

Name \_\_\_\_\_  
8A; Algebra 1

Date \_\_\_\_\_  
Period \_\_\_\_\_

### Review for your Statistics Test

Please make sure to focus on the following questions to be successful on your Statistics Exam;

#'s:

3,4,7,9,

12,13,15,18,

19,20,21,22,

23,26a & c,28,

30, \*\*\*\*this is the one most people get wrong.

31a,33,34,35,



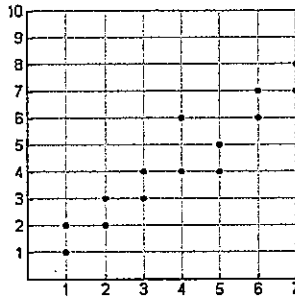
Name: \_\_\_\_\_

Date: \_\_\_\_\_

## Review for Statistics #1

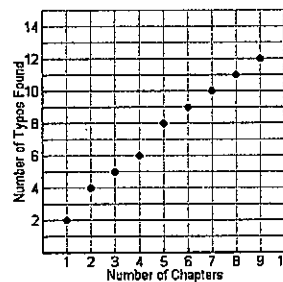
1) The correlation seen in the graph at the right would be best described as:

- (1) high positive correlation
- (2) low positive correlation
- (3) high negative correlation
- (4) low negative correlation



2) If you were asked to interpolate information from this graph, you would have to be careful to limit the number of chapters to:

- (1) between 2 and 8 chapters
- (2) between 1 and 9 chapters
- (3) chapters 2, 4, 6, and 8 only
- (4) there is no need to limit chapters



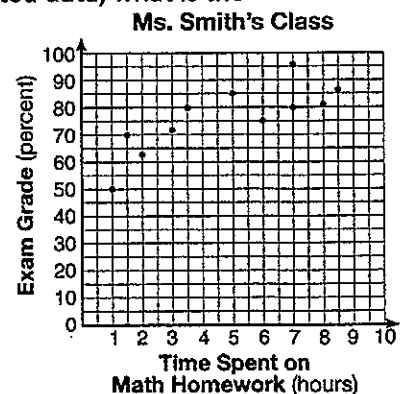
- 3) Which of the following does not describe a causal relationship?
- (1) the faster the runner's pace, the quicker he will finish the race.
  - (2) the rooster crows, the sun rises.
  - (3) the more miles driven, the more gas will be used.
  - (4) the more powerful the microwave, the faster the food cooks.

4) Which situation describes a negative correlation?

- (1) the amount of gas left in a car's tank and the amount of gas used from it
- (2) the number of gallons of gas purchased and the amount paid for the gas
- (3) the size of a car's gas tank and the number of gallons it holds
- (4) the number of miles driven and the amount of gas used

5) The number of hours spent on math homework during one week and the math exam grades for eleven students in Ms. Smith's algebra class are plotted below. Based on the plotted data, what is the correlation between the time spent on homework and the exam grade?

- (1) positive
- (2) negative
- (3) no correlation
- (4) cannot be determined



6) Identify data as univariate or bivariate data:

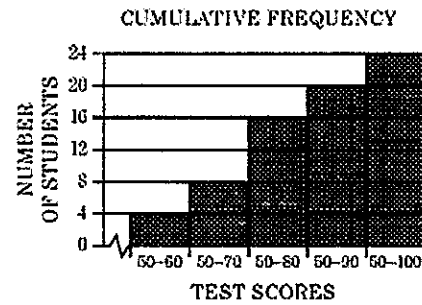
a)

Age (yr)	Frequency
14	12
15	21
16	14
17	19
18	15

b)

Time Spent Studying (hr)	Test Grade (%)
1	65
2	72
3	83
4	85
5	92

7) The accompanying frequency histogram shows the scores that 24 students received on a science test. How many students have scores of 61-70?



8) The table below indicates the grades from a pre-assessment on graphing in Ms. Cress's class.

- Complete the cumulative frequency table.
- What percent of students failed the pre-assessment if a passing score was 65%?  
(Round to the nearest percent)

Interval	Frequency	Cumulative Intervals	Cumulative Frequency
75-79	5		
70-74	6		
65-69	3		
60-64	10		
55-59	1		
50-54	5		

9) In Australia, a study of all farms with 30 or fewer sheep produced the following data.

- What is the mean for the number of sheep per farm?
- Find the standard deviation to the nearest tenth.
- Find the variance to the nearest hundredth.

