1. **2 Components of Shape Distribution**	1. Skewness 2. Kurtosis
2. **2 sources of data**	primary and secondary
3. **2 Types of random variables**	discrete random and continuous random
4. **4 categories from where data is found**	1. data distributed by an organization/individual 2. designed experiment 3. survey 4. observational study
5. **6 categories for each variable, so the contingency table is called a _____ table**	6x6
6. **6 Hypothesis Testing Steps**	1. Null & alternative hypothesis 2. α=_____ and n=_____ 3. Relationship? (nominal/ordinal data? normal?) 4. Decision ruled based on α=.05 & ____ degrees of freedom 5. $x^2= \Sigma(fo-fe)^2/fe$ 6. Make a decision.
7. **About how many of the techniques require normal distribution?**	80-90%
8. **Advantage of a pivot table**	in excel, can easily convert to an overall percentages table can easily add variables to an existing table can easily change the statistic displayed
9. **advantage of the standard deviation**	uses the same units as the original data
10. **All Continuous distributions are normal. True or False?**	False
11. **All of the items or individuals about which you want to draw a conclusion about are known as**	a population
12. **another name of an "ogive" is a:**	cumulative percentage polygon
13. **Another word for "ogive"**	cumulative percentage polygon
14. **Anytime you have mean and standard deviation, it means you have what kind of data?**	interval/ratio data
15. **Applied Research**	activity that has relevancy for current practice or policy Helps you make a decision --->Ultimate reason for business research
16. **Basic Research**	increase knowledge about a topic by advancing the theory
17. **The boundaries of class groupings for a frequency distribution should be overlapping. True or false?**	False
18. **The Boxplot**	a graphical display of the data based on the five number summary
19. **Box Plot Skewdness**	...
20. **the branch of statistics that collects, summarizes and presents data is known as**	descriptive statistics
21. **the branch of statistics that uses sample data to draw conclusions about an entire population is known as**	inferential statistics
22. **business managers use statistics to do all of the following:**	to improve business processes, to present and describe business data and information properly, make reliable forecasts about a business activity
23. **a business student applying to an MBA program must compete a standardized exam known as the GMAT, which is measured on an interval scale. True or fale?**	True...difference between measurements important but no true zero
24. **Categorical Data**	also called qualitative placed into defined categories

25. **Categorical data is organized with?**	tables 1 variable--summary table 2+ variables--contingency table
26. **Causation Implies _____ but _____ alone does not imply _____**	causation implies CORRELATION, but CORRELATION alone does not imply CAUSATION
27. **The central limit theorem says that as the sample size gets large enough...**	the sampling distribution of the sample mean becomes almost normal regardless of shape of population
28. **Central Limit Theorem states...**	that the sample mean from the population WILL BE APPROXIMATELY NORMAL as long as the sample size is large enough
29. **Changing μ does what to a normal distribution?**	shifts distribution left or right
30. **Changing σ does what to a normal distribution?**	increases or decreases the spread (narrow or wider)
31. **A characteristic of an individual or item is known as**	a variable
32. **The characteristic shape of the cumulative percentage polygon, or ogive is a bell curve. True or false.**	False...it is an "S" shape
33. **Characteristics of a Normal Distribution**	1. Bell Shaped 2. Symmetrical 3. Mean, Median and mode are equal 4. Location is determined by the mean, μ 5. Spread is determined by the standard deviation, σ 6. Random variable has an infinite theoretical range: $+\infty$ to $-\infty$
34. **Chebyshev Rule**	at least $(1-1/k^2)$ *100% of the values will fall within k standard deviations of the mean (for k>1) REGARDLESS OF HOW THE DATA ARE DISTRIBUTED
35. **Chi-Square Statistic**	...
