Exam 1 Stat 100 Spring 2023

EXAM 1: Statistics 100

READ THE DIRECTIONS BELOW TWICE! Form A (Stat) Key

Cover Sheet Questions

1) What's your **name**? _____ _____

(Last name) (First name)

2) What's your **net ID** (email)? _____@illinois.edu

3) Which **section** are you in? ***Circle one:***

i) L2 (In Person Section) **ii)** O1 (Online Section)

This test is ALL multiple choice. **Circle all answers on this exam and fill in the corresponding bubble on your orange scantron**. All questions have exactly one answer. If you circle/bubble in more than one answer, you will automatically be marked wrong. Make sure to circle the answers on this test and fill out your scantron. If you don't do both, you will get a 0.

SCANTRON Directions

- Print and bubble in your LAST NAME with **no spaces** starting in the left most column. Print your FIRST INITIAL in the right-most column.
- Print and bubble in your UIN number in the Student Number box.
- Print and bubble in your NET ID with **no spaces** in the NETWORK ID box.
- *Write Stat 100 on the COURSE line.*
- *Write your instructor's name (Karle Flanagan) on the INSTRUCTOR line.*
- *Write your section (L2 or O1) on the SECTION line.*
- *Sign your name, and right underneath the student signature line PRINT your name.*

READ THIS: Failure to fill out your scantron correctly will result in a loss of 2 points on your exam!

WARNING- The exams look alike but you are sitting next to people who actually have a different version than you. Copying from anyone is equivalent to giving a signed confession.

All cheating including being caught with a non-permissible calculator or formula sheet will result in a 0 and an academic integrity violation on your university record.

Make sure you have all 7 pages including the normal table (63 questions).

There is NO CLASS on Friday this week!

Scores will be posted on Canvas by Monday at noon. Students may pick up their exam in 171 Computing Applications Building during office hours next week.

Exam 1 Stat 100

Questions 1-7 pertain to the following situation. We compared Exam 1 scores of two groups of Stat 100 students: Those who followed the directions and filled out their scantron correctly and those who didn't. The average exam score for the 1300 students who filled out their scantrons correctly was significantly higher than the average exam score of the 100 students who did not.

1. What type of study was this?
   a) a non-randomized experiment **b)** a randomized controlled experiment **c)** an observational study

2. Based only on the given information, can you conclude that filling out scantrons correctly caused the higher exam scores?
   a) No, we can only conclude that it is *associated* with higher exam scores. It definitely couldn't cause them because association rules out causation.
   b) No, we can only conclude that it is *associated* with higher exam scores. It may or may not be the cause.
   c) Yes, we can conclude causation because the data arose naturally. We didn't set out to prove or disprove that how students fill out their scantron mattered so the data is unbiased and can be trusted.
   d) Yes, there's strong evidence that it was the scantron that caused the higher exam scores since everything else was the same.

3. Which technique would be best to control for any potential confounders in this study?
   a) Blocking **b)** Stratification **c)** There should not be any confounders in this study.

Below are either possible causal links, confounders or neither. (Choose answer based only on given info.)

4. Serious Students who follow directions: Serious Students who follow directions are both more likely to do well on exams and more likely to fill out their scantrons correctly.
   a) Causal Link **b)** Confounder **c)** Neither

5. Getting in the proper mind set: Reading the cover sheet carefully and properly filling out your scantron puts you in the proper frame of mind to perform better on the actual exam.
   a) Causal Link **b)** Confounder **c)** Neither

6. Higher Homework Average: Students with a higher homework average also tend to have higher exam scores.
   a) Causal Link **b)** Confounder **c)** Neither

7. Professor Bias: Scantron mistakes are annoying and waste professors' time so they may grade exams with incorrect scantrons more harshly.
   a) Causal Link **b)** Confounder **c)** Neither

Questions 8-11 pertain to the following situation. A group of 500 men on Medicare (age 65+) participated in a study for a new pill to treat high blood pressure. Half of the men were randomly assigned to take the new drug and half were given a plain sugar pill. Neither the patients nor the doctors who evaluated them knew who was in which group. Both groups said they felt better, but there was no difference between the 2 groups in average blood pressure.

