Multiple Choice: (Questions 1 – 20) Answer the following questions on the scantron provided using a #2 pencil. Bubble the response that best answers the question. Each multiple choice correct response is worth 3 points. For your record, also circle your choice on your exam since the scantron will not be returned to you. Only the responses recorded on your scantron will be graded.

1. A researcher was interested in whether Starbucks or Dunkin Donuts is preferred among all college students in the United States. A random sample of 1,028 students from universities across the country were asked which of the two they preferred. The results of the survey revealed that 58% of the students preferred Starbucks over Dunkin Donuts.

   Which of the following statements is <u>not</u> true?

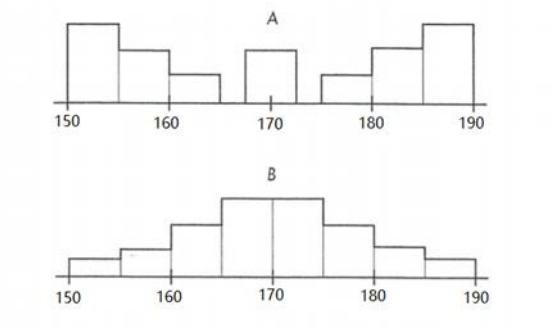
   A. The population of interest is all college students in the United States.

   B. This is an example of a designed experiment.

   C. The type of data collected is qualitative and nominal.

   D. 58% is a statistic.

2. Carefully observe the following two data sets.



   Which of these statements about the data sets is / are true?

   I. Both data sets have the same mean.

   II. Both data sets are symmetric.

   III. Both data sets have the same mode.

   A. I only
   B. I and II
   C. I and III
   D. All of the statements are true

Below is the frequency distribution of the Gas Prices for 20 local gas stations. Use this information to answer questions 3 and 4.

Gas Price ($/Gallon)	Number of Gas Stations
$3.50 - $3.74	3
$3.75 - $3.99	5
$4.00 - $4.24	7
$4.25 - $4.49	5

3. **What is the cumulative relative frequency of the $3.75 - $3.99 bin?**

A. 0.40
B. 5
C. 8
D. 0.25

4. **What percentage of gas stations have a gas price of exactly $4.00?**

A. 0%
B. 35%
C. 1%
D. Cannot be determined

5. Zip codes in the Clemson area include 29630, 29631, 29670, and 29634.

What type of variable describes zip codes?

A. Quantitative and Discrete
B. Qualitative and Interval
C. Qualitative and Nominal
D. Quantitative and Continuous

6. A local newspaper reporter wanted to ask Clemson fans about their opinions on Clemson football team's recent performance against Texas A&M. She chose for the sample 100 students wearing Clemson Tigers T-shirts.

What type of sampling method was used?

A. Simple Random Sampling
B. Convenience Sampling
C. Systematic Sampling
D. Cluster Sampling

7. A Fertilizer Company is interested in determining the proportion of people who need new fertilizer for their garden in Imalone, Wisconsin. The company representative Anna Cornkernel decides to poll a random sample of 120 households in town.

Identify the population of interest.

A. The 120 households polled by Anna
B. All households in the U.S.
C. All households in Imalone, Wisconsin
D. The gardens getting the fertilizer

8. In 2018, YouGov surveyed 8,125 U.S. adults asking them whether they believe the Earth is flat. Of the respondents, 2% reported that they "have always believed the Earth is flat."

Which of the following statements about this study is correct?

A. The population is all U.S. adults and the parameter is 2%.
B. The population is the 8,125 U.S. adults and the statistic is 2%.
C. The sample is the 8,125 U.S. adults and the statistic is 2%.
D. The sample is all U.S. adults and the statistic 2%.

9. Steve has had a relatively rough semester. In one course, the scores from his first three exams are 65, 72, and 47. In order to maintain his financial aid, Steve must maintain an exam average of 70. Assume that there are four equally weighted exams in Steve's course.

What score does Steve need on the fourth exam in order to achieve an exam average of 70?

A. Cannot be determined from the data provided.
B. 70
C. 96
D. 80

Use the following information to answer questions 10 and 11.

Suppose the following relative frequency distribution represents the proportion of majors for STAT 3090 students.

Major	**Relative Frequency**
Business	0.429
Finance	0.179
Computer Science	0.146
Marketing	0.211
Other	0.035

10. There are 950 STAT 3090 students during the Fall 2019 semester.

Approximately how many of these students are Computer Science Majors?