Number of Sheep per Farm	Number of Farms
15	6
20	3
22	5
25	4
30	2

10) If the variance of a set of data is 25, what is the standard deviation?

11) A rapidly growing bacteria has been discovered. Its growth rate is shown in the chart.

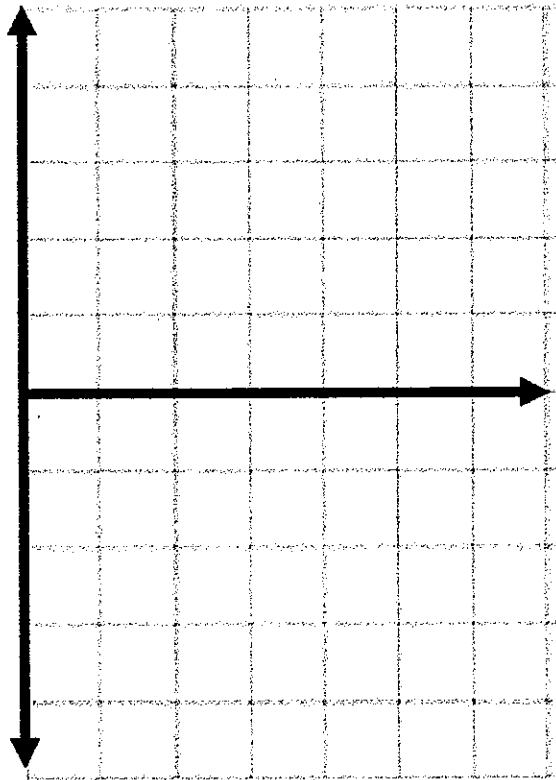
a. Calculate the Linear Regression Model for the Data. Round to the nearest hundredth.

b. Find the residuals for the predicted values not completed. Round to the nearest tenth.

Hours since observation began	Number of bacteria in the sample
0	20
1	40
2	75
3	150
4	297
5	510

Hours since observation began	# of bacteria	Predicted Value	Residual Value
0	20		
1	40		
2	75		
3	150		
4	297		
5	510		

c. Create a residual plot for the model.

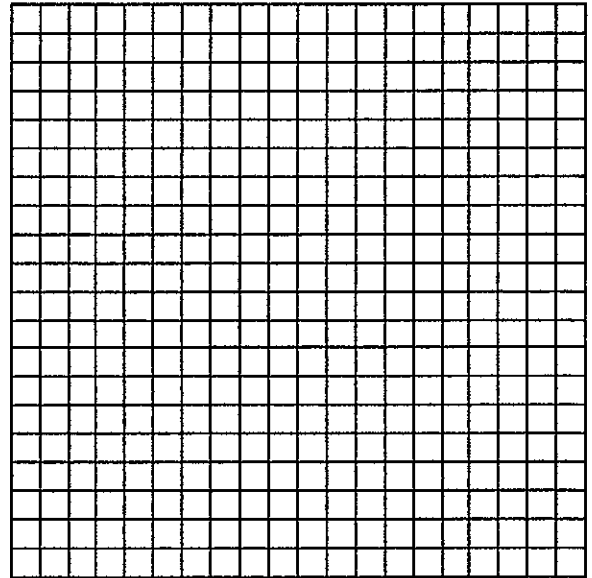


d. Use the residual plot to determine if the Regression line (Line of Best Fit) is a good fit for this data. Explain.

12) The table shows study times of 8 students and their related test scores.

a) Draw a scatter plot for the following data.

Study Time (min)	Test Score
19	66
24	65
26	70
32	73
34	71
39	74
44	75
46	80



b) Find the best trend line (line of best fit), rounding values to the nearest hundredth.

c) Using the equation you found in part b, predict how many minutes a student has to study to receive a grade of 84. Round to the nearest whole number.

d) Using the equation you found in part b, predict the test score of a student who studied for 90 minutes. Round to the nearest whole number.

e) Were you interpolating or extrapolating data for part d? Explain.

f) What is the value of the correlation coefficient to the nearest ten thousandth? Explain the meaning.

13) Which set of data of temperatures has the largest dispersion as measured by its interquartile range?

