AP Stats Lunchtime review #2:
Univariate Data

Variables
WHAT WE'RE MEASURING

Variables
QUANTITATIVE:
numerical response

Variables
CATEGORICAL:
non-numerical response
(interpreted as percentage in inference)

DISPLAYS

DESCRIBE DISTRIBUTIONS USING SOCS
- Shape
- outliers
- center
- spread

- STEMPLOT
- DOTPLOT
- HISTOGRAM
- BOX PLOT

BAR GRAPH
PIE CHART

WHAT IS THE SHAPE?

SKEWED (has outliers)
- BOX PLOT: Min, Q_1, Med, Q_3, Max
- OUTLIERS: $Q_1 - 1.5IQR$, $Q_3 + 1.5IQR$

BEST DISPLAY

SYMMETRIC
- HISTOGRAM

MEASURE OF CENTER
- MEDIAN: RESISTANT TO OUTLIERS
- MEAN: IT IS NOT RESISTANT TO

MEASURE OF SPREAD
- IQR: Q3 − Q1
- STANDARD DEVIATION: $\sqrt{\frac{\sum (x-\bar{x})^2}{n}}$

DESCRIPTIVE STATISTICS

WHY DO WE HOPE FOR SYMMETRIC DATA?
- BECAUSE IN APPLYING THE STANDARDIZED VALUE $\left(z = \frac{x-\mu}{\sigma}\right)$ FOR ANY VALUE IN THE DATA SET FOR NORMALLY DISTRIBUTED DATA CAN LEAD US TO FIND PROBABILITY. USING THE NORMAL CURVE LEADS TO INFERENCE....

... INFERENCE ...

WHEN WE USE STATISTICS (A SAMPLE) AND THE THEORY OF SAMPLING DISTRIBUTIONS TO PREDICT

CONFIDENCE INTERVALS
- ESTIMATE PARAMETER

UNIVARIATE DATA

one sample $< \mu <$	t	
$< p <$ (2 sample)	z	
$< \mu_1 - \mu_2 <$	t	
$< p_1 - p_2 <$	z	
χ^2		

SIGNIFICANCE TESTS
- CHANGE IN PARAMETER

one sample	$H_0: \mu =$ $H_A: \mu \neq$
	$H_0: p =$ $H_A: p \neq$
2 SAMPLE	$H_0: \mu_1 - \mu_2 = 0$ $H_A: \mu_1 - \mu_2 \neq 0$
	$H_0: p_1 - p_2 = 0$ $H_A: p_1 - p_2 \neq 0$
χ^2:	H_0: distribution has no change H_A: distribution has change

BIVARIATE DATA

Slope $< \beta <$

$H_0: \beta = 0$
$H_A: \beta \neq 0$ or $\beta > 0$

AP Stats Lunchtime review #2:
Univariate Data

1. For their final project, a group of AP Statistics students wanted to compare the texting habits of males and females. They asked a random sample of students from their school to record the number of text messages sent and received over a two-day period. Here are their data:

Males: 127 44 28 83 0 6 78 6 5 213 73 20 214 28 11
Females: 112 203 102 54 379 305 179 24 127 65 41 27 298 6 130 0

(a) Determine if there are any outliers in either data set using the $1.5 \times IQR$ rule.

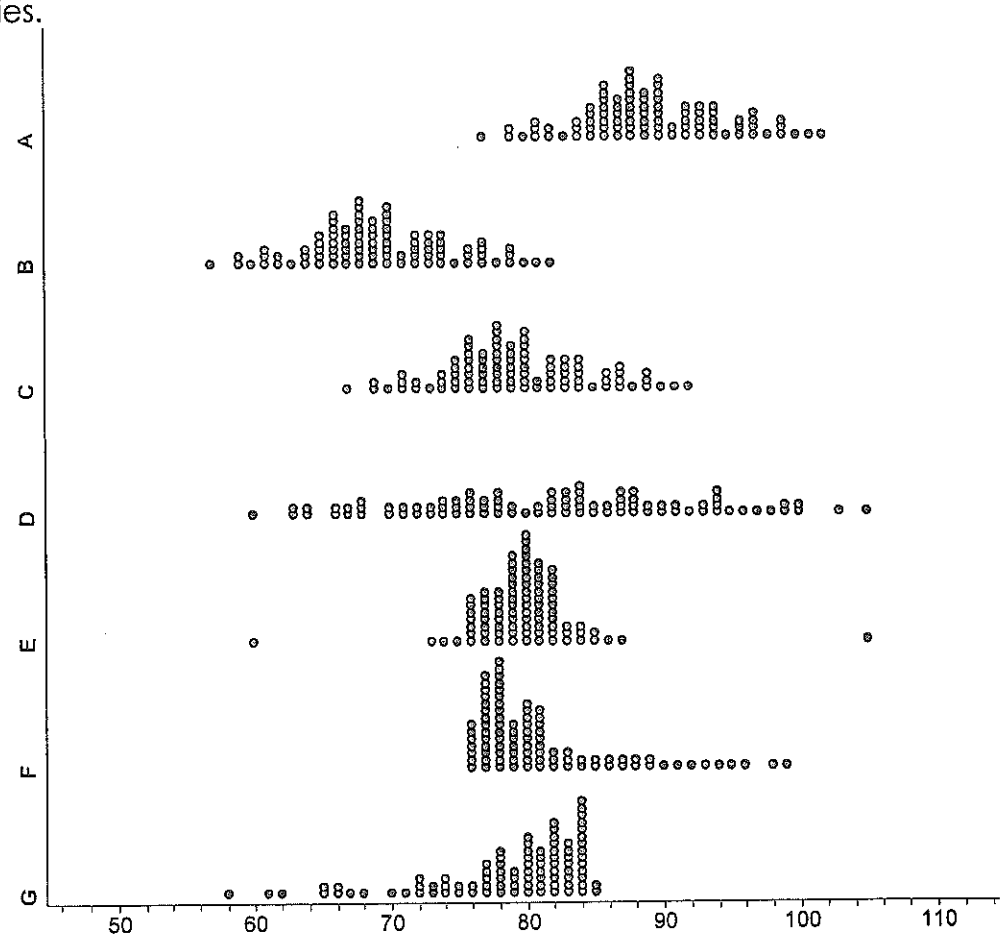
(b) Construct boxplots to compare the male and female data.