36. **coefficient of correlation indicates what?**	the linear relationship, or association, between two numerical variables when it is closer to −1 or +1, the relationship is stronger
37. **Coefficient of Correlation Symbol**	population coefficient: p (rho) sample coefficient: r
38. **Coefficient of Variation**	Measures Relative Variation Always shown in % Shows VARIATION RELATIVE TO MEAN can be used to compare the variability of two or more sets of data measured in different units CV = (standard deviation/mean)*100 If below 50%, more likely to be normal distribution
39. **a company is considering 10 college graduates for a management trainee position and must rank them from most suitable (rank =1) to least suitable (rank =10). These rankings represent what type of measurement scale?**	ordinal scale
40. **the complete listing of names of all persons that currently own GM automobiles represents...?**	population
41. **Compute the Z Score**	Subtract the mean and divide by the standard deviation z= X-sample mean / sample standard deviation
42. **Computing the Average Proportion**	p bar = x_1 + x_2/n_1+n_2 = x/n

43. **contingency tables are best used when**	comparing multiple population proportions used to classify sample observations according to two or more characteristics also called a "cross-classification table"
44. **Continuous Variables**	come from a measurement (ex: annual salary, weight) decimals, fractions
45. **The correlation of hypothesis to subproblem is ?**	1 to 1
46. **Covariance Formula**	...
47. **Covariance(X,Y) < 0 means...**	X & Y tend to move in OPPOSITE directions
48. **Covariance(X,Y) = 0 means...**	X & Y are independent
49. **Covariance(X,Y) > 0 means...**	X & Y tend to move in the SAME direction
50. **Data from a categorical variable are measured on what kind of scale(s)?**	nominal or ordinal ordinal if the categories have an implied ranking
51. **D.C.O.V.A.**	Define the variables Collect data Organize data by developing tables Visualize data by developing charts Analyze data by examining charts and tables to reach conclusions
52. **The Decision Rule**	If x^2stat > $x^2\alpha$, REJECT H0, otherwise do not reject H0 1 degree of freedom
53. **Decision rule for x^2 test of independence**	if x^2stat > $x^2\alpha$, REJECT H0, otherwise, do not reject H0 Where $x^2\alpha$ is from the chi-squared distribution with (r-1)(c-1) degrees of freedom
54. **Definition of a Random Variable**	represents a possible numerical value from an uncertain event
55. **Definition of Central Tendency**	the extent to which all the data values group around a typical or central value 'middle of the data'
56. **Definition of Shape**	The pattern of the distribution of values from lowest value to highest value
57. **Definition of Variation**	the amount of dispersion or scattering of values
58. **Descriptive Statistics consists of?**	collecting, summarizing, presenting and analyzing data collect data-survey present data-tables & graphs characterize data-the sample mean
59. **Descriptive statistics uses ____ to _____?**	numerical measures to describe features of a set of data
60. **Difference between sample and population variance formulas**	sample variance and standard deviation uses (n-1) in denominator while population variance uses N
61. **Different sample sizes from the same population will yield _____?**	different sample means
62. **Discrete Variables**	come from a counting process (ex: number of classes you are taking) countable number of items whole #'s

63. **The Empirical Rule**	approximates the variation of data in a bell-shaped curve distribution approximately 68% of the data in a bell shaped distribution is within 1 standard deviation of the mean or μ±1σ approximately 95% of the data in a bell shaped distribution is within 2 standard deviations of the mean or μ±2σ approximately 99.7% of the data in a bell shaped distribution is within 3 standard deviations of the mean or μ±3σ If normally distributed*
64. **Empirical Rules for Normal Distribution**	μ±1σ covers about 68.26% of X's μ±2σ covers about 95% of X's μ±3σ covers about 99.7% of X's
65. **Evaluating Normality using charts or graphs**	small/moderate size data sets: -construct a stem-and-leaf plot display or boxplot to check for symmetry large date sets: -does the histogram or polygon appear bell-shaped?
66. **Evaluating Normality using Descriptive Summary Measures**	1. Mean, median and mode are all ≈ 2. IQR ≈ 1.33σ 3. Range ≈6σ
67. **Evaluating the Normal Probability Plot**	does the normal probability plot appear linear with a positive slope? (straight line?)