(1) 15, 17, 19, 21, 21, 22, 28

(3) 10, 19, 22, 23, 23, 29, 44

(2) 21, 23, 36, 37, 44, 48, 50

(4) 42, 47, 49, 50, 52, 59, 60

14) Calculate the range for the data: 15, 20, 22, 25, 30

15) Mark off the appropriate box with a check or X.

a) Determine which of the following would represent **qualitative** or **quantitative** data

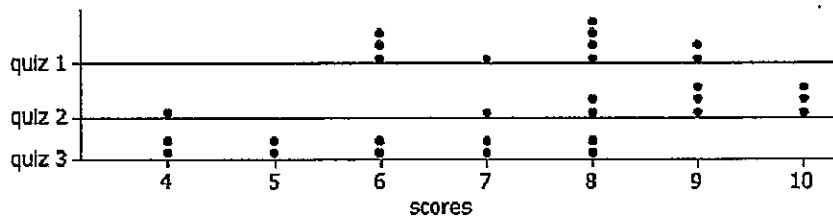
	Qualitative	Quantitative
SAT Scores		
Nationality		
Color		
Cholesterol Level		

b) Which group being surveyed is **biased** or **unbiased** when determining a schools favorite sport?

	Biased	Unbiased
Football Team		
Randomly selecting students from a lunch line		
Every 3 <sup>rd</sup> girl that enters the building		
Surveying 5 random classes from each grade level		

16) The scores of three quizzes are shown in the following data plot for a class of 10 students.

Each quiz has a maximum possible score of 10. Possible dot plots of the data are shown below.



a) Without performing any calculations, which quiz has the lowest standard deviation in the students' scores?

Justify your response.

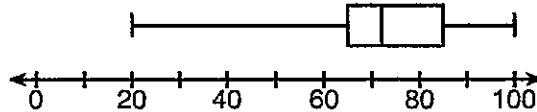
b) If you were to calculate a measure of central tendency for Quiz 1, would you recommend using the mean or median? Explain your choice.

17) Which statement is true about the data set 4, 5, 6, 6, 7, 9, 12?

- (1) mean = mode
- (2) mode = median
- (3) mean < median
- (4) mode > mean

18) a. The box-and-whisker plot below represents the results of tests scores in a math class.  
What do the scores 65, 85, and 100 represent?

- (1)  $Q_1$ , median,  $Q_3$
- (2)  $Q_1$ ,  $Q_3$ , maximum
- (3) median,  $Q_1$ , maximum
- (4) minimum, median, maximum



b. What percent of the data do the test scores from 20 to 72 represent? Explain how you arrived at your answer.

19) Carrie has grades of 84, 65, and 76 on three social studies tests. What grade must she obtain on the next test to have an average of exactly 80 for the four tests? Show how you arrived at your answer algebraically.

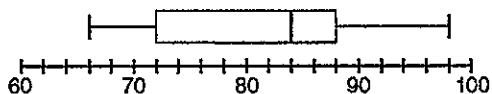
20) Create a box plot (box and whiskers) for the following data: 65, 90, 100, 72, 88, 55, 73



21) Which situation should be analyzed using bivariate data?

- 1) Ms. Saleem keeps a list of the amount of time her daughter spends on her social studies homework.
- 2) Mr. Benjamin tries to see if his students' shoe sizes are directly related to their heights.
- 3) Mr. DeStefan records his customers' best video game scores during the summer.
- 4) Mr. Chan keeps track of his daughter's algebra grades for the quarter.

22) The box-and-whisker plot below represents the math test scores of 20 students.



What percentage of the test scores are *less than 72*?

- 1) 25
- 2) 50
- 3) 75
- 4) 100

23) Which data set describes a situation that could be classified as qualitative?

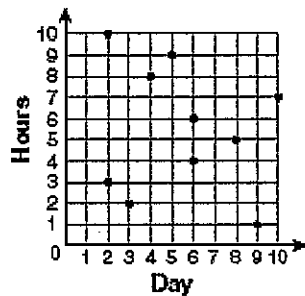
- 1) the elevations of the five highest mountains in the world
- 2) the ages of presidents at the time of their inauguration
- 3) the opinions of students regarding school lunches
- 4) the shoe sizes of players on the basketball team

24) The mean of five consecutive integers is 55. What are the integers?

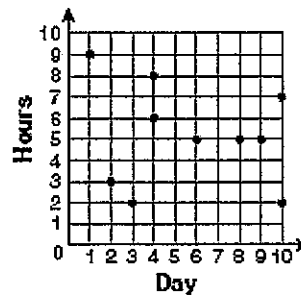
25) For 10 days, Romero kept a record of the number of hours he spent listening to music. The information is shown in the table below.