(c) Compare the distributions of text messages sent and received by males and females.

AP Stats Lunchtime review #2:
Univariate Data

2. Describing distributions of quantitative data

Brian and Jessica have decided to move and are considering seven different cities. The dotplots below show the daily high temperatures in June, July, and August for each of these cities.



a) What is the most important difference between cities A, B, and C?

b) What is the most important difference between cities C and D?

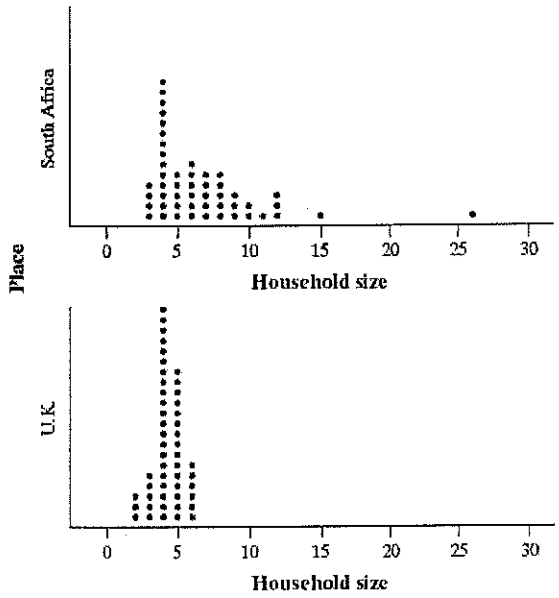
c) What is the most important difference between cities D and E?

d) What is the most important difference between cities C, F, and G?

AP Stats Lunchtime review #2:
Univariate Data

3. **Making and interpreting dotplots**

How do the numbers of people living in households in the United Kingdom (U.K.) and South Africa compare? To help answer this question, we used Census At School's "Random Data Selector" to choose 50 students from each country. Here is a comparative dot plot of the household sizes reported by the survey respondents.



Problem: Compare the distributions of household size for these two countries.

4. Making and interpreting stem-and-leaf plots

To become president of the United States, a candidate does not have to receive a majority of the popular vote. The candidate does have to win a majority of the 538 electoral votes that are cast in the Electoral College. Data on the number of electoral votes for each of the 50 states and the District of Columbia in 2015 are shown in the table below.

AL 9	AK 3	AZ 11	AR 6	CA 55	CO 9	CT 7	DE 3	DC 3	FL 29
IL 20	IN 11	IA 6	KS 6	KY 8	LA 8	ME 4	MD 10	MA 11	MI 16
MT 3	NE 5	NV 6	NH 4	NJ 14	NM 5	NY 29	NC 15	ND 3	OH 18
RI 4	SC 9	SD 3	TN 11	TX 38	UT 6	VT 3	VA 13	WA 12	WV 5
WI 10	WY 3								

Make a stemplot. Describe what you see.

5. People say that it takes a long time to get to work in New York State due to the heavy traffic near big cities. What do the data say? The stemplot shows the travel times in minutes of 20 randomly chosen New York workers. How would you summarize the center of the distribution?

0	5
1	005555
2	0005
3	00
4	005
5	
6	005
7	
8	5

Key: 4|5 is a New York worker who reported a 45-minute travel time to work.

Chapter 1 Multiple Choice Practice

Directions. *Identify the choice that best completes the statement or answers the question. Check your answers and note your performance when you are finished.*

1. You measure the age (years), weight (pounds), and breed (beagle, golden retriever, pug, or terrier) of 200 dogs. How many variables did you measure?

A. 1
B. 2
C. 3
D. 200
E. 203

2. You open a package of Lucky Charms cereal and count how many there are of each marshmallow shape. The distribution of the variable "marshmallow" is:

A. The shape: star, heart, moon, clover, diamond, horseshoe, balloon.
B. The total number of marshmallows in the package.
C. Seven—the number of different shapes that are in the package.
D. The seven different shapes and how many there are of each.
E. Since "shape" is a categorical variable, it doesn't have a distribution.