8. How would you best describe this study? **a)** Randomized Double-Blind Experiment **b)** Observational Study

9. What kind of bias exists in these results?
   a) Subject Bias- the patients could tell which group they were.
   b) Evaluator Bias- the evaluators could tell which group the subjects were in
   c) Both Subject and Evaluator Bias
   d) No Bias

10. Choose the best conclusion:
    a) Only the drug works. **b)** Only the sugar tablet works. **c)** They both work equally well.

11. Suppose 50 of the men were very old and the researchers want to make sure that these very old men were exactly evenly divided between the treatment and control groups, but they don't want to introduce bias. What should they do?

a) They should divide the men into 2 groups (the 50 men who are very old and the 450 men who aren't). Then randomly assign half of each group to treatment and half to control.

b) They should randomly assign half of the 500 men to treatment and half to control. This will ensure that the men will be evenly divided on all characteristics relevant to the response including age.

c) Randomly assign half of the 500 men to treatment and half to control. Check to see if the very old men are evenly divided. If not, you can rearrange them without introducing bias as long as you do it before treatment starts

Questions 12-15 pertain to the following situation. Would students in Stat 100 learn better if they were allowed Formula Sheets during exams? To answer that question, we did 2 studies.

Study A—We **randomly** assigned half the Stat 100 students to the Formula Sheet Group and half to the No Formula Sheet group.

Study B—We acted like a doctor and "prescribed" formula sheets to those students I thought really needed them and didn't prescribe them to those I thought would do well without them.

All students took the same exams and here are the results:

	Study A--- Randomized		Study B---Non-Randomized	
	# of Students	Average Exam Score	# of Students	Average Exam Score
Formula Sheet Group	500	80%	400	74%
No Formula Sheet Group	500	92%	600	94%

12. Both studies found that the No Formula group did much better than the Formula group, but the randomized design saw only a 12% difference whereas the Non-Randomized Design showed a 20% difference. What possible reason could account for that?

a) In the non-randomized study, we chose the stronger students to be in the No Formula group and the weaker students to be in the Formula group, so the No Formula group did better both because they were stronger students to begin with and because not having a Formula sheet made them learn better.

b) In the non-random study, we showed how tailoring the study method to fit the student works better and therefore causes a more dramatic improvement.

c) In the non-randomized studies, more students were assigned to the No-Formula group than to the Formula group so that could account for the increased difference.

13. Which study is more likely to have confounders? **a)** Study A **b)** Study B **c)** They're equally likely

Judging from both studies, would you conclude that there is good evidence for the following statements?

14. Students learn better when they are allowed formula sheets based on their needs. **a)** Yes **b)** No

15. Students seem to learn better when they are not allowed to rely on formula sheets. **a)** Yes **b)** No

Questions 16. A study was done to compare the effectiveness of high dose vs low dose chemotherapy to treat prostate cancer. The subjects were 2000 adults. Half were randomly assigned to take a high dose pill daily and half assigned to take the low dose pill daily. In every other way the 2 groups received the same medical care. The table below gives the 5-year survival rate for "adherers" and "non-adherers" in the high and low dose groups. Adherers regularly took the drug at least 2/3 of the time while non-adherers took the drug less than 2/3 of the time.

	High Dose		**Low Dose**	
	Number	**5-year survival rate**	**Number**	**5-year survival rate**
Adherers	500	84%	900	78%
Non-Adherers	500	60%	100	68%
Total	1000	72%	1000	77%

16. To assess which dosage is more effective, which two percents in the table above should you compare?

a) 84% vs. 60% **b)** 84% vs. 78% **c)** 84% vs. 77% **d)** 72% vs. 77%

Questions 17-21 pertain to the following study: A Swedish study found that heart attack patients have lower death rates when they eat chocolate. The study compared the records of 1,169 patients recovering from a heart attack and tracked them for eight years. Those who reported eating chocolate regularly were less likely to die after 8 years than those who ate no chocolate. And the more chocolate they ate the higher the benefit.

17. Based only on the information above, this study is an example of…
    a) A randomized controlled double-blind experiment.
    b) A non-randomized controlled experiment with historical controls.
    c) An observational study
    d) A randomized controlled experiment that was not double blind and did not have a placebo.