Day	1	2	3	4	5	6	7	8	9	10
Hours	9	3	2	6	8	6	10	4	5	2

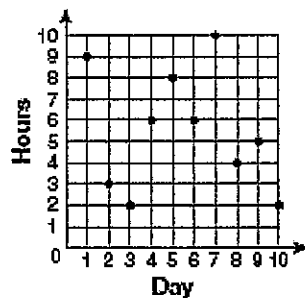
Which scatter plot shows Romero's data graphically?



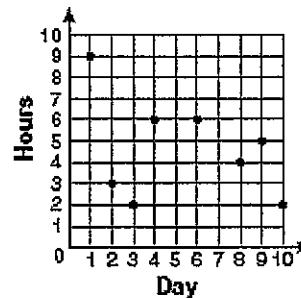
(1)



(3)



(2)



(4)

26) Twenty students were surveyed about the number of days they played outside in one week. The results of this survey are shown below.

{6,5,4,5,0,7,1,5,4,4,3,2,2,3,2,4,3,4,0,7}

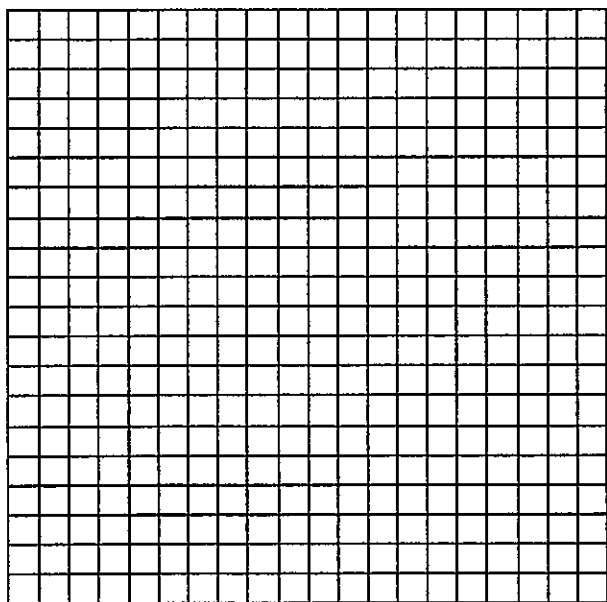
(a) Complete the frequency table below for these data.

Interval	Tally	Frequency
0-1		
2-3		
4-5		
6-7		

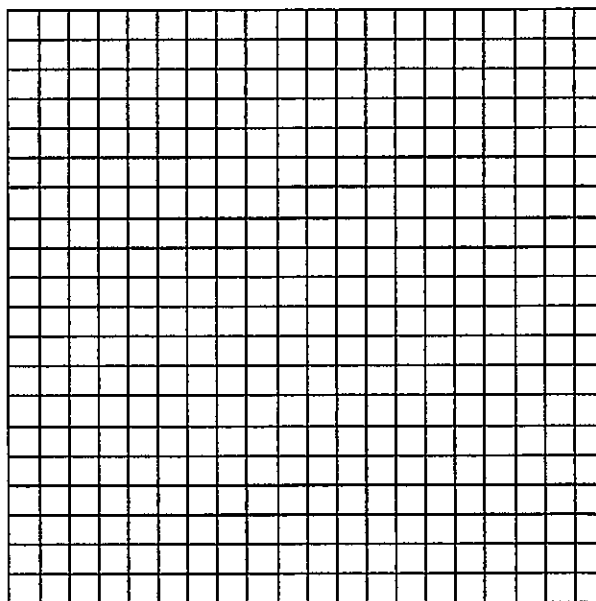
(b) Complete the cumulative frequency table below using these data.

Interval	Cumulative Frequency
0-1	
0-3	
0-5	
0-7	

c) create a frequency histogram



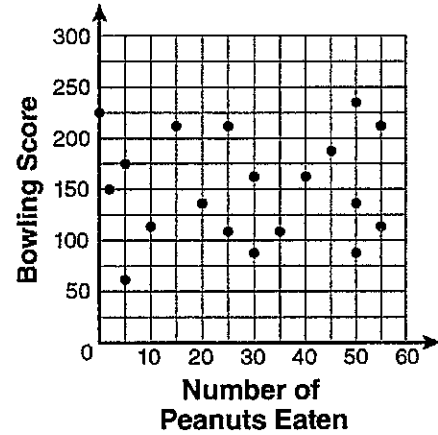
d) create a cumulative frequency histogram



c) which interval contains the 40th percentile?