3. A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.

	Male	Female
Democrat	300	600
Republican	500	300
Other	200	100

The proportion of males that are registered as Democrats is:

A. 300
B. 30
C. 0.33
D. 0.30
E. 0.15

4. For a physics course containing 10 students, the maximum point total for the quarter was 200. The point totals for the 10 students are given in the stemplot below.

```
11 | 6  8
12 | 1  4  8
13 | 3  7
14 | 2  6
15 |
16 |
17 | 9
```

Which of the following statements is NOT true?

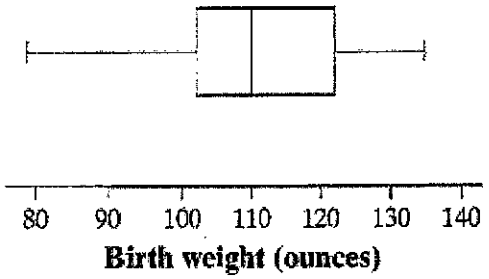
A. In a symmetric distribution, the mean and the median are equal.
B. About fifty percent of the scores in a distribution are between the first and third quartiles.
C. In a symmetric distribution, the median is halfway between the first and third quartiles.
D. The median is always greater than the mean.
E. The range is the difference between the largest and the smallest observation in the data set.

5. When drawing a histogram it is important to

A. have a separate class interval for each observation to get the most informative plot.
B. make sure the heights of the bars exceed the widths of the class intervals so that the bars are true rectangles.
C. label the vertical axis so the reader can determine the counts or percent in each class interval.
D. leave large gaps between bars. This allows room for comments.
E. scale the vertical axis according to the variable whose distribution you are displaying.

6. A set of data has a mean that is much larger than the median. Which of the following statements is most consistent with this information?
A. The distribution is symmetric.
B. The distribution is skewed left.
C. The distribution is skewed right.
D. The distribution is bimodal.
E. The data set probably has a few low outliers.

7. The following is a boxplot of the birth weights (in ounces) of a sample of 160 infants born in a local hospital.



About 40 of the birthweights were below
A. 92 ounces.
B. 102 ounces.
C. 112 ounces.
D. 122 ounces.
E. 132 ounces.

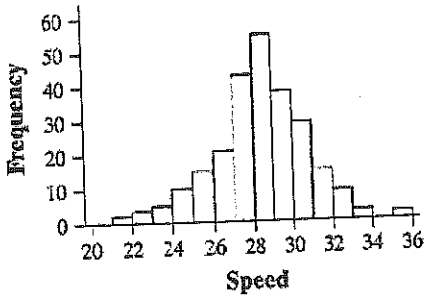
8. A sample of production records for an automobile manufacturer shows the following figures for production per shift:

705 700 690 705

The variance of the sample is approximately
A. 8.66.
B. 7.07.
C. 75.00.
D. 50.00.
E. 20.00.

9. You catch 10 cockroaches in your bedroom and measure their lengths in centimeters. Which of these sets of numerical descriptions are *all* measured in centimeters?
A. median length, variance of lengths, largest length
B. median length, first and third quartiles of lengths
C. mean length, standard deviation of lengths, median length
D. mean length, median length, variance of lengths.
E. both (B) and (C)

10. A policeman records the speeds of cars on a certain section of roadway with a radar gun. The histogram below shows the distribution of speeds for 251 cars.



Which of the following measures of center and spread would be the best ones to use when summarizing these data?

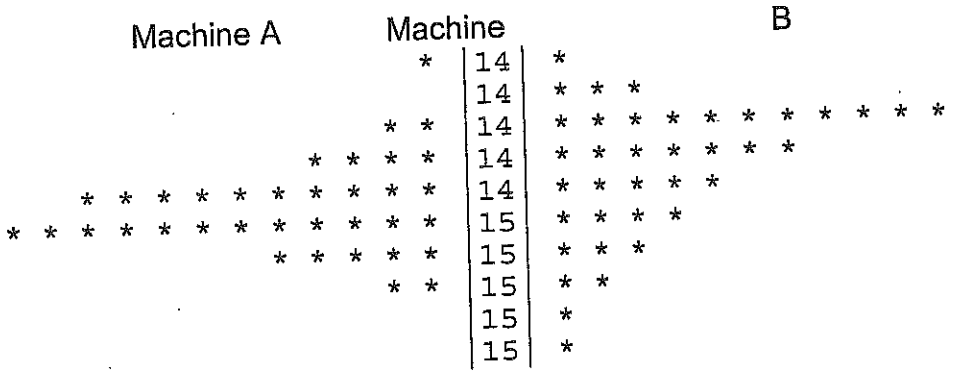
A. Mean and interquartile range
B. Mean and standard deviation
C. Median and range
D. Median and standard deviation
E. Median and interquartile range

Problem	Answer	Concept	Right	Wrong	Simple Mistake?	Need to Study More
1	C	Variables				
2	D	Categorical Variables				
3	D	Two-way table				
4	D	Distribution basics				
5	C	Constructing histograms				
6	C	Skewed distributions				
7	B	Interpreting boxplots				
8	D	Variance				
9	E	Summary statistics units				
10	B	Choosing statistics				

FRAPPY! Free Response AP Problem, Yay!