68. **Examples of Ordinal Scale**	student class designation product satisfaction faculty rank standard & poor's bond ratings student grades
69. **Expected Cell Frequencies Formula**	fe= row total * column total / n
70. **Expected Value of a Discrete Random Variable**	μ=E(X)=Σ XiP(Xi) Sum X * Probability X
71. **Features of the Coefficient of Correlation**	Either p or r have: 1. unit free 2. ranges between -1 and 1 3.The closer to -1, the stronger the negative linear relationship 4. the closer to 1, the stronger the positive linear relationship 5. the closer to 0, the weaker the linear relationship Ho: p = 0 Ha: p≠ 0
72. **First Quartile Position**	Q1 = (n+1)/4 value for which 25% of the observations are smaller and 75% are larger
73. **The Five Number Summary**	1. xSmallest 2. First Quartile (Q1) 3. Median (Q2) 4. Third Quartile (Q3) 5. xLargest
74. **For fairly symmetric distributions, n should be > _?**	n > 15
75. **For most distributions, n > ____ will give a sampling distribution that is nearly normal**	n > 30
76. **Formula to standardize p to a Z value**	Z = p-π/σp = p-π
77. **For normal population distributions, the sampling distribution of the ___ is always _____**	the sampling distribution of the MEAN is always NORMALLY DISTRIBUTED
78. **Frequency Distributions do what?**	arrange the data in numerically ordered classes
79. **graph with a larger standard deviation appears**	flatter
80. **graph with a smaller standard deviation appears...**	steeper
81. **Height, weight, age and salary are an example of what kind of data**	Ratio Scale

82. **Highest level of measurement**	data measured on an interval or ratio scale because you can determine not only which observed value is larger but also by how much
83. **Histogram**	-Visualize numerical data -data in a frequency distribution -NO GAPS BETWEEN ADJACENT BARS -horizontal axis-class boundaries -vertical axis-frequency, relative frequency or percentage -% histogram-vertical axis would show % of observations per class
84. **How do we obtain different normal distributions?**	varying the parameters μ (mean) and σ (standard deviation)
85. **How do you determine the width of a class interval**	Range of data / # of class groupings desired
86. **How to compute proportion or relative frequency?**	frequency in each class / total number of values
87. **How to Construct a normal probability plot**	1. arrange data into an ordered array 2. find corresponding standardized normal quantile values (z) 3. plot the pairs of points with observed data values (x) on the vertical axis and the standardized normal quantile values (Z) on the horizontal axis 4. evaluate the plot for evidence for linearity
88. **How to determine degrees of freedom**	(row-1)*(column-1) ex: (2-1)(2-1)=1 degree of freedom
89. **How to translate to the standardized normal distribution (formula)**	Z = X − μ / σ
90. **If the population is normal with mean μ and standard deviation σ, the sampling distribution of x bar is also _____**	normally distributed. μ x bar = μ and σ x bar = $\sigma/\sqrt{n}$
91. **If the population is not normal and we want to do sample mean distribution we can apply the ____?**	central limit theorem
92. **If you have a non-normal distribution and want to know the measure of central tendency you should use the ___ because _____?**	use the median because not affected by extreme values (outliers) like the mean is
93. **In a contingency table, the number of rows and columns:**	...
94. **In a histogram, there are no gaps between adjacent bars as there are in a bar chart of categorical data. True or false**	True
95. **In analyzing categorical data, the following graphical device is NOT appropriate**	stem and leaf display
96. **In a X^2 problem, if no relationship between variables, what would you expect X^2 to be?**	?
