the researcher surveys a representative sample of age-18 persons who attend college and a representative sample of age-18 persons who do not attend college. Four years later, in 2023, the researcher surveys each person again. Suppose that the researcher's data is in the table below, in which political ideology is measured from 0 for extremely liberal to 10 for extremely conservative.

Group	Mean ideology at age 18	Mean ideology at age 22
Persons not in college	5.0	4.5
Persons in college	4.5	4.0

    If the researcher used a difference-in-differences design that compared persons in college to persons not in college, the researcher's (better) estimate of the effect of college on political ideology would be that college ___.
    A. made persons in the sample about 0 units more liberal on average
    B. made persons in the sample about 0.5 units more liberal on average
    C. made persons in the sample about 1.0 unit more liberal on average
    D. made persons in the sample about 2.0 units more liberal on average

89. Based on the table below, which group has the highest per capita rate of college graduates?

Group	Population	Number of college graduates
A	900	600
B	900	300
C	800	600

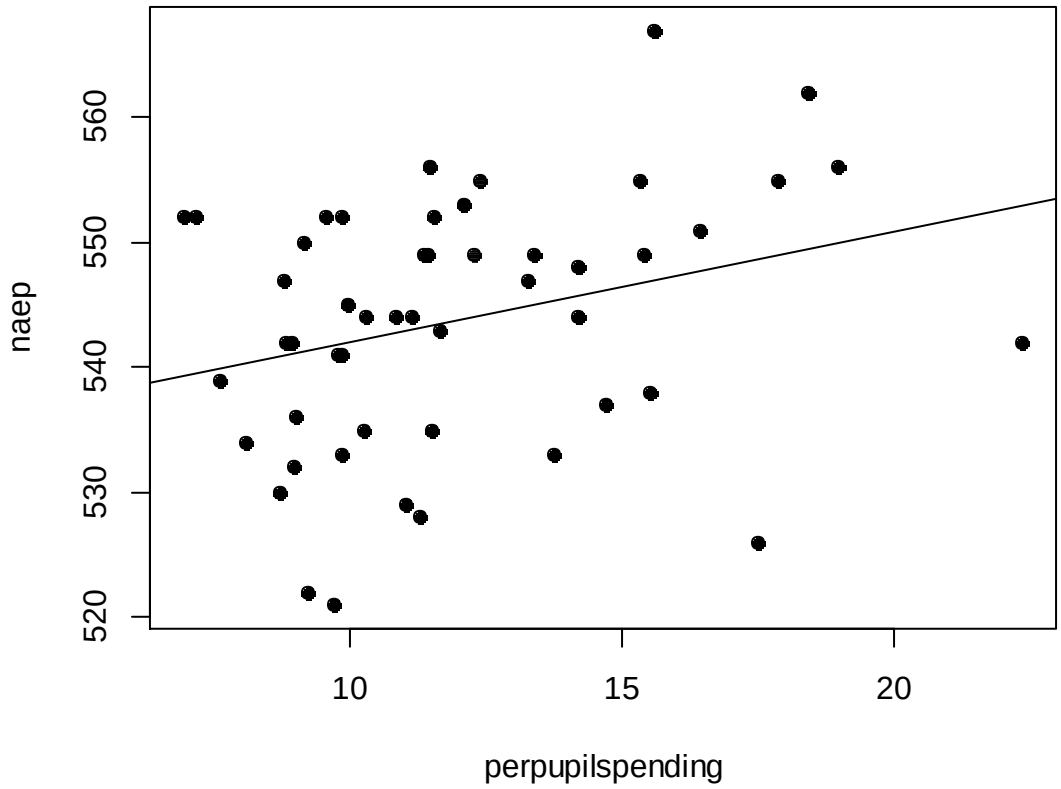
A. Group A
B. Group B
C. Group C

90. Data below are from seven employees of a company.

Employee	**Gender**	**Pay**	**Hours**	**Employee**	**Gender**	**Pay**	**Hours**
1	Male	$70	40	5	Female	$65	40
2	Male	$50	20	6	Female	$65	40
3	Male	$50	20	7	Female	$45	20
4	Male	$50	20				