The following problem is modeled after actual Advanced Placement Statistics free response questions. Your task is to generate a complete, concise response in 15 minutes. After you generate your response, view two example solutions and determine whether you feel they are "complete," "substantial," "developing" or "minimal." If they are not "complete," what would you suggest to the student who wrote them to increase their score? Finally, you will be provided with a rubric. Score your response and note what, if anything, you would do differently to increase your own score.

SugarBitz candies are packaged in 15 oz. snack-size bags. The back-to-back plot below displays the weights (in ounces) of two samples of SugarBitz bags filled by different filling machines. The weights ranged from 14.1 oz. to 15.9 oz.



(a) Compare the distributions of weights of bags packaged by the two machines.

(b) The company wishes to be as consistent as possible when packing its snack bags. Which machine would be *least* likely to produce snack bags of SugarBitz that have a consistent weight? Explain.

(c) Suppose the company filled its bags using the machine you chose in part (b). Which measure of center, mean or median, would be closer to the advertised 15oz.? Explain why you chose this measure.

How would you score these?

Student Response 1:

a) Machine A has a slightly higher center than Machine B. Machine B has a much larger range. Machine A is approximately symmetric and Machine B is slightly skewed right. Neither machine has any extreme values.
b) Machine B would be least likely to produce bags containing 15 oz of SugarBitz because it has a much wider range than Machine A.
c) The company should report the mean weight of Machine B. Since the distribution is skewed to the right, the mean will be pulled higher towards the tail. Therefore, the mean will be higher than the median and will be closer to 15.

How would you score this response? Is it substantial? Complete? Developing? Minimal? Is there anything this student could do to earn a better score?

Student Response 2:

a) Machine A is normal and Machine B is skewed. Both have a single peak and wide ranges.
b) Machine B usually fills bags with about 14.6 oz of candy. Machine A usually fills bags with 15 oz of candy. Machine B is least likely to fill the bags with 15 oz. of candy.
c) The mean because it is about 15.

How would you score this response? Is it substantial? Complete? Developing? Minimal? Is there anything this student could do to earn a better score?

Scoring Rubric

Use the following rubric to score your response. Each part receives a score of "Essentially Correct," "Partially Correct," or "Incorrect." When you have scored your response, reflect on your understanding of the concepts addressed in this problem. If necessary, note what you would do differently on future questions like this to increase your score.

Intent of the Question

The goals of this question are (1) to determine your ability to use graphical displays to compare and contrast two distributions and (2) to evaluate your ability to use statistical information to make a decision.

Solution

(a) Both distributions are unimodal (single-peaked). However, Machine A's distribution is roughly symmetric while Machine B's is skewed to the right. The center of the weights for Machine A (median A = about 15) is slightly higher than that of Machine B (median B = about 14.5). There is more variability in the weights produced by Machine B. Machine A has one low value (14.1) that does not fall with the majority of weights. However, it does not appear to be extreme enough to be considered an outlier.

(b) Both machines produce bags of varying weight. However, Machine B has a higher variability as evidenced by a wider overall range. Machine B would be least likely to produce a consistent weight for the snack bags.

(c) The mean would be closer to the advertised 15 oz. weight. This is because in a skewed distribution, the mean is pulled away from the median in the direction of the tail. In Machine B's distribution, the peak is at about 14.5 oz so we would expect the mean to be higher and closer to 15 oz.

Scoring:

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is essentially correct if you correctly identify similarities and differences in the shape, center, and spread for the two distributions.
Part (a) is partially correct if you correctly identify similarities and differences in two of the three characteristics for the two distributions.
Part (a) is incorrect if you only identify one similarity or difference of the three characteristics for the two distributions.

Part (b) is essentially correct if Machine B is chosen using rationale based on its measure of spread of the packaged weights.
Part (b) is partially correct if B is chosen, but the explanation does not refer to the variability in the weights.
Part (c) is incorrect if B is chosen and no explanation is provided OR if A is chosen.

Part (c) is essentially correct if the mean is chosen and the explanation addresses the fact that the mean will be greater than the median in a skewed right distribution.
Part (c) is partially correct if the mean is chosen, but the explanation is incomplete or incorrect.
Part (c) is incorrect if the mean is chosen, but no explanation is given OR if the median is chosen.
NOTE: If Machine A was chosen in part (b) and the solution to part (c) indicates either the mean or median would be appropriate due to the fact that they will be approximately equal in a symmetric, mound-shaped distribution, part (c) should be scored as essentially correct.