27) The scatter plot below represents the relationship between the number of peanuts a student eats and the student's bowling score. Which conclusion about the scatter plot is valid?

- (1) There is almost no relationship between eating peanuts and bowling score.
- (2) Students who eat more peanuts have higher bowling scores.
- (3) Students who eat more peanuts have lower bowling scores.
- (4) No bowlers eat peanuts.



28) Jenna has bought a new hybrid car. Each week for a year, she has noted the fuel efficiency (in miles per gallon) of her car. The results for 7 weeks are shown below:

45 44 43 44 45 44 43

- a) Determine if this data set is a population or a sample
- b) Find the standard deviation to the nearest hundredth.
- c) Find the variance to the nearest hundredth

29) The data below shows the comparison of the price of a particular product with the number of units sold.

If the least-squares line is  $y = 2x + 53$ , determine the residuals for the last two data sets and complete the table accordingly. Show all work below.

Price Per Unit (Dollars)	Units Sold	Predicted Units Sold	Residual
32	100	117	-17
34	144	121	23
36	130		
38	118		

30) The table shows math scores from Smith's math class:

Interval	Frequency	Cumulative Frequency
61-70	2	
71-80	3	
81-90	8	
91-100	7	

a) Complete the table.

b) What is the modal interval?

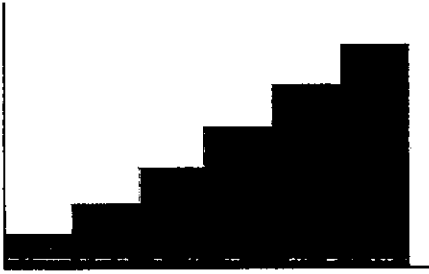
c) What is the interval containing the median?

d) What interval contains the upper quartile?

e) What percent of the students scored 81-100?

31) For the following graphs, a) describe the distribution as symmetrical, skewed, uniform, or U-shaped.  
b) state whether you would use the mean, median, or neither to represent a typical value.

a)



b)

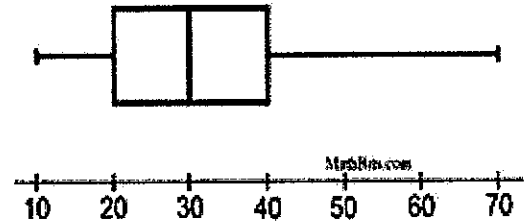


32)

The box plot shown at the right is described as being a \_\_\_\_\_.

Choose:

- a) symmetric distribution
- b) skewed left distribution
- c) skewed right distribution

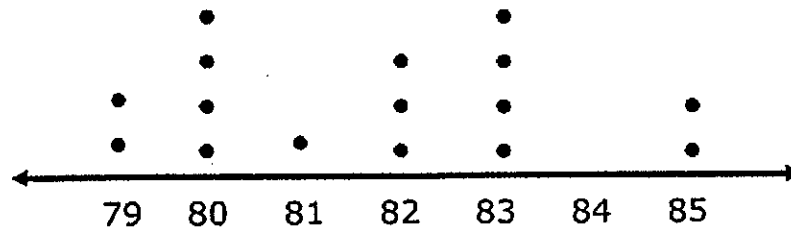


33) A radio station did a survey to determine what kind of music to play by taking a sample of middle school, high school, and college students. They were asked which of the three different types of music they prefer on the radio: hip-hop, alternative, or classic rock. The results are summarized in the table below.

	Hip-Hop	Alternative	Classic Rock
Middle School	28	18	4
High School	22	22	6
College	16	20	14

What percent of college students prefer classic rock?

34) The resting pulse rates were recorded for 16 boys in gym class before they exercised. The dot plot here shows the results.