18. The study reported that they controlled for gender. This means they thought gender might be a confounder, so they eliminated its confounding effect. How did they do that?
    a) At the beginning of the study, they divided the patients into males and females and then randomly divided the males and females equally between the chocolate and no chocolate groups.
    b) At the end of the study, they stratified on gender, and compared the death rate of chocolate eaters to non-chocolate eaters within each gender.
    c) Throughout the study they kept track of those who failed to adhere in both groups whether they were male or female and made sure to compare the death rate of everyone in the original chocolate group to everyone in the original non-chocolate group.

Circle whether the following are confounders, causal links, or neither:

19. Flavonoid Antioxidants: Chocolate contains flavonoid antioxidants that are widely believed to have beneficial cardiovascular effects. **a)** confounder **b)** causal link **c)** neither

20. Eating Disorders – People with eating disorders are less likely to eat chocolate and eating disorders take a toll on your cardiovascular health. **a)** confounder **b)** causal link **c)** neither

21. Chocolate Type- Dark chocolate is said to have more health benefits than white chocolate.
    a) confounder **b)** causal link **c)** neither.

Questions 22-26 pertain to the following list of 6 numbers: 2, -3, -1, 5, 3, 6

22. The average is… **a)** 2 **b)** 2.5 **c)** -1 **d)** 3.33 **e)** none of the above

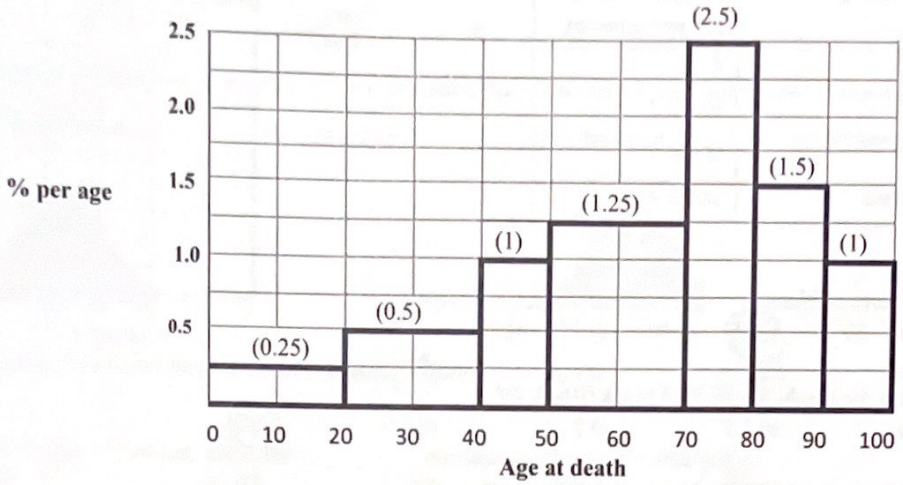
23. The median is… **a)** 2 **b)** 2.5 **c)** -1 **d)** 3.33 **e)** none of the above

24. The **deviations** from the **average** are…
    a) -3, -1, 2, 3, 5, 6 **b)** -0.5, -5.5, -3.5, 2.5, 0.5, 3.5 **c)** 0, -5, -3, 3, 1, 4

25. The sum of the deviations from the average should = _____.
    a) 0 **b)** 1 **c)** 2 **d)** a different answer for every set of 6 numbers

26. Compute the Standard Deviation.
    a) 2 **b)** 10 **c)** 7.75 **d)** 3.16 **e)** 0

Questions 27-36 pertain to the histogram below showing the age at death of a large population. The height of each block is given in parentheses. (Assume an even distribution throughout each interval.)