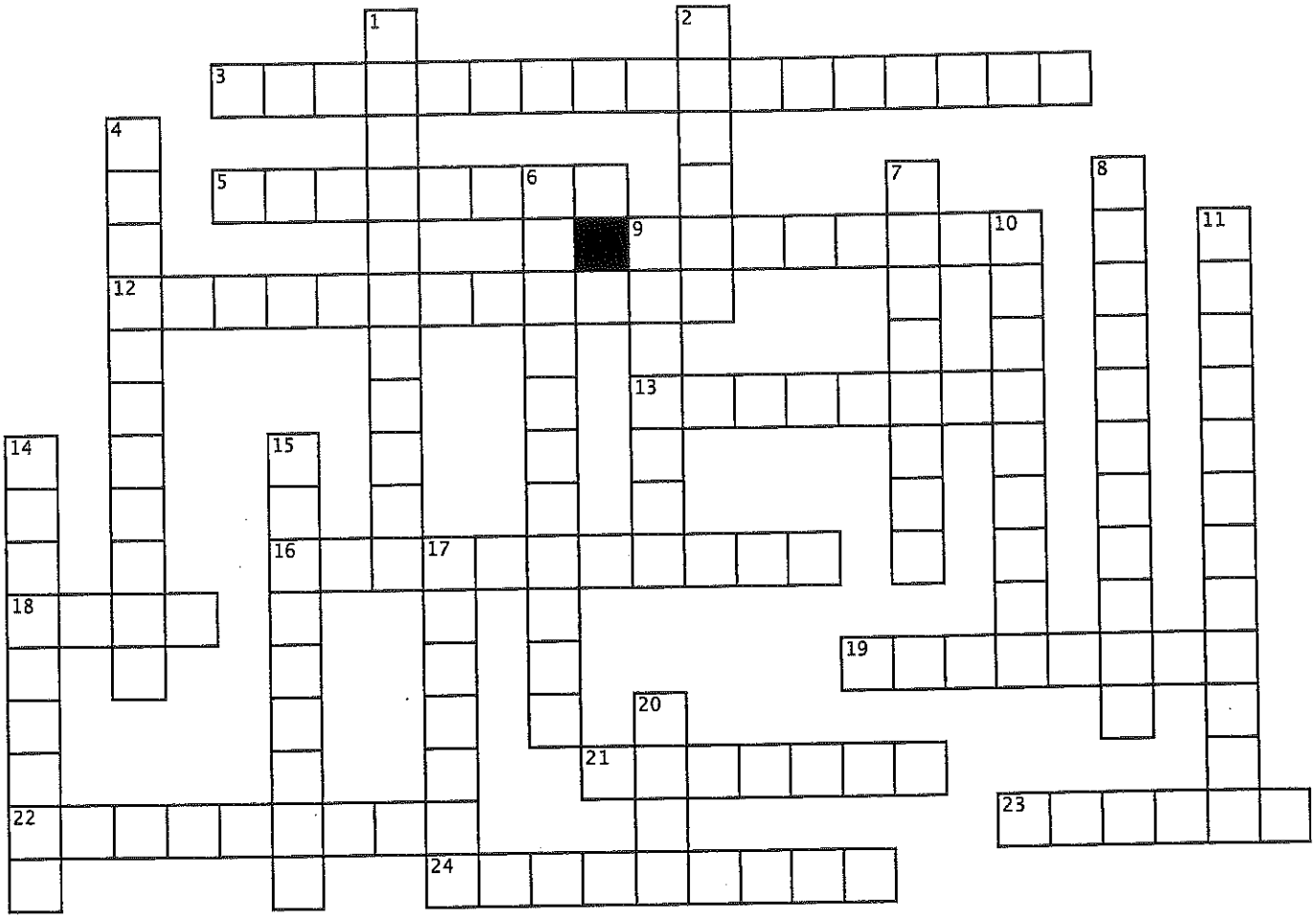
4 Complete Response
All three parts essentially correct

3 Substantial Response
Two parts essentially correct and one part partially correct

2 Developing Response
Two parts essentially correct and no parts partially correct
One part essentially correct and two parts partially correct
Three parts partially correct

1 Minimal Response
One part essentially correct and one part partially correct
One part essentially correct and no parts partially correct
No parts essentially correct and two parts partially correct

Chapter 1: Exploring Data



Across

3. The average distance of observations from their mean (two words)
5. The average squared distance of the observations from their mean
9. Displays the counts or percents of categories in a categorical variable through differing heights of bars
12. Tells you what values a variable takes and how often it takes these values
13. Displays a categorical variable using slices sized by the counts or percents for the categories
16. When specific values of one variable tend to occur in common with specific values of another
18. A measure of center, also called the average
19. A graphical display of quantitative data that involves splitting the individual values into two components
21. One of the simplest graphs to construct when dealing with a small set of quantitative data
22. Drawing conclusions beyond the data at hand
23. The shape of a distribution if one side of the graph is much longer than the other
24. What we call a measure that is relatively unaffected by extreme observations

Down

1. The objects described by a set of data
2. The midpoint of a distribution of quantitative data
4. A _____ distribution describes the distribution of values of a categorical variable among individuals who have a specific value of another variable.
6. A variable that places an individual into one of several groups or categories
7. A characteristic of an individual that can take different values for different individuals
8. When comparing two categorical variables, we can orgainze the data in a ___-___ _____.
9. A graphical display of the five-number summary
10. A graphical display of quantitative data that shows the frequency of values in intervals by using bars
11. A variable that takes numerical values for which it makes sense to find an average
14. The shape of a distribution whose right and left sides are approximate mirror images of each other
15. These values lie one-quarter, one-half, and three-quarters of the way up the list of quantitative data
17. A value that is at least 1.5 IQRs above the third quartile or below the first quartile
20. When exploring data, don't forget your ____

Chapter 2 Multiple Choice Practice

Directions. *Identify the choice that best completes the statement or answers the question. Check your answers and note your performance when you are finished.*

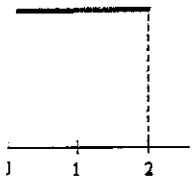
1. The 16th percentile of a Normally distributed variable has a value of 25 and the 97.5th percentile has a value of 40. Which of the following is the best estimate of the mean and standard deviation of the variable?