97. **Inferential statistics allows you to ____ through ____ and _____?**	say something about the population from which the sample came through estimation and hypothesis testing
98. **Inferential statistics used _____ to _____**	data collected from a small group to DRAW CONCLUSIONS about a larger group
99. **Interquartile Range**	Q3 - Q1 Measures the spread in the middle 50% of the data Also called "midspread" Measure of variability that is NOT influenced by extreme values or outliers (resistant measures)

100.	**the intersection of a column and row in an Excel worksheet form boxes known as**	cells
101.	**Interval Scale**	Difference between measurements is meaningful but measurements DO NOT have a true zero point
102.	**a japanese employee's annual salary in japanese yen represents which scale of measurement?**	ratio scale
103.	**Kurtosis**	measures the relative concentration of values in the center of a distribution as compared with the TAILS
104.	**Left Skewed**	Mean < Median Negatively skewed
105.	**the lowest level of measurement is**	nominal scale
106.	**The main principle behind the Pareto diagram is the ability to track the "vital few" from the "trivial many" True or false.**	True
107.	**The Mean**	most common measure of central tendency CAN ONLY CALCULATE MEAN WITH INTERVAL AND RATIO DATA sum of values/number of values AFFECTED BY EXTREME VALUES (outliers) "the average"
108.	**Measures of Central Tendency**	1. arithmetic mean 2. median 3. mode 4. geometric mean
109.	**Measures of Variation**	1. range 2. variance 3. standard deviation 4. coefficient of variation *Give information on the SPREAD or VARIABILITY or DISPERSION of the data values how close/far data values are from the middle
110.	**Measures of Variation: Summary of Characteristics**	the more the data are spread out, the greater the range, variance and standard deviation the more the data are concentrated, the smaller the range, variance and standard deviation if the values are all the same (no variation), all these measures will be zero None of these measures will ever be negative
111.	**The Median**	"the middle" of an ordered array **not affected by extreme values CAN USE ORDINAL, INTERVAL AND RATIO DATA
112.	**The Mode**	the value that occurs most often in a set of data most frequently used measure of central tendency can be calculated with nominal, ordinal, interval and ratio data may be no mode or several modes numerical or categorical data not affected by extreme outliers
113.	**most common measure of variation**	sample standard deviation

114.	**Normally distributed data should approximate the theoretical normal distribution:**	1. Normal distribution is bell shaped (symmetrical) where mean is equal to the median 2. Empirical rule applies to the normal distribution 3. The IQR is 1.33 standard deviations
115.	**Numerical Data**	also called quantitative 1. discrete (counting process) 2. continuous (decimals)
116.	**A numerical measure that describes a characteristic of a population is known as**	a parameter
117.	**the # of classes in a frequency distribution depend on?**	the values in the data the more data, the more number of classes As data size increases, the impact of changing class boundaries greatly decreases
118.	**One advantage of the stem-and-leaf display over histograms is that it maintains the original values for further analysis. True or False.**	True
119.	**one of the advantages of a bar chart is that it clearly shows the total of all the categories of the bar chart adds to 100%. True or false.**	False...this is an advantage of the pie chart, not the bar chart
120.	**Ordered Array**	sequence data in rank order from small to large shows the range of the data helps identify outliers numerical data
121.	**An ordered array has the ability to separate the "vital few" from the "trivial many". True or false?**	False...a pareto diagram seperates the vital few from the trivial many
122.	**The percentage of responses for each category in a pie chart must sum to 100%. True or false?**	True
123.	**percentage polygon uses ?**	midpoint of each class to represent the data
124.	**the PHStat2 software that accompanies the textbook is known as an addd in...**	that seamlessly creates model Excel solutions students can examine and incorporate into their own Excel solutions
125.	**p is approximately distributed as a normal distribution when _____ is _____**	n is large
126.	**Pivot Tables**	can be used to discover possible patterns and relationships in multidimensional data excel tool can be used to change and/or add variables to a table
127.	**pivot tables used with?**	multidimensional data used to discover possible patterns and relationships