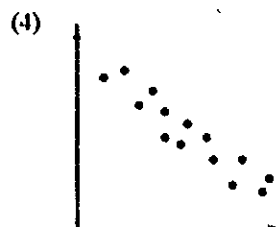
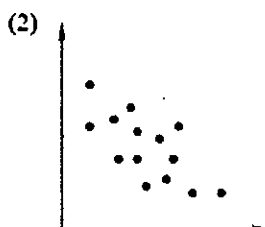
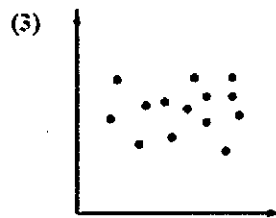
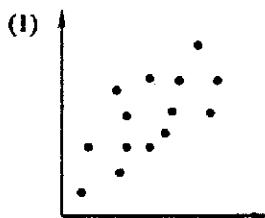


a) What is the range (spread) of the pulse rates?

b) What is the mode of the pulse rates?

c) What is the median of the pulse rates?

35) Which of the following scatter plots would have a correlation coefficient closest to -1?







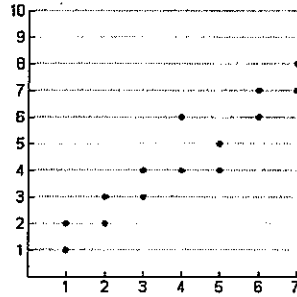
Name: Key

Date: \_\_\_\_\_

# Review for Statistics #1

1) The correlation seen in the graph at the right would be best described as:

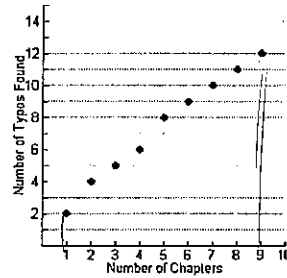
- (1) high positive correlation
- (2) low positive correlation
- (3) high negative correlation
- (4) low negative correlation.



2) If you were asked to interpolate information from this graph, you would have to be careful to limit the number of chapters to:

- (1) between 2 and 8 chapters
- (2) between 1 and 9 chapters
- (3) chapters 2, 4, 6, and 8 only
- (4) there is no need to limit chapters

fall within the plotted values



3) Which of the following does not describe a causal relationship?

- (1) the faster the runner's pace, the quicker he will finish the race.
- (2) the rooster crows, the sun rises.
- (3) the more miles driven, the more gas will be used.
- (4) the more powerful the microwave, the faster the food cooks.

100% causal

4) Which situation describes a negative correlation?

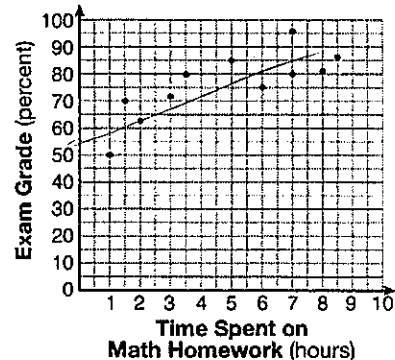


As the x increases, the y decreases

- (1) the amount of gas left in a car's tank and the amount of gas used from it
- (2) the number of gallons of gas purchased and the amount paid for the gas
- (3) the size of a car's gas tank and the number of gallons it holds
- (4) the number of miles driven and the amount of gas used

5) The number of hours spent on math homework during one week and the math exam grades for eleven students in Ms. Smith's algebra class are plotted below. Based on the plotted data, what is the correlation between the time spent on homework and the exam grade?

Ms. Smith's Class



- (1) positive
- (2) negative
- (3) no correlation
- (4) cannot be determined

6) Identify data as univariate or bivariate data:

a)

Age (yr)	Frequency
14	12
15	21
16	14
17	19
18	15

Univariate

b)

Time Spent Studying (hr)	Test Grade (%)
1	65
2	72
3	83
4	85
5	92

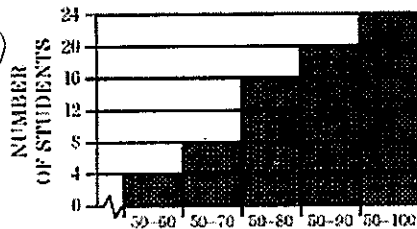
bivariate

7) The accompanying frequency histogram shows the scores that 24 students received on a science test.

How many students have scores of 61-70? when going from Regular to Cumulative you add the #?

Cumulative regular  
 $50-60 = 4$   
 $50-70 = 8$   
 $50-80 = 16$   
 $50-90 = 20$   
 $50-100 = 24$

61-70 = 4  
 Subtract  $8-4 = 4$   
4 students



when going from Cumulative to regular you subtract the #?

8) The table below indicates the grades from a pre-assessment on graphing in Ms. Cress's class.