A. Mean ≈ 32.5; Standard deviation ≈ 2.5
B. Mean ≈ 32.5; Standard deviation ≈ 5
C. Mean ≈ 32.5; Standard deviation ≈ 10
D. Mean ≈ 30; Standard deviation ≈ 2.5
E. Mean ≈ 30; Standard deviation ≈ 5

2. The proportion of observations from a standard Normal distribution that take values larger than 0.75 is about

A. 0.2266
B. 0.2500
C. 0.7704
D. 0.7764
E. 0.8023

nsity curve below takes the value 0.5 on the interval $0 < x < 2$ and takes the value 0 everywhere at percent of the observations lie between 0.4 and 1.08?



A. 25%
B. 34%
C. 50%
D. 68%
E. 70%

listribution of the heights of students in a large class is roughly Normal. The average height is 68 inches, and approximately 99.7% of the heights are between 62 and 74 inches. Thus, the variance of the height distribution is approximately equal to

A. 2
B. 3
C. 4
D. 6
E. 9

5. The mean age (at inauguration) of all U.S. Presidents is approximately Normally distributed with a mean of 54.6. Barack Obama was 47 when he was inaugurated, which is the 11th percentile of the distribution. George Washington was 57. What percentile was he in?

A. 6.17
B. 65.17
C. 62.92
D. 34.83
E. 38.9

6. Which of the following statements are false?
I. The standard Normal table can be used with z-scores from any distribution
II. The mean is always equal to the median for any Normal distribution.
III. Every symmetric, bell-shaped distribution is Normal
IV. The area under a Normal curve is always 1, regardless of the mean and standard deviation.

A. I and II
B. I and III
C. II and III
D. III and IV
E. None of the above gives the correct set of true statements.

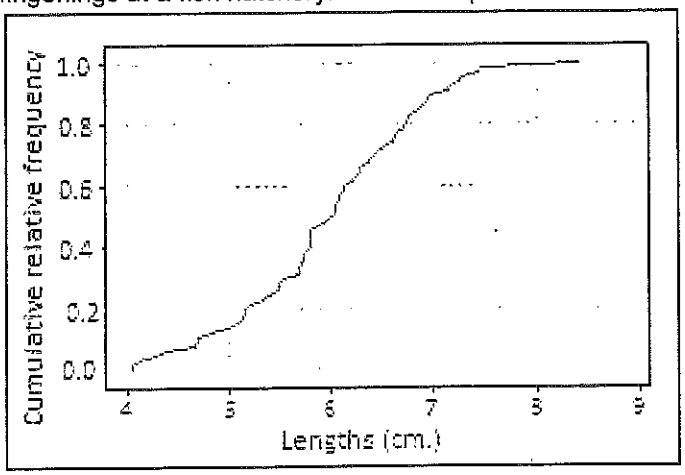
7. High school textbooks don't last forever. The lifespan of all high school statistics textbooks is approximately Normally distributed with a mean of 9 years and a standard deviation of 2.5 years. What percentage of the books last more than 10 years?

A. 11.5%
B. 34.5%
C. 65.5%
D. 69%
E. 84.5%

8. The distribution of the time it takes for different people to solve their Strive for a Five chapter crossword puzzle is strongly skewed to the right, with a mean of 10 minutes and a standard deviation of 2 minutes. The distribution of z-scores for those times is

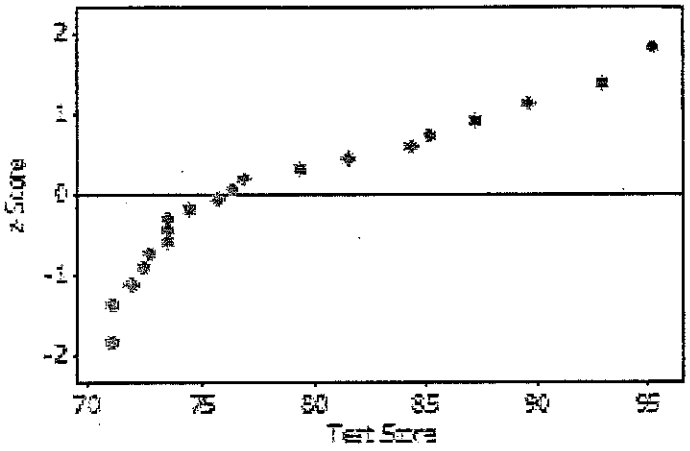
A. Normally distributed, with mean 10 and standard deviation 2.
B. Skewed to the right, with mean 10 and standard deviation 2.
C. Normally distributed, with mean 0 and standard deviation 1.
D. Skewed to the right, with mean 0 and standard deviation 1.
E. Skewed to the right, but the mean and standard deviation cannot be determined without more information.

