Stats chapter 1,2, 3 review

What is the term for a characteristic or attribute that can assume different values? VariableStatistics is the science of conducting studies tocollect, organize, summarize, analyze, and draw conclusions from dataVariables with values that are determined by chance are calledrandom variablesEach value in a data set may be referred to as either a data value or a(n)datumlf a weather center monitors and calculates the average number of tornadoes that pass through Topeka, Kansas each year, what type of variable would they be investigating? random variableWhich branch of statistics would employ probability to predict how many miles one would be able to drive a 2000 Toyota Celica during its lifetime? inferential statisticsWhich branch of statistics would buy a hundred Toyotas, drive them into the ground, record the final mileage, and then write a report for Car and Driver? descriptive statisticsWhich of the following correctly describes the relationship between a sample and a population? A sample is a group of subjects selected from a population to be studiedBased on the following graph, what conclusion could you make comparing how well students did on their statistics exam as a function of how many hours they spent preparing for the exam? There is a possible relationship between grades and time spent preparing for the examif you were told that four students from a class of twenty were questioned for a grade versus test preparation poll, this would be an example ofsamplingWhat level of measurement classifies date into mutually exclusive (non overlapping), exhausting categories in which no order or ranking can be imposed on the data? nominalAn independent variable can also be called a(n)explanatory variableWhat level of measurement possesses all the characteristics of interval measurement, and there exists a true zero? ratiolf a researcher manipulates one of the variables and tries to determine
how the manipulation influences other variables, the researcher is conducting a(n)experimental studylf you classified the fruit in a basket as apple, orange, or banana, this would be an example of which level of measurement? nominalWhich of the following best defines the relationship between confounding, dependent, and independent variables? The confounding variable influences the dependent variable, but cannot be separated from the independent variableln a true experimental study, the subjects should be assigned to groups randomly. If this is not possible and a researcher uses intact groups, they are performing aquasi-experimental studyWhat would be the boundaries on the average age for high school graduates if they were reported to be 18 years old? 17. 5-18. 5The amount of time needed to run the Boston marathon is an example of which type of variablecontinuousWhat type of sampling is being employed if the country is divided into economic classes and a sample is chosen from each class to be surveyedstratified samplingThe four basic methods used to obtain samples are: random, irregular, cluster, and stratified. Falselnferential statistics is based on probability theoryTrueWhen running an experimental study, the group that is manipulated is called the treatment group. TrueData can be classified as qualitative, continuous, or non-sequentialFalseThe number of birds in a tree is an example of a continuous variableFalseln the following chart, the height is the independent variable and the age of the tree is the dependent variableFalseA dependent variable can also be referred to as an outcome variableTrueRating a restaurant by a number of stars is an example of an ordinal level of measurementTrueAlthough it is much easier to perform long statistical computations on a calculator or computer, a student still needs to learn how these computations are done in order to understand the
data. TrueBased on Mrs. Smith's electric bill for last year she expects that she will be paying $\$ 75 /$ month this year. This is an example of descriptive statisticsFalseA person's hair color would be an example of a quantitative variableFalself every 14th customer leaving a movie was surveyed, this would be an example of systematic samplingTrueThe variable of height is an example of a quantitative variableTrueln a research study, it is always preferable for the researcher to choose his participants as carefully as possible rather than randomly accept samplesFalseThe $\qquad$ level of measurement classifies date into categories that can be ranked; however, precise differences between the ranks do not exist. ordinalThe number of people form the state of Alaska who voted for a Republican in the last election is an example of the $\qquad$ level of measurement. ratioA $\qquad$ consists of all subjects that are being studied. population $\qquad$ is a decision making process for evaluating claims about a population based on information obtained from samples. hypothesis testingA $\qquad$ variable assumes values that can be counted. discreteOne advantage of $a(n)$ $\qquad$ study is that it occurs in a natural settingobservational $\qquad$ sampling is used when the population is large or when it involves subjects residing in a large geographic area. clusterThe $\qquad$ variable influences the $\qquad$ variableindependent, dependentHow much a telephone survey performed between the hours of 8AM and 5PM be biased? Because they are only interviewing people available during standard working hours. What level of measurement would be applied when doing a survey on the average American's shoe size? intervalHow are statistics important in our everyday lives, and why do we need to understand them? They are used to analyze results of surveys, and especially in sports and insurance. It is important to understand them