27. What percent of the population died in the 70-80 interval?
a) 10% **b)** 15% **c)** 20% **d)** 25% **e)** 30%

28. What percent of the population died in the 50-70 interval?
a) 10% **b)** 15% **c)** 20% **d)** 25% **e)** 30%

29. What percent of the population died between the ages of 10 and 20?
a) 2.5% **b)** 5% **c)** 10% **d)** 0.25% **e)** 20%

30. The median is closest to …
a) 40 **b)** 50 **c)** 60 **d)** 70 **e)** 80

31. The median is _____ the average.
a) less than **b)** greater than **c)** equal to **d)** cannot be determined

32. The 25th percentile is….
a) 20 **b)** 25 **c)** 30 **d)** 40 **e)** 50

33. The percent of the population who died at exactly 75 years is closest to…..
a) 1% **b)** 1.25% **c)** 1.5% **d)** 2% **e)** 2.5%

34. If everyone lived 1 more year, the average would ….
a) Increase by 1 year. **b)** Increase by 0.01 years **c)** Increase 15% **d)** Stay the Same **e)** Decrease

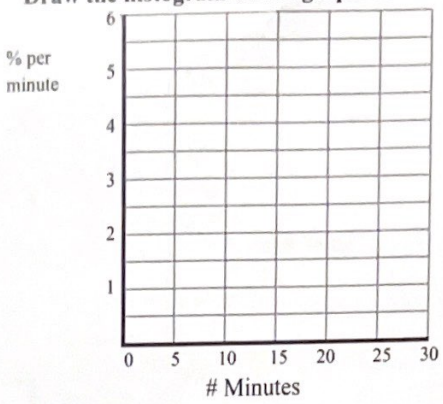
35. If everyone lived 1 more year, the SD would ….
a) Increase by 1 year. **b)** Increase by 0.01 years **c)** Increase 15% **d)** Stay the Same **e)** Decrease

36. If you knew the average and SD of the ages displayed in the histogram above, would it be appropriate to use the normal approximation to figure what percentage of the ages fell within various intervals?
a) Yes, because we know that the histogram represents the age at death of a large population.
b) Yes, because the ages at death range from 0 to 100.
c) No, because the histogram of the ages is not close enough to following the normal curve; it has a long left-hand tail.
d) Maybe, depending on whether the ages were randomly drawn from a larger population.

Questions 37-47 pertain to this situation: 3,000 teenage Americans were asked the question: "How much time, in minutes, elapses between when your alarm first goes off and you get out of bed?" The results are summarized in the table below. Fill in the 4 blanks in the table below:

Minutes	Area %	Height of Block (% per minute)
0-5	20	4
5-15	30	Blank 2
15-20	25	Blank 3
20-30	Blank 1	Blank 4

Draw the histogram on the graph below.



37. What goes in Blank 1?
a) 20 **b)** 25 **c)** 50 **d)** 15 **e)** 10

38. What is the height for the 5-15 block (Blank 2)?
a) 1 **b)** 1.5 **c)** 2 **d)** 2.5 **e)** 3

39. What is the height for the 20-30 block (Blank 4)?
a) 1 **b)** 1.5 **c)** 2 **d)** 2.5 **e)** 3

40. This histogram has a…. **a)** long right hand tail **b)** long left hand tail **c)** neither

41. What is the median number of minutes? **a)** 5 **b)** 15 **c)** 20 **d)** 50

42. The number of people who answered 0-5 minutes is _____ the number of people who answered 15-20 minutes.
a) less than **b)** more than **c)** the same as

Questions 41-45 pertain to this: A list of 10 numbers has an average = 6, median= 4, and SD = 2. Fill out the table below, then answer the questions. Calculate the new average, median, and SD after the original list has been changed.

	New Average	New Median	New SD
4 is added to every number on the original list.	Write a **number.** __Blank 5___	Write a **number.** __Blank 6___	Write a **number.** ___Blank 7____
Every number on the original list is multiplied by **negative** 3.	Write a **number.** __Blank 8___	Write a **number.** __Blank 9____	Write a **number.** __Blank 10____
Change every number to a Z score by subtracting 6 and dividing by 2.	Write a **number.** __Blank 11___	Write a **number.** __Blank 12___	Write a **number.** ___Blank 13_____

43. What goes in Blank 5? **a)** 6 **b)** 4 **c)** 8 **d)** 10 **e)** 14
44. What goes in Blank 9? **a)** -12 **b)** 12 **c)** 4 **d)** -18 **e)** 18
45. What goes in Blank 11? **a)** 6 **b)** 3 **c)** 0 **d)** 2 **e)** impossible to tell
46. What goes in Blank 13? **a)** 2 **b)** -2 **c)** -4 **d)** 0 **e)** 1