9. The cumulative relative frequency graph below shows the distribution of lengths (in centimeters) of fingerlings at a fish hatchery. The third quartile for this distribution is approximately:



A. 6.7 cm
B. 7 cm
C. 6 cm
D. 5.5 cm
E. 7.5 cm

10. The plot shown below is a Normal probability plot for a set of test scores. Which statement is true for these data?



A. The distribution is clearly Normal.
B. The distribution is approximately Normal.
C. The distribution appears to be skewed.
D. The distribution appears to be uniform.
E. There is insufficient information to determine the shape of the distribution.

Problem	Answer	Concept	Right	Wrong	Simple Mistake?	Need to Study More
1	E	68-95-99.7 Rule				
2	A	Standard Normal Table				
3	B	Area under a Density Curve				
4	C	68-95-99.7 Rule and Variance				
5	B	Standard Normal Calculations				
6	B	Properties of Normal Distributions				
7	B	Standard Normal Calculations				
8	D	Standardized Scores				
9	A	Cumulative Relative Frequency Graph				
10	C	Normal Probability Plots				

FRAPPY! Free Response AP Problem, Yay!

The following problem is modeled after actual Advanced Placement Statistics free response questions. Your task is to generate a complete, concise response in 15 minutes. After you generate your response, view two example solutions and determine whether you feel they are "complete," "substantial," "developing," or "minimal." If they are not "complete," what would you suggest to the student who wrote them to increase their score? Finally, you will be provided with a rubric. Score your response and note what, if anything, you would do differently to increase your own score.

Final exam grades are determined by the percent correct on the exam. A teacher's records indicate the performance on the exam is Normally distributed with mean 82 and standard deviation 5. The grades on her exam are assigned using the scale below.

Grade	Percent Correct
A	$94 \leq$ percent ≤ 100
B	$85 \leq$ percent < 94
C	$76 \leq$ percent < 85
D	$65 \leq$ percent < 76
F	$0 \leq$ percent < 65

(a) Use a sketch of a Normal distribution to illustrate the proportion of students who would earn a B. Calculate this proportion.

(b) Students who earn a B, C, or D, are considered to "meet standards." Based on this grading scale, what percent of students will receive a score that places them in a category other than "meets standards"?

(c) What grade would the student who scored at the 25th percentile earn on this chapter? Justify your answer.

How would you score these?

Student Response 1:

a) P(B) = 0.9918 – 0.7257 = 0.2661. 26.61 % of students earn a B.
b) P(does not meet standards) = P(F) = P(z<3.4) = 0.0003
c) z = 0.25 so the score would be 82 + 0.25(5) = 83.25

How would you score this response? Is it substantial? Complete? Developing? Minimal? Is there anything this student could do to earn a better score?

Student Response 2:

a) P(B) = P($0.6 \leq z < 2.4$) = 0.9793 – 0.5239 = 0.4554
b) P(A or F) = P($z > 2.4$) + P($z < 3.4$) = 0.0085
c) A z-score of -0.6745 corresponds to the 25^{th} percentile. So, the score would be 82+ (-0.6475)(5) = 78.8.

How would you score this response? Is it substantial? Complete? Developing? Minimal? Is there anything this student could do to earn a better score?

Scoring Rubric

Use the following rubric to score your response. Each part receives a score of "Essentially Correct," "Partially Correct," or "Incorrect." When you have scored your response, reflect on your understanding of the concepts addressed in this problem. If necessary, note what you would do differently on future questions like this to increase your score.

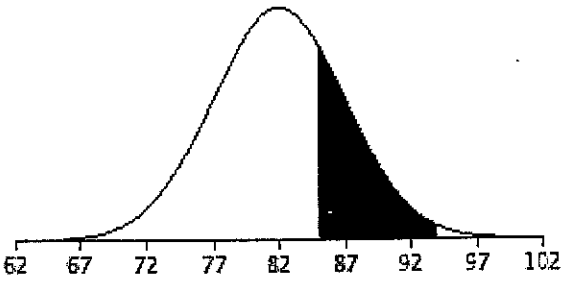
Intent of the Question

The goal of this question is to determine your ability to perform and interpret Normal calculations.

Solution

(a)

$85 \le \text{percent} < 94$)
$= P((85-82)/5 \le z < (94-82)/5)$
$= P(0.6 \le z < 2.4)$
$= 0.9918 - 0.7257$
$= 0.2661$



(b) P(A or F) = P(A) + P(F)
$= P(z \ge (94-82)/5) + P(z < (65-82)/5)$
$= P(z \ge 2.4) + P(z < 3.4)$
$= 0.0082 + 0.0003$
$= 0.0085$

(c) A z-score of -0.6745 corresponds to the 25th percentile.
x = mean + z(std dev)
x = 82+ (-0.6475)(5)
x = 78.8

Scoring:

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is essentially correct if the response (1) recognizes the need to look at grades of A and F and (2) correctly computes the tail probabilities and adds them together.
Part (a) is partially correct if the response
Considers only an A or an F and calculates the corresponding tail area correctly
Recognizes the need to look at A and F but only calculates one of the tail areas correctly
Approximates the probabilities using the Empirical rule
Computes the proportion that will "meet standards"
States the correct answer 0.0085 without supporting work

Part (b) is essentially correct if (1) the appropriate probability is illustrated using a labeled Normal curve and (2) the proportion is correctly computed.
Part (b) is partially correct if only one of the above elements is correct.