- Complete the cumulative frequency table.
- What percent of students failed the pre-assessment if a passing score was 65%? (Round to the nearest percent)

$\frac{16}{30} = 53\%$

Interval	Frequency
75-79	5
70-74	6
65-69	3
60-64	10
55-59	1
50-54	5

Cumulative Intervals	Cumulative Frequency
50-79	30
50-74	25
50-69	19
50-64	16
50-59	6
50-54	5

9) In Australia, a study of all farms with 30 or fewer sheep produced the following data.

- What is the mean for the number of sheep per farm?

$\bar{x} = \frac{420}{20} = 21$

- Find the standard deviation to the nearest tenth.

STAT → CALC  
 1σ VarStat |  $\sigma x = 4.7$   
 VARS | 5: STATISTICS  
 $(\sigma x)^2 = 22.50$

population  $\sigma x$  (all farms)

Number of Sheep per Farm	Number of Farms	Frequency
15	6	= 90
20	3	= 60
22	5	= 110
25	4	= 100
30	2	= 60

- Find the variance to the nearest hundredth.

standard deviation has to do with the spread of the #s

When putting this into the calc you have two choices

10) If the variance of a set of data is 25, what is the standard deviation?

$V = SD^2$   
 $\sqrt{25} = SD$   
 $SD = 5$

Ex 15 } 6 times  
 15 }  
 15 }  
 15 }  
 15 }  
 15 }  
 20 } 3 times  
 20 }  
 20 }  
 20 }

- Put # of sheep in  $L_1$  + # of farms in  $L_2$  + go to 1-var stats + put  $L_2$  into frequency OR
- Put each # of sheep the amount of freq. into  $L_1$  + go to 1-var stats +

STAT 1:Edit put #'s into L1 + L2

11) A rapidly growing bacteria has been discovered. Its growth rate is shown in the chart.

a. Calculate the Linear Regression Model for the Data. Round to the nearest hundredth.

STAT CALC 4: LinReg

$y = 94.17x - 53.43$

b. Find the residuals for the predicted values not completed. Round to the nearest tenth.

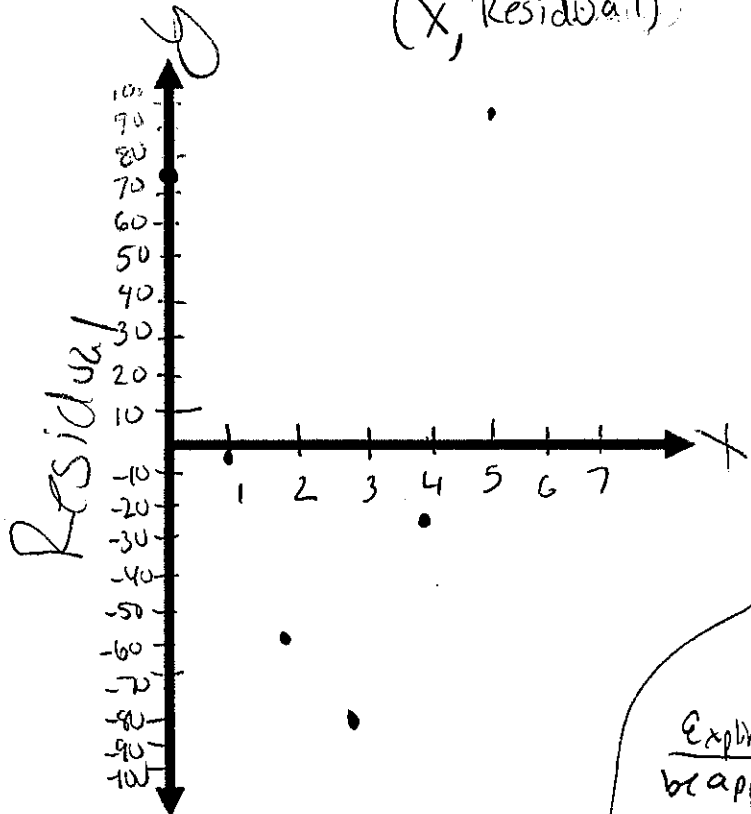
$y_2$  plug into  $y = 94.17x - 53.43$   
 $y_1 - y_2$

L1 Hours since observation began	L2 Number of bacteria in the sample
0	20
1	40
2	75
3	150
4	297
5	510

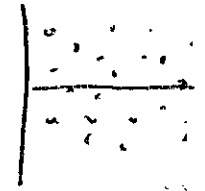
X	$y_1$ # of bacteria Actual	$y_2$ Predicted Value	Residual Value Actual - Predicted
0	20	-53.4	73.4
1	40	40.7	-.7
2	75	134.9	-59.9
3	150	229.1	-79.1
4	297	323.3	-26.3
5	510	417.4	92.6

c. Create a residual plot for the model.