because it helps us predict the future and understand graphs and tables. Explain the difference between qualitative, quantitative, discrete and continuous variables. Qualitative variables: can be placed in categories, but not rankedQuantitative variables: can be rankedDiscrete: assigned value, generally an integer. can be countedContinuous: assume value between 2 specific values. An ad for an exercise product states:" Using this product will burn 74\% more calories." This is an example ofdetached statisticsAn advertiser states that its brand of energy pills gets into the user's blood stream faster than a competitor; s and shows the following graphs to prove its claim. Why is the comparison misleading? There are no labels or scales, so graph 1 looks faster. Which of the following should not be done when constructing a frequency distribution? use a class width with an even numberDetermine the range for this data: $4,7,3,16,5,22$, and 8 . 19 If a frequency distribution had class boundaries of $132.5-147.5$, what would be the class width? 15A grouped frequency distribution is used when the range of the data values is relatively small. FalseThe lower class limit represents the smallest data value that can be included in the class. TrueThe $\qquad$ is the number of values in a specific class of distribution. frequencyWhen data are collected in original form, they are called $\qquad$ . raw dataGreg wants to construct a frequency distribution for the political affiliation of the employees at Owen's Hardware Store. What type of distribution should he use? categoricalWhat is the lower class limit in the class 13-17? 13What is the midpoint of the classes $13.5-17.3$ ? 15.4 Using the class $23-35$, what is the upper class boundary? 35. 5The $\qquad$ is obtained by first adding the lower and upper limits and then dividing by 2. midpointlf the limits for a class were 20-38, the boundaries would be 19. 5-38. 5. TrueWhen the range is
large and classes that are several units in width are needed, a $\qquad$ frequency distribution is used. groupedFor the class 16. 3-23. 8, the width is 7. FalseWhat are the boundaries of the class 1. 87-3. 43? 1. 865-3. 435Find the class with the least number of data values. $(70,90,60,40) 90$ Find the class with the greatest number of data values (70, 90, 60, 40)60An ogive graph is also called a cumulative frequency graph. TrueThe cumulative frequency is the sum of the frequencies accumulated to the upper boundary of a class in the distribution. TrueThe three most commonly used graphs in research are the histogram, the $\qquad$ , and the cumulative frequency graph (ogive). Frequency PolygonThe graphs that have their distributions as proportions instead of raw data as frequencies are calledrelative frequency graphs. Which type of graph represents the data by using vertical bars of various heights to indicate frequencies? histogramThe frequency polygon is a graph that displays the data by using lines that connect points plotted for the frequencies at the midpoints of the classes. TrueA histogram uses the midpoints for the $x$ values and the frequencies as the $y$ values. FalseA histogram is a graph that represents the cumulative frequencies for the classes in a frequency distribution. FalseGiven the following frequency distribution, how many pieces of data were less than 28.5? Class Boundaries - Frequencies13. 5-18. 5 - 418. 5-23. 5 - 923. 5-28. 5 - 1228. 5-33. 5 1533. 5-38. $5-1725$ If data is clustered at one end or the other, it indicates that there is a $\qquad$ . skewed distributionA weatherman records the amount of rain that has fallen in Portland, Oregon during each day. What type of graph should he use? time series graphGraphs give a visual representation that enables readers to analyze and interpret data more easily than they could simply by looking at numbers. TrueA time series graph
