1. **Don’t even open this until you are told to do so.**

2. All graphs are on the last page which you may remove.

3. There are 20 multiple-choice questions on this exam, each worth 5 points. There is partial credit. Please mark your answers **clearly**. Multiple marks will be counted wrong.

4. You will have 60 minutes to finish this exam.

5. If you have questions, please write out what you are thinking on the back of the page so that we can discuss it after I return it to you.

6. If you are caught cheating or helping someone to cheat on this exam, you both will receive a grade of **zero** on the exam. You must work alone.

7. This exam is worth the same as a regular exam (this may differ from section to section).

8. When you are finished please make sure you have marked your CORRECT section (Tuesday 12:45 is 508, 2:20 is 509, and 3:55 is 510) and FORM and 20 answers, then turn in JUST your scantron.

9. Good luck!
1. The size of a sample, $n$, affects
   A. the spread of the sampling distribution of the statistic
   B. the shape of the sampling distribution of the statistic
   C. the center of the sampling distribution of the statistic
   D. All of the above are true.
   E. Only two of the above are true.

2. Which of the following is/are true for the histogram?
   A. The 5-Number-Summary is 0,1,2,3,4.
   B. The median and the standard deviation should be used to describe this distribution.
   C. The mean is greater than 1.
   D. Two of the above are true.
   E. None of the above are true.

3. Which of the following statements is/are true?
   A. The further an $(x,y)$ data point is from the mean, the less it affects the sample statistics.
   B. Influential points should be deleted since they change the linear relationship between $x$ and $y$.
   C. The intercept is affected by both shift and scale changes in $x$ and $y$, but the slope is only affected by scale changes.
   D. Two of the above are true.
   E. None of the above are true statements.

4. If the normal quantile plot represents data with a mean, $\mu = 30$, and standard deviation, $\sigma = 5$, which of the following is/are true?
   A. $Q_1 = 30 - 0.675 \times 5 = 26.625$ and $Q_3 = 30 + 0.675 \times 5 = 33.375$, so the IQR = 6.75 covers 50% of the observations.
   B. The range of the of the middle 95% of the observations is (20,40).
   C. The approximate minimum and maximum are 15 and 45.
   D. All of the above are true.
   E. None of the above are true.

5. The first thing you must ALWAYS do when given data to analyze is
   A. calculate the mean and the standard deviation
   B. plot the data
   C. determine what type of data you’ve been given
   D. find any outliers since they will affect the statistics you’ll need to calculate
   E. decide if there could be lurking variables

6. Which of the following statements is/are true above the scatterplot?
   A. The correlation is weak.
   B. The point (6,2) is a possible influential point.
   C. The point at (0.8,0) is an outlier.
   D. All of the above are true.
   E. Only two of the above are true.

7. Suppose we weigh 50 people. 49 of them are between 180 and 280 lbs. One weighs only 150 lbs. Which of the following is true for this sample?
   A. The sample mean, $\bar{x}$, is NOT the best measure of center for this data.
   B. The sample mean, $\bar{x}$, will be larger than the sample median.
   C. The data cannot be normally distributed since there is an outlier.
   D. All of the above are true for this data.
   E. Only two of the above are true for this data (excluding D.).

8. Numbering them from left to right, 1 through 6, which of the boxplots are normally distributed?
   A. 5 only
   B. 4 and 5
   C. 4, 5 and 6
   D. 4 only
   E. 6 only

9. Assume that each boxplot represents 50 data points. Using the definition of standard deviation, which distribution most likely has the largest standard deviation?
   A. 1 since it would have the largest outlier (furthest from the mean)
   B. 3 since it has the most outliers
   C. 4 since it has the largest range
   D. 6 since it has the largest IQR
   E. It’s impossible to say since the standard deviations aren’t given.

10. One more about the boxplots. Which boxplot(s) most likely has a mean less than 50?
    A. You can’t determine the mean from a boxplot.
    B. 1
    C. 1 and 2
    D. 1, 2 and 3
    E. 1, 2, 3 and 5

11. Suppose you want to take a survey to determine how Texans will vote in the upcoming election. Which of the following should you include?
    A. an equal number of Republicans and Democrats so you don’t bias your result
    B. blinding so that the voters don’t know what your survey is about because that could influence their answers to your survey
    C. an incentive to answer your questions to reduce nonresponse bias
    D. a stratified sample to ensure that all regions of the state are represented
    E. randomize the gender of the poll taker since voters may respond differently to a male or a female
12. Which of the following is FALSE?

A. People involved in an experiment might act differently just because they are participating in the experiment.

B. Telling a patient that they are taking a placebo (a sugar pill, e.g.), might influence the experiment's results.

C. Observational studies are useful because they help us determine what causes certain diseases.

D. Confounding variables affect the response variable, and they're effects cannot be separated from the explanatory variables effects.

E. Randomized experiments are preferred since they are better at determining cause and effect.

13. Suppose you have a bell-shaped distribution with mean, \( \mu = 40 \), and standard deviation, \( \sigma = 4 \). What approximate percent of the observations fall between 44 and 52?

A. approximately 95%

B. about 16%

C. a little less than 16%

D. about 81.5%

E. about 13.5%

14. Which of the following is/are true for bivariate, \((x,y)\), data?

A. Multiplying the \( x \)'s by 5 will increase the slope, \( b_1 \), by 5 but will not change the intercept, \( b_0 \).

B. Adding 25 to both the \( x \)'s and the \( y \)'s will not change the intercept unless the slope is 1.

C. Removing outliers will always increase the correlation.

D. All of the above are true.

E. None of the above are true.

15. Suppose you need to convert a distribution, \( X \), with mean 50 and standard deviation 3, to a new distribution, \( Y = 14 - 8X \). What are the new mean and standard deviation?

A. \( Y = 50 \) and \( s_Y = 3 \), shift and scale changes don't affect the new distribution

B. \( Y = -386 \) and \( s_Y = -10 \)

C. \( Y = -386 \) and \( s_Y = 10 \), standard deviations can't be negative

D. \( Y = -386 \) and \( s_Y = 24 \)

E. \( Y = -386 \) and \( s_Y = -24 \)

16. Which of the following is/are true?

A. Categorical data can't be numbers

B. Categorical data can never be normal in shape

C. Categorical data is rarely used since you can't calculate a mean or standard deviation

D. All of the above are true.

E. None of the above are true.

17. The proportion of students represented above who are junior Geoscience majors is

A. 6/78

B. 6/11

C. 6/41

D. 11/41

E. 11*41/78

18. Which of the following is/are true about the previous Two-way table?

A. 4/17 is a conditional probability conditioning on being a senior.

B. 21/41 is a conditional probability conditioning on being a junior.

C. 14/78 is a marginal probability of classification.

D. All of the above are true.

E. Only two of the above are true.

19. Suppose your high school grading system used an 8 point scale with a mean of 6.5 and a standard deviation of 1.2 and your average was 7. What is your equivalent average in a 100 point system with a mean of 82 and a standard deviation of 14?

A. 82.417

B. 87.838

C. 76.162

D. 81.583

E. 83.2

20. Which of the following best describes a 95\(^{th}\) percentile for a given distribution?

A. the value of 95% of the observations

B. a number that is 2 standard deviations above the mean of the distribution

C. a number that would be equivalent to 95 on a 100 point scale

D. a number such that 95% percent of the observations in the distribution have that value or less

E. Two of the above are true.