128.	**Portfolio Expected Return & Risk**	risk is standard deviation return is mean
129.	**Primary Data**	data collector is the one using the data for analysis examples: political survey, collected from experiment, observed data

130.	**Principles of Excellent Graphs**	1. not distort the data 2. not contain unnecessary adornments (chart junk) 3. scale on the VERTICAL AXIS should begin at zero 4. all axes should be properly labeled 5. should contain a title 6. simplest possible graph should be used for a given set of data
131.	**Probability As Under the Curve**	total area under the curve is 1.0 half above the mean, half below the mean
132.	**Probability Distribution for Discrete Random Variable**	mutually exclusive listing of all possible numerical outcomes for that variable and a probability of occurrence associated with each outcome MUST EQUAL 1
133.	**Proportions means ____ data**	nominal (maybe ordinal) data
134.	**Quartiles**	Q1 and Q3 are measures of non-central location Q2=median, is a measure of central tendency Q4 always the highest value in a data set
135.	**Quartiles are ____ not _____**	position not value
136.	**Quich Rule**	abs(r) > (2/√n) the smaller the n, the value for r has to be closer to 1 or -1 the larger the n, the value for r has to be farther away from 1 or -1
137.	**Ratio Scale**	difference between measurements is meaningful and measurements DO HAVE a true zero point
138.	**The relative frequency distribution is obtained by dividing the frequency in each class by the total number of observations. True or false.**	True
139.	**Responses to the question, "How much money do you expect to spend on stereo and consumer electronics equipment in the next 12 months" represents what type of variable?**	Continuous and numerical
140.	**Responses to the question, "how tall are you" represent what type of variable?**	numerical continuous
141.	**Responses to, "was this your first purchase at good tunes?" represents what kind of variable?**	categorical
142.	**A retail chain wants to measure employee sales but one division gathered dollar amount per sale, while a second gathered the number of sales per employee. The discrepancy in data gathering is due to a failure to use the same**	operational definition
143.	**Right-Skewed**	Mean > Median Positively skewed
144.	**The Sample is**	the portion of the population selected for analysis
145.	**Sample Proportion (p) falls between**	$0 \leq p \leq 1$
146.	**Sample Proportion (p) Formula**	p = X/n = # of items in the sample having the characteristic of interest/sample size
147.	**Sample Proportion (p) provides an estimate of ?**	...

148. **sample standard deviation**	shows variation about the mean (related to the distance from the mean)...a measure of the "average" scatter around the mean the square root of the variance HAS THE SAME UNITS AS THE ORIGINAL DATA MUST BE INTERVAL/RATIO
149. **The Sample Variance**	average squared deviations of values from the mean S^2
150. **The sample will tend to be ____ if population not normally distributed?**	normally distributed
151. **The sampling distribution becomes normal as ___ does _____**	the sample size (n) increases
152. **Sampling Distribution of p is normally distributed if...**	$n\pi \geq 5$ and $n(1-\pi) \geq 5$
153. **Sampling Distribution Properties: As n increases....**	σ x bar decreases
154. **Sampling Distribution Properties: Larger sample size looks like...**	steeper, narrow graph
155. **Sampling Distribution Properties: Smaller sample size looks like**	wider, less steep graph
156. **Sampling Distribution Properties: The larger the sample size the less ____ and ____ the graph appears**	the larger the sample size the less VARIATION and NARROWER the graph appears
157. **Scatter Diagram/Plot**	used to visualize two numerical variables examine possible relationships between 2 numerical variables
158. **Scores on a management aptitude test represents a categorical variable? True or False?**	False
159. **Secondary Data**	person perfomring analysis did not collect data examples: -analyze census data -examine data from print/electronic sources -collected prior to study -time saving*
160. **Second quartile position**	Q2 = (n+1)/2 same as the median--50% of the observations are smaller and 50% are larger
161. **a sequence of data in rank order, from the smallest value to the largest is known as**	an ordered array
162. **Shape of Kurtosis**	Flatter than bell shaped (negatively skewed) Bell shaped (normal distribution) Sharper peak than bell curved (positively skewed)
163. **Skewness**	Measures the amount of asymmetry in a distribution
164. **Standard Deviation of a Discrete Random Variable is most associated with**	financial risk