Based on these data, the mean pay is $55.00 among the three male employees and is $58.33 among the three female employees. Controlling for only the number of hours worked would suggest that ___.
A. there is no gender gap in mean pay
B. there is a gender gap in mean pay that disadvantages male employees on average
C. there is a gender gap in mean pay that disadvantages female employees on average

[Items 91 through 94] The plot below contains one point for each of the 50 states. The x-axis is per pupil spending in the state in thousands of dollars, and the y-axis is the combined math and reading test score in the state on the federal NAEP test.



The output below is from a linear regression using the x variable to predict the y variable.

```
Coefficients:
                 Estimate   p-value
(Intercept)        533.14     < 0.01 *
perpupilspending     0.89       0.04 *
```

91. Does the plot and statistical output contain sufficient evidence to conclude at the conventional level in political science that higher per pupil spending in a state caused that state to have a higher NAEP test score, at least on average?
    A. Yes
    B. No

92. Which of the following is a correct linear regression equation for the output in the image, using X and Y?
    A. Y = 0.89 + 533.14X
    B. Y = X + 533.14 * 0.89
    C. Y = 533.14 + 0.89X

93. The 533.14 intercept means that___
   A. the mean NAEP score was 533.14 across all states
   B. the predicted NAEP score is 533.14 in a state that had zero per pupil spending
   C. the predicted NAEP score is 533.14 in a state that had the mean per pupil spending
   D. for each extra thousand dollars in per pupil spending, the mean NAEP score in a state is predicted to increase by 533.14
   E. the predicted NAEP score is 0 in a state that spend $533.14 per pupil

94. The 0.89 coefficient for perpupilspending means that ___.
   A. the mean NAEP score was 0.89 across all states
   B. the predicted NAEP score is 0.89 in a state that had zero per pupil spending
   C. the predicted NAEP score is 0.89 in a state that had the mean per pupil spending
   D. for each extra thousand dollars in per pupil spending, the mean NAEP score in a state is predicted to increase by 0.89
   E. the predicted NAEP score is 0 in a state that spend $0.89 per pupil

---

[Items 95 and 96] Below is output from a linear regression using data from the ANES 2020 Time Series Study, predicting respondent ratings about feminists (FTFEMINISTS), using a predictor for the political party of the respondent, with categories of Democrat, Republican, and Independent, with Republican as the omitted category.

FTFEMINISTS	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
PARTY						
1. Democrat	31.04	0.68	45.75	0.000	29.71	32.37
3. Independent	15.41	0.70	21.98	0.000	14.04	16.78
_cons	42.89	0.49	86.73	0.000	41.92	43.85

95. What does the 42.89 coefficient estimate for the intercept indicate?
   A. The mean rating about feminists is predicted to be 42.89 among the average respondent.
   B. The mean rating about feminists is predicted to be 42.89 among Democrat respondents.
   C. The mean rating about feminists is predicted to be 42.89 among Independent respondents.
   D. The mean rating about feminists is predicted to be 42.89 among Republican respondents.

96. What does the 31.04 coefficient estimate for the Democrat category indicate?
   A. The mean rating about feminists is predicted to be 31.04 among Democrat respondents.
   B. The mean rating about feminists is predicted to be 31.04 among Republican respondents.
   C. The mean rating about feminists is predicted to be 31.04 among Independent respondents.
   D. The mean rating about feminists is predicted to be 31.04 higher among Democrat respondents than among Republican respondents.
   E. The mean rating about feminists is predicted to be 31.04 higher among Republican respondents than among Democrat respondents.
   F. The mean rating about feminists is predicted to be 31.04 higher among Democrat respondents than among all other respondents.