Part (c) is essentially correct if the student recognizes the situation as binomial and identifies p from part (b) and shows work to calculate the correct probability.
Part (c) is partially correct if the student recognizes the situation as binomial and identifies p, but does not compute the correct probability OR gives the correct probability but does not show work.

NOTE: If the student makes an error in part (b) and correctly uses that probability in part (c) to compute a reasonable probability, part (c) is essentially correct.

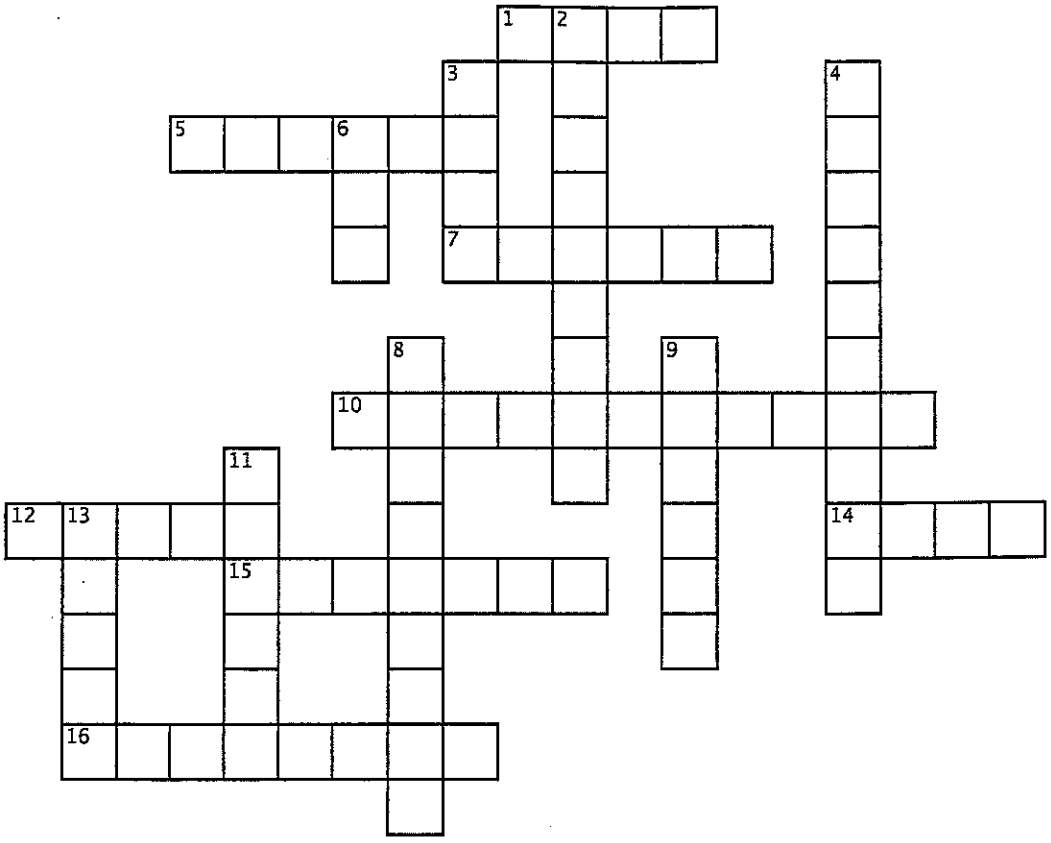
4 Complete Response
All three parts essentially correct

3 Substantial Response
Two parts essentially correct and one part partially correct

2 Developing Response
Two parts essentially correct and no parts partially correct
One part essentially correct and two parts partially correct
Three parts partially correct

1 Minimal Response
One part essentially correct and one part partially correct
One part essentially correct and no parts partially correct
No parts essentially correct and two parts partially correct

Chapter 2: Modeling Distributions of Data



Across

1. The balance point of a density curve, if it were made of solid material
5. The standardized value of an observation
7. These common density curves are symmetric and bell-shaped
10. A Normal _____ plot provides a good assessment of whether a data set is approximately Normally distributed
12. Another name for a cumulative relative frequency graph
14. The standard Normal table tells us the area under the standard Normal curve to the ___ of z
15. A ___ curve is a smooth curve that can be used to model a distribution
16. This Normal distribution has mean 0 and standard deviation 1

Down

2. The ____ rule is also known as the 68-95-99.7 rule for Normal distributions
3. To standardize a value, subtract the ___ and divide by the standard deviation
4. The value with p percent of the observations less than it
6. The area under any density curve is always equal to
8. We ___ data when we change each value by adding a constant and/or multiplying by a constant.
9. If a Normal probability plot shows a _____ pattern, the data are approximately Normal
11. The point that divides the area under a density curve in half
13. This mathematician first applied Normal curves to data to errors made by astronomers and surveyors