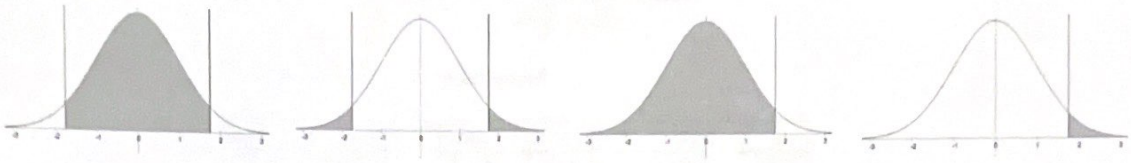
47. If everyone number on the original list remains the same, EXCEPT that 20 is added to the largest number, what happens to the average, median, and SD?
a) They all increase.
b) The average and the median increase, but the SD stays the same.
c) The average and SD increase, but the median stays the same.
d) Only the average increases.
e) None of the above

Questions 48-63 pertain to this situation. According to Bonus Survey 1, the weights of Stat 100 females follow the normal curve with average = 135 pounds and an SD=24 pounds.

What percentage of the female students weigh over 177 pounds?

48. First, convert 177 pounds to a z score.
a) 1 **b)** 1.5 **c)** 1.75 **d)** 2

49. Which histogram's shaded region correctly depicts the percentage of women who weigh over 177 pounds?
a) Histogram A **b)** Histogram B **c)** Histogram C **d)** Histogram D



50. What percentage of women weigh over 177 pounds? **a)** 4% **b)** 8% **c)** 16% **d)** 92%

51. Women who weigh 177 pounds are at the _____ percentile of the weight distribution.
a) 4th **b)** 8th **c)** 16th **d)** 92nd **e)** 96th

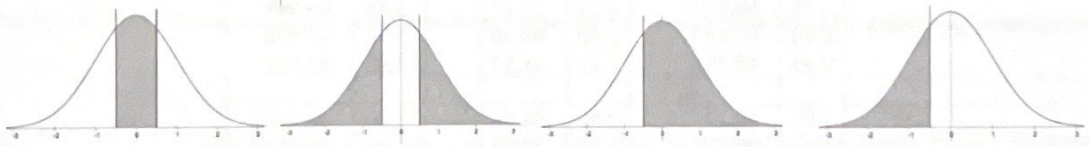
If a student is 0.5 SD's ***below*** average in weight. How much does she weigh and what percentile is she in?

52. What's her z-score? **a)** 0 **b)** -0.5 **c)** 0.5 **d)** 1 **e)** none of the above

53. She weighs _____ pounds. **a)** 135 pounds **b)** 111 pounds **c)** 147 pounds **d)** 123 pounds

54. She's in the _____ percentile **a)** 31st **b)** 34th **c)** 38th **d)** 69th

55. Which histogram's shaded region correctly depicts the percentile?
a) Histogram A **b)** Histogram B **c)** Histogram C **d)** Histogram D



If a student is in the 60th percentile how much does she weigh?

56. To find the z score, you should look at which middle area? **a)** 60% **b)** 40% **c)** 20% **d)** 80%

57. What is the z score for the 60th percentile? **a)** 0.2 **b)** 0.25 **c)** 0.5 **d)** 0.55

58. How much does the student who is in the 60th percentile weigh?
a) 135 pounds **b)** 111 pounds **c)** 141 pounds **d)** 147 pounds

59. What middle area corresponds to the 15th percentile? **a)** 15% **b)** 30% **c)** 35% **d)** 70%

60. If you're below average in weight your z score is? **a)** positive **b)** negative **c)** not enough info

61. If you're exactly at the 50th percentile in weight, then your z score is? **a)** 0 **b)** 1 **c)** 0.5 **d)** not enough info

62. If you're exactly at the 50th percentile in weight, then you weigh?
a) 24 pounds **b)** 135 pounds **c)** 159 pounds **d)** 150 pounds

63. If 2 people have the same Z scores in absolute value but opposite signs, then their percentiles must sum to
a) 0 **b)** 25 **c)** 50 **d)** 100 **e)** Not enough info.