A. 146
B. 811
C. 139
D. 475

11. **Which of the following graphs is most appropriate for this data and why?**

A. Since the variable Major is quantitative, we should use a histogram.
B. Since the variable Major is qualitative, we should use a pie chart.
C. Since the variable Major is qualitative, we should use a stem and leaf plot.
D. Since the variable Major is quantitative, we should use a dot plot.

12. Björg was asked by his employer to report a five number summary for a sample.

Which of the following statistics should <u>not</u> be included in a five number summary?

A. The smallest value
B. The largest value
C. The median
D. The mean

13. At a local high school, GPA's have a bell or mound shaped symmetric distribution with a mean of 2.9 and standard deviation of 0.6.

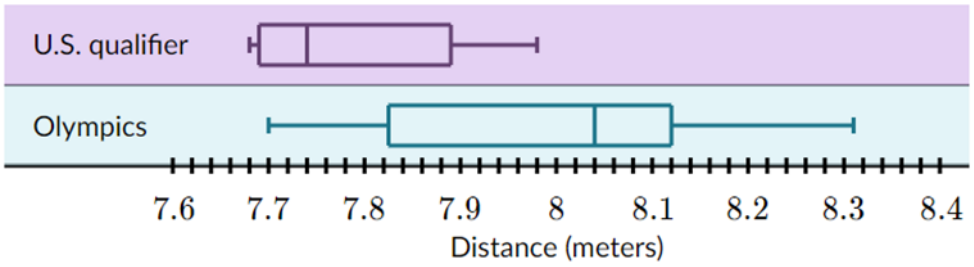
What percentage of students at the high school have a GPA between 2.3 and 3.5?

A. 68%
B. 99.7%
C. 95%
D. 84%

14. **Which symbol is used to represent the standard deviation of a sample?**

A. σ^2
B. μ
C. s
D. $\bar{x}$

15. Below are two box plots from Men's Long Jump. The upper box plot shows the distance (in meters) for the top 12 men's long jumpers at the U.S. qualifying meet. The lower box plot shows the distance (in meters) achieved in the men's long jump at the 2012 Olympic games.



Select the correct statement about these box plots.

A. The longest jump in the U.S. qualifier was shorter than the median jump in the Olympics.
B. The IQR of the U.S. qualifier is approximately 0.3 meters and the IQR of the Olympics is 0.6 meters
C. All of the Olympic jumps were further than all of the U.S. qualifier jumps.
D. The median of the U.S. jumps is above the 1st quartile of the Olympic jumps.

16. Before applying to undergraduate studies in the US, non-native speakers of English need to take exams called the TOEFL (Test of English as a Foreign Language) or IELTS (International English Language Testing System). Here are some summary statistics for each exam:

Exam	Mean	Standard Deviation
TOEFL	98	12.5
IELTS	6.2	1.2

Olga took both tests. She scored 102 on the TOEFL and 7.2 on the IELTS. **On which test did she score better relative to others who took the test? (Hint: is there a way to standardize the test scores?)**

A. TOEFL.
B. IELTS.
C. Not enough information.
D. She did equally well in both exams.

17. **Which of the following statements best describes the shape of the distribution of the following data?**

Stem	Leaf
1	1
2	5 6
3	0 3
4	7 8
5	0 1 8
6	5 7 8 8
7	0 0 0 1 2 4 4 4
8	5 5 6

Key 1|1 = 1.1

A. Skewed left
B. Skewed right
C. Symmetric with one mode
D. Symmetric with more than one mode

18. The following table displays the 100 senators of the 114th U. S. Congress on June 23, 2016, classified by political party affiliation and gender.

	Male	**Female**	**Total**
Democrat	30	14	44
Republican	48	6	54
Independent	2	0	2
Total	80	20	100

A senator is selected at random from this group. What is the probability that the senator is a Democrat or a female?

A. 0.14
B. 0.50
C. 0.64
D. 0.74

19. 175 college students were asked what their favorite Marvel movie was. The results are summarized in the table below.

	Freshman	Sophomore	Junior	Senior
Avengers: Endgame	12	14	19	6
Thor: Ragnarok	4	20	3	12
Black Panther	8	1	0	9
Captain America: Civil War	6	16	10	10
Spiderman: Homecoming	8	0	5	12

Suppose that a student from this group is randomly selected. **Are the events "The Student is a Freshman" and "The Student's Favorite Movie was Black Panther" mutually exclusive?**

A. No, they are not mutually exclusive since the probability of selecting a student who is a Freshman and who liked Black Panther is equal to 0.046.

B. Yes, they are mutually exclusive since the probability of selecting a student who is a Freshman is 0.217 and the probability of selecting a student who liked Black Panther is only 0.103.