165. **Standard Error of Mean**	a measure of the variability in the mean from sample to sample STANDARD ERROR OF THE MEAN DECREASES AS SAMPLE SIZE INCREASES
166. **Standard Error of Mean Formula**	σ x bar = $(\sigma/\sqrt{n})$
167. **Standardized normal distribution has a mean of ? and a standard deviation of ?**	mean of 0 standard deviation of 1
168. **Standardized test scores are an example of what kind of data**	interval scale

169.	**Stats are used in business to?**	1. Summarize business data (descriptive methods used to create charts and tables) 2. Draw conclusions from business data (inferential methods used to reach conclusions about a large group based on data from a small group) 3. Make reliable business forecasts (inferential stats used to develop, quantify, and improve the accuracy of predictive models) 4. Improve business processes (managerial approaches such as six sigma)
170.	**The stem-and leaf display is an appropriate method for graphic categorical data. True or false?**	False...the stem and leaf display is an appropriate method for graphing numerical data
171.	**stem and leaf displays show?**	how the data are distributed and where concentrations exist when visualizing numerical data
172.	**Steps in Defining the Problem**	1. Statement of hypothesis 2. Set research limits 3. Define terms 4. State assumptions 5. Importance of Study
173.	**Steps to Find P (a < X < b) when X is distributed normally:**	1. Draw the normal curve for the problem in terms of X 2. Translate x values into z values 3. use the standardized normal table
174.	**Steps to find the X value for a given probability**	1. Find the Z value for the known probability 2. Conver to x units using the formula: $X=\mu + Z\sigma$
175.	**A student's class designation, shown as Freshman, sophmore, junior, or senior represents which scale of measurement?**	ordinal scale
176.	**A student's standardized exam score represents which scale of measurement?**	interval scale
177.	**The study of patterns that may exist between two numerical variables can be accomplished by cross-tabulating the data. True or false.**	False...relationships between two numerical variables are explored by drawing scatter diagrams
178.	**SubProblems**	Each subproblem should be a stand-alone researchable unit together, should be a part of the problem and together address the main problem 2-3 MAX
179.	**A summary table in which the data are arranged numerically ordered class groups known as...**	a frequency distribution
180.	**The sum of a relative frequency column should always be equal to the sum of the frequencies. True or false?**	False...the sum of the frequencies will equal the total number of observations
181.	**A survey question which asks whether or not the respondent is "for" or "against" a proposal represents**	a categorical variable
182.	**A survey that asks consumers to indicate the brand of their primary automobile is measured on a nominal scale of measurement, true or fale?**	True...no ranking is implied
183.	**Symbol for Mean**	population parameter: μ $\bar{}$ sample statistic: X
184.	**Symbol for Standard Deviation**	population parameter: σ sample statistic: Σ (S)
185.	**Symbol for Variance**	population parameter: σ^2 sample statistic: Σ^2 (S^2)
186.	**Symmetric**	Mean = Median Normal Distribution

187. **Tables used to organize numerical data**	1. ordered array 2. frequency distribution 3. cumulative distribution
188. **Temperature is an example of what kind of data**	interval scale
189. **Third quartile position**	Q3 = 3(n+1)/4
190. **Time Series Plot**	Visualizing two numerical variables -used to study patterns in values of a numeric variable over time -vertical axis:numeric variable -horizontal axis:time period
191. **to determine the appropriate width of each class interval in a grouped frequency distribution, we:**	divide the range of the data by the number of desired class groupings
192. **A type of bar chart in which the categories are plotted in descending rank order of the magnitude of their frequencies is called a:**	pareto diagram
193. **the values located at the intersections of the rows and columns of a contingency table are called**	cells
194. **the values to be analyzed using either descriptive or inferential statistics are known as**	the data
195. **Visualizing Categorical Data**	1 variable -bar chart -pie chart -pareto chart 2 variables -side by side bar chart
196. **Visualizing Numerical Data**	ordered array--->stem & leaf display Frequency distribution & cumulative distribution-->historgram, polygon, ogive
197. **the weakest form or scale of measurement that does not allow any ranking across the various categories is**	nominal scale
198. **What does a summary table indicate about items in a set of categories?**	...
