(1) The following terms measure location. Make sure you know the definitions, how they differ from each other, how to calculate them, and when you’d want to use each one: mean, median.

(2) The following terms measure spread. Make sure you know the definitions, how they differ from each other, how to calculate them, and when you’d want to use each one: variance, IQR, standard deviation.

(3) The following terms measure shape. Make sure you know what they measure: modality, skew, kurtosis.

(4) The following terms are relative measures. Make sure you know the definitions, how they differ from each other, how to calculate them, and when you’d want to use each one: correlation, covariance, coefficient of variation, z-score.

(5) Describe the qualities of, and give a simple example of, each of the following types of distributions: uniform discrete, uniform continuous, Bernoulli, binomial, Poisson, normal, exponential.

(6) What’s the difference between a probability function and a density function?

(7) What’s the difference between a population and a sample?

(8) Suppose a tech stock price has an average of $50/share with a standard deviation $10, and an automotive stock has an average of $25/share with a standard deviation of $7.50. If you want a less risky investment, which one should you choose?

(9) Suppose you know that 95% of all US college students love statistics, only 50% of non-college students love statistics, and 10% of the US population is enrolled in college. Suppose you meet someone who loves statistics. What’s the probability that he or she is a college student?

(10) Translate the following into statistical terminology: The median age of an invoice is 40 days, and about three-quarters of the invoices remain unpaid for 30 or more days. About 25% of the invoices are paid within 40 and 55 days.
(11) Consider the histogram depicted below, which describes the number of minutes after 6:30 AM when the first flash of lightning occurred each day.

![Histogram of first daily lightning strike times]

(a) Comment on the shape of the histogram.
(b) Make a relative frequency table for the data.
(c) Make a cumulative percent frequency table for the data.
(d) Estimate which bin(s) the mean and median will fall in. Do you have reason to pick one over the other as a measure of location? Justify your response.
(e) If the mean of the data is $\bar{x} = 300$ and the standard deviation is $s = 60$, compute the $z$-score of the largest bin value (i.e., the largest value represented by this data). Is this an outlier? Why or why not?
(f) If asked to give a report based on this histogram, what could you say about times of lightning strikes?

(12) Consider the following data about a stress in test-takers. In 19 out of 20 cases where the test-taker was stressed, he or she passed the exam. In 8 out of 10 cases where the test-taker was not stressed, he or she failed the exam.

(a) Make a crosstabulation for this data.
(b) Given the test-taker is passes, what is the relative frequency of being stressed?
(c) What is the relative frequency of not being stressed and failing?
(d) Based on this data, would you say that passing and being stressed are independent? Justify your answer numerically.

(13) You read in the newspaper that a man has been convicted of murder. Apparently, the person was wearing a clown hat on the night of the murder. In the general population, very few people wear clown hats, but it is known that 80% of murderers wear clown hats. The jury claimed that because the defendant was wearing a clown hat, it was too unlikely that he was innocent. What is your reaction to the article?
(14) You are given the choice of playing two gambling games involving rolling a fair die. Each costs $5 to play. Which game should you play if you want to maximize your expected net winnings? Support your answer.

**Game 1:** If you roll 1, 2, or 3, you win nothing. If you roll 4, you win $3. If you roll 5, you win $5. If you roll 6, you win $10.

**Game 2:** You win the dollar amount shown on the die (so rolling 1 pays $1, 2 pays $2, etc.).

(15) Consider the scatter plot below, relating the number of plants planted per acre with the yield in bushes per acre. A trendline has been added to the plot.

(a) What kind of curve best describes the data?

(b) Using the graph, guess the values for the mean yield and the mean number of plants per acre.

(c) Would you say that the variance of plants/acre or the variance of yield is larger? Why?

(d) What can you say about the correlation of these two variables?

(e) If you were planting, how many plants/acre would you plant?

(16) I report to my department chair that 100% of my students love statistics. I know this because I asked each of you as you left class one evening and marked down your response. Should my department chair give me a raise based on my results? Why or why not?
(17) Match each type of distribution with the most appropriate scenario below. Write your responses in the spaces provided.

(a) Poisson ____
(b) normal ____
(c) binomial ____
(d) uniform discrete ____
(e) uniform continuous ____
(f) Bernoulli ____
(g) exponential ____

(i) You are playing a video game where aliens pop up. Each time, you have a 70% chance of killing it.
(ii) One of 10 office telephones rings when a customer calls. For any call, each phone has probability 0.10 of being the phone to ring.
(iii) The density function looks bell-shaped.
(iv) You are playing a video game where you have to kill as many as possible of the 4000 aliens that will pop up. You are interested in the number of aliens you kill in each game.
(v) You are playing a video game where, on average, 12 aliens pop up per minute. You trim your garde shrubbery every week. Each time, the shrubs are trimmed to some height between 5 and 5.5 ft. and any height is equally likely to occur.
(vi) You are playing a video game where, on average, an alien pops up every 5 seconds.

Extra Credit: During WWII, some planes came back from strikes riddled with bullet holes while others didn’t come back. Each time a plane went out and returned, it had more bullet holes. If you were an airplane pilot, where would you want the most reinforcement on your plane? Why?