C. No, they are not mutually exclusive since the probability of selecting a student who is both a Freshman and their favorite movie is Black Panther is equal to 0.

D. Yes, they are mutually exclusive since the probability of selecting a student who is both a Freshman and their favorite movie is Black Panther is not equal to 0.

20. Test scores for a small class of statistics students are recorded below.

90 82 100 43 78 99 66 78 50 74

What is the range of these test scores?

A. 76

B. 57

C. 19

D. 43

(This space was intentionally left blank. **Exam continues on next page.**)

Free Response: The Free Response questions will count as 39% of your total grade. Read each question carefully. In order to receive full credit, you must show logical (relevant) justification which supports your final answer. You MUST show your work. Answers with no justification will receive no credit.

1. Microzide is a prescription pill used to help patients control their blood pressure. However, doctors are concerned about the potential negative side effects of the drug including yellowing of the skin or eyes (jaundice), severe allergic reactions, and rashes.

   A study of the safety and effectiveness of Microzide included 1,493 patients who volunteered to participate in a clinical trial of the drug. Of the 751 patients randomly assigned to take Microzide over a 6-week period, 28 experienced serious side effects. Of the 742 patients randomly assigned to take a placebo over the same period, 16 experienced serious side effects. Doctors were interested in whether the data provide convincing statistical evidence that taking Microzide would increase the risk of serious side effects among all patients similar to the volunteers.

   A. Does this study represent an experiment or observational study? Explain your answer. (2pt)

This study represents an experiment since the researchers control the environment of the experiment. (administration the use of Microzide or some other similar answers). Stating that the participants were randomly assigned to treatments would also be a correct answer.

2 pts	1 pt	0 pts
Identifies the study as an experiment with a correct justification	Identifies the study as an experiment with an incorrect justification	Identifies the study as an observational study OR Identifies the study as an experiment with no justification.

B. Fill in the table below by identifying the two variables in this study and determining which variable is the explanatory variable and which is the response variable. For each variable also determine if it is qualitative or quantitative, and discrete, continuous or neither.

Variables (2pt)	Which drug the person received, Microzide or the Placebo	Whether or not the person experienced serious side effects
Explanatory or Response (2pt)	Explanatory	Response
Qualitative or Quantitative (2pt)	Qualitative	Qualitative
Discrete, Continuous or Neither (2pt)	Neither	Neither

2 pts	1 pt	0 pts
Both variables are correct	One variable is correct	Neither variable is correct
Variables are correctly labeled as explanatory and response. Note follow work if original variables are incorrect	Variables are reversed OR Only one variable is correctly labeled	Not answered OR Completely incorrect
Both variables are correctly identified as qualitative OR Follow work if variables are incorrect but the student correctly chooses the type of variable.	Only one variable is correctly identified as qualitative OR Follow work and only one variable is correctly identified.	Incorrect identification of the variables as qualitative or quantitative
Both variables are identified as "neither". OR Follow work and student correctly chooses discrete or continuous	Only one variable is identified as "neither" OR Follow work and only one variable is correct	Both variables are incorrectly labeled

2. Consider a dataset which contains 13 data points.

{11, 22, 24, 24, 25, 25, 27, 30, 30, 32, 32, 37, 38}

A. Find the 5-number summary by hand. **Show your work for calculating quartiles.** (5pts)

Min	11
Q1	24 location =13(.25)=3.25 use position 4
Median	27 location = 13(.5)=6.5 use position 7
Q3	32 location =13(.75)=9.75 use position 10
Max	38

1 pt	½ pt	0 pts
Correct minimum	n/a	Incorrect
Correct Q1 and supporting work (students may use an alternative method for calculating Q1, but work must be shown)	Correct value and no supporting work OR Incorrect supporting work	Incorrect
Correct Median and supporting work	Correct value and no supporting work OR Incorrect supporting work	Incorrect OR Mean reported along with the median OR Mean reported instead of the median
Correct Q3 and supporting work (students may use an alternative method for calculating Q3, but work must be shown)	Correct value and no supporting work OR Incorrect supporting work	Incorrect
Correct Max	n/a	Incorrect