199. **What does Covariance measure?**	the strength of the linear relationship between random variables X and Y positive covariance indicates positive relationship negative covariance indicates negative relationship
200. **What does the Z Score Do/Show?**	Tells you about what distribution looks like It is the number of standard deviations a data value is from the mean if less than -3.0 or greater than +3.0, considered an outlier if normal distribution, z=0
201. **What do you assume about the contingency table in a Chi-square proportion test**	each cell in the contingency table has expected frequency of at least 5
202. **What is an ordinal scale level of measurement**	data is classified into distinct categories in which RANKING IS IMPLIED
203. **What is a sampling distribution?**	a distribution of all the possible values of a sample statistic for a given size sample selected from a population

204.	**What is important about the standardized normal distribution?**	any normal distrbution can be transformed into STANDARDIZED NORMAL DISTRIBUTION (Z) need to transform x units into z units
205.	**what is nominal scale measurement**	classifies data in categories in which NO RANKING is implied
206.	**What is the investment objective with a portfolio?**	maximize return (mean) while minimizing risk (standard deviation)
207.	**What is the major flaw with covariance?**	No limits...not possible to determine the relative strengths of the relationship from the size of the covariance
208.	**What makes up a strong hypothesis?**	1. adequate for its purpose 2. testable 3. better than your rivals
209.	**What should you use to study patterns that may exist between two or more categorical variables**	contingency tables
210.	**What table has class boundaries that are important?**	frequency distributions must not be overlapping
211.	**what type of chart is recommended to illustrate categorized responses in descending order according to their frequencies?**	the pareto diagram
212.	**when a private company uses data provided by the United States Bureau of Labor Statistics, the source of the company's data is considered to be**	a secondary source
213.	**When comparing 2 or more groups with different sample sizes you must use _____?**	relative frequency or a percentage distribution
214.	**when constructing charts, which of the following chart types is polotted at the class midpoints?**	percentage polygons
215.	**when grouping data into classes it is recommended that we have**	not fewer than 5 and no more than 15 classes
216.	**when polygons or histograms are constructed, which axis must show the true zero or "origin"?**	the vertical axis
217.	**When should you use the Chebyshev Rule?**	for heavily skewed data sets and those not appearing bell-shaped
218.	**When the number of frequencies in the two groups is not the same, you need to use ___ or ___ in order to compare the groups?**	proportions or relative frequencies and percentages
219.	**which of the four categories of data sources is used when people are asked their opinion about a particular product?**	A survey
220.	**which of the four categories of data sources is used when researchers watch shoppers browsing the aisle of a clothing retailer?**	Observation
221.	**Which of the four data sources is being used when market researcher conducts a focus group to elicit unstructured responses to open-ended questions?**	Observation
222.	**Which type of graph is appropriate when you have two numerical variables?**	scatter diagram
223.	**Why can the range be misleading?**	Ignores the way in which data are distributed Only tells you about 2 values...both of which may be extreme SENSITIVE TO OUTLIERS

224.	**Why should you use a frequency distribution?**	condenses raw data into more useful form gives a quick visual interpretation *shows where data are concentrated/clustered enables determination of major characteristics of data set
225.	**The width of each bar in a histogram corresponds to the**	width of the class boundaries in a frequency distribution
226.	**With nominal data, what techniques can you use?**	x^2 test and z-proportion test
227.	**x^2 test of independence assumption**	each cell in the contingency table has expected frequency of at least 1
228.	**x^2 Test of Independence--Hypothesis**	H0: The two categorical variables are independent. (no relationship) H1: The twow categorical variables are dependent. (relationship)
229.	**x^2 Test of Indepenence**	similar to x^2 test for equality of more than two proportions, but extends the concept to contingency tables with r rows and c columns
230.	**"Z" distribution always has...**	Mean is 0 Standard Deviation is 1 Values above the mean have POSITIVE z values Values below the mean have NEGATIVE z values
231.	**Z Value for Sampling Distribution of the Mean**	((sample mean - population mean) / (pop. std. deviation/$\sqrt{}$sample size))
232.	**α means?**	...