represents data that occur over a specific period. TruePareto charts have units that are used for the frequency that areequal in size. An automobile dealer wants to construct a pie graph to represent types of cars sold in July. He sold 72 cars; 16 of which were convertibles. The convertibles will represent how many degrees in the circle? $80^{\circ}$ In a pie chart, if pepperoni pizza were $24 / 72$ of the distribution, how many degrees would be needed to represent pepperoni? $120^{\circ}$ Which graph should be used to represent the frequencies that certain types of classes are taken at Highlands Middle School? Pareto chartExaggerating a one-dimensional increase by showing it in two dimensions is an example of a(n)misleading graph. A Pareto chart arranges data from largest to smallest according to frequencies. TrueWhen two sets of data are compared on the same graph using two lines, it is called a compound time series graph. TrueA pie graph would best represent the number of inches of rain that has fallen in Ohio each day for the past 2 months. FalseThe percentage of white, wheat, and rye bread sold at a supermarket each week is best shown using a $\qquad$ graph. pieA
$\qquad$ would most appropriately represent the number of students that were enrolled in Statistics for the past ten yearstime series graphA pie graph was created showing the number of children per family. If 234 families were in the survey and the section depicting families with three children represented $120^{\circ}$, the number of families with three children was 78. TrueKaren is constructing a pie graph to represent the number of hours her classmates do homework each day. She found that 8/24 did homework for three hours each day. In her pie graph, this would represent how many degrees? $120^{\circ}$ What is the term for a characteristic or measure obtained by using all the data values for a specific population? parameterWhich of the
following is the correct mean for the given data? $7,8,13,9,10,119.7$ Which of the following is the mode for the given data? $5,4,3,4,5,6,5,5,3,45$ Find the mode for the number of police officers in selected city districts. 24, 26, $24,30,23,28,19,31,24,26,1924$ The following data set could also be referred to as a data array. 3, 4, 2, 7FalseA weighted mean is used when the values of the data set are not all equally represented. TrueFind the median for the following data. $6,7,4,5,3,7,45$ In a unimodal, symmetrical distribution as shown in the figure below. The mean, the median, and the mode are the same. The median can be a more appropriate measure of central tendency if the distribution of the data is extremely skewed. Truelf a distribution is negatively skewed as shown in the figure below, the mean will fall to the right of the median and the mode will be on the left of the median. FalseFor the sample $1,8,7,2,9,15$, and 18 , the mean is 7 . 6 . FalseA is the midpoint in a data array. MedianThe $\qquad$ is the mode for grouped data. Modal ClassFind the mean, mode, median, and midrange value for the following data set. $12,15,18,18,15,22,15,30$, 12 Mean $=17.4$, Median= 15, Mode $=15$, Midrange $=21$ Given that the mean of a set of data is 25 and the standard deviation is 3 , what would be the coefficient of variation? $12 \%$ Given that the variance for a data set is 1.20 , what would be the standard deviation? 1. 10The variance is the square root of the standard deviation of a set of data. FalseThe range of a data set is the distance between the highest value and the lowest value. TrueChebyshev's theorem can be used to find the minimum percentage of data values that will fall between any two given values. TrueThe coefficient of variation is the mean divided by the standard deviation expressed as a percentage.

FalseThe $\qquad$ and $\qquad$ are used to determine the
$\qquad$ applies to any distribution regardless of its shape. Chebyshev's TheoremThe grades for a trigonometry exam follow. Find the range. 85, 76, 93, 82, 84, 90, 7518The unbiased estimator is included in the formula for calculating the variance of a sample because without it, the computed variance usually underestimates the population variance. TrueThe $\qquad$ is the average of the squares of the distance each value is from the mean. Variance $\qquad$ divide the distribution into four groups, and
$\qquad$ divide the distribution into ten groups. Quartiles, Deciles $\qquad$ are either extremely high or extremely low data values compared with the rest of the data. OutliersThe interquartile range or IQR is found by subtracting the mean from the maximum value of a data set.

FalseThe percentile corresponding to a given value $X$ is computed by adding the 0.5 to number of values below $X$ and dividing this value by the total number of values within the data set. TrueGiven the following data set, find the value that corresponds to the 75 th percentile. $10,44,15,23,14,18,72$, 5644A stem and leaf plot is a data plot that uses part of a data value as the stem and part of the data value as the leaf to form groups or classes. Trueln exploratory data analysis, $\qquad$ are used instead of quartiles to construct boxplots. HingesA five-number summary of a data set consists of the minimum, , the mean, , and the maximum. FalseChoose the correct statement describing the following stem and leaf plot for grades on a linear algebra exam. Of the 29 students who took the exam, nine scored between 80 and 89.