B. Calculate the IQR (Inner Quartile Range). (1 pt)

IQR=32-24=8

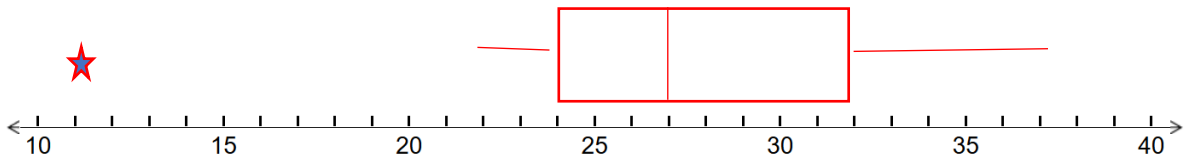
1 pt	½ pt	0 pts
Correct Value Follow work if student has the wrong values in part A.	Minor arithmetic error	Incorrect

C. Determine which data values, if any, are outliers (you must show your calculations). (4 pts)

Q1- 1.5(IQR)= 24 - 1.5(8)= 12 11 is a low outlier
Q3 + 1.5(IQR)= 32 + 1.5(8)= 44 No high outliers

2 pt	1 pt	0 pts
Correct value for lower fence Follow work if parts A and B are incorrect	Correct value for lower fence but no work shown OR Correct work but does not indicate the outlier OR Wrong calculation but a correct identification of outliers	Incorrect Student makes claims about outliers with no justification
Correct value for upper fence and student identifies the high outlier Follow work if parts A and B are incorrect	Correct value for upper fence but no work shown OR Wrong calculation but a correct identification of outliers	Incorrect Student makes claims about outliers with no justification

D. Construct a box plot using the number line below. (6 pts)



1 pt	0 pts
Correct left whisker ending at last point before fence Follow work	Incorrect
Correct Q1 with left edge of box Follow work	Incorrect
Correct Median inside box Follow work	Incorrect
Correct Q3 with right edge of box Follow work	Incorrect
Correct right whisker ending at last point before fence Follow work	Incorrect
Outlier Follow work	Incorrect

3. Below are the results of a sample of size 5 from the Frog Jumping contest held at the Calaveras County fair.

Frog Name	Jump Length
Skeeter Eater	231 in.
Warped	230 in.
Greg	229 in.
R.G.	227 in.
The Well Ain't Dry	221 in.

A. Calculate the mean distance jumped. Round your answer to an integer. Be sure use the correct symbol, show your work, and include units in your answer. (4 pts)

$$\bar{x} = \frac{231 + 230 + \cdots + 221}{5} = 228 \text{ in}$$

1 pt	0 pts
Uses $\bar{x}$	Missing or uses μ
Correctly fills in formula	Incorrect
Correct answer	Incorrect
Uses units	No units or incorrect units

B. Calculate the variance of the distances. Use the correct symbol in your answer and show your work by substituting values into the correct formula. Round your answer to an integer and be sure to include units. (4 pts) You may use either of the tables below to assist in calculating the variance. <u>You are not required to use the tables.</u>

$$s^2 = \frac{(231 - 228)^2 + + \cdots + (221 - 228)^2}{5 - 1} = 16 \text{ in}^2$$

OR

$$s^2 = \frac{(231^2 + \cdots + 221^2) - \frac{(231 + \cdots + 221)^2}{5}}{5 - 1} = 16 \text{ in}^2$$

1 pt	½ pt	0 pts
Uses s^2	Uses s	Missing or uses σ^2 or σ
Correctly fills in formula	Correct formula for standard deviation	Incorrect
Correct answer	Correct answer for standard deviation	Incorrect
Uses units	Uses units for standard deviation	No units or incorrect units

4. Suppose a group of high technology stocks has an average earnings per share of $8.52 with a standard deviation of $1.53.

   A. **Fill in the blanks.** If the data regarding the earnings per share has a bell-shaped distribution, out of the two methods discussed in Chapter 4, (the) ____empirical_____ ____rule_____ gives the best approximation of the percentage of the data that lies in the interval $3.93 and $13.11. (2 pts)

2 pt	0 pts
Uses empirical rule	Does not use empirical rule

   B. If the data does **not** have a bell-shaped distribution, state the range in which at least 75% of the data will reside. Show your calculations. (3 pts)

$1 - \frac{1}{k^2} = 0.75$

$1 - 0.75 = \frac{1}{k^2}$

$.25 = \frac{1}{k^2}$

$.25k^2 = 1$

$k^2 = 4$

$k = 2$

Lower limit: 8.52-2(1.53) = 5.46

Upper limit: 8.52+2(1.53) = 11.58

At least 75% of the data will lie in the interval $5.46 to $11.58

1 pt	½ pt	0 pts
Finds correct value for k	n/a	Incorrect value of k
Correct lower limit Follow work	Correct process with minor arithmetic error Follow work	Incorrect
Correct upper limit Follow work	Correct process with minor arithmetic error Follow work	Incorrect