Graph:  
(X, Residual)



★ If it was random scatter that would be a good fit



Explanation: A linear regression model would be appropriate for the data b/c there is NO pattern (random scatter) in the residual plot

d. Use the residual plot to determine if the Regression line (Line of Best Fit) is a good fit for this data. Explain.

The linear regression model is not a good fit. (NOT appropriate) for the data because there is a pattern in the residual plot. A non-linear model would be a better fit.

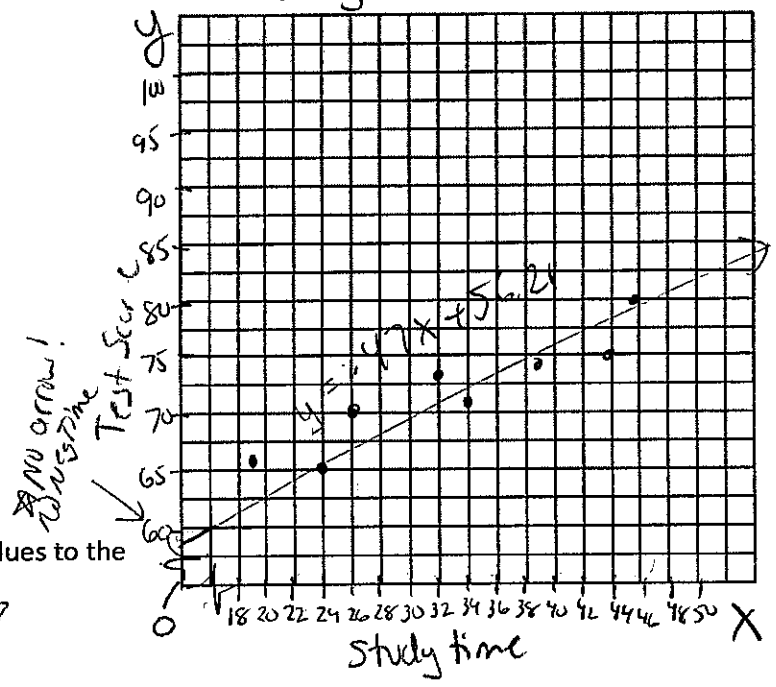
12) The table shows study times of 8 students and their related test scores.

Study time vs Test Score

a) Draw a scatter plot for the following data.

Study Time (min)	Test Score
19	66
24	65
26	70
32	73
34	71
39	74
44	75
46	80

L1 → STAT 1:Edit



b) Find the best trend line (line of best fit), rounding values to the nearest hundredth.

STAT →  
CALC  
4: LinReg

Linear Regression model  
 $y = .47x + 56.21$

c) Using the equation you found in part b, predict how many minutes a student has to study to receive a grade of 84. Round to the nearest whole number.

$$y = .47x + 56.21$$

$$84 = .47x + 56.21$$

$$\begin{array}{r} 84 \\ -56.21 \\ \hline 27.79 \end{array} = \frac{.47x}{.47} \quad x = 59.12765952 \dots$$

≈ 59 minutes

d) Using the equation you found in part b, predict the test score of a student who studied for 90 minutes. Round to the nearest whole number.

$$y = .47x + 56.21$$

$$y = .47(90) + 56.21$$

$$y = 42.3 + 56.21$$

$$y = 98.51$$

≈ 99

e) Were you interpolating or extrapolating data for part d? Explain.

extrapolating 90 is outside the plotted data (look at table to see)  
(falls outside the values of the table)

f) What is the value of the correlation coefficient to the nearest ten thousandth? Explain the meaning.

RTU turn on the "r"  
press MODE  
↓ until you see  
STAT 7:AB

$r = .9306$

STATS CALC 4: LinReg

The correlation coefficient shows a strong positive linear relationship between the students study time and their related test scores. B/C the correlation coefficient is close to 1.

13) Which set of data of temperatures has the largest dispersion as measured by its interquartile range?

Spread

$Q_3 - Q_1$

(1) 15, 17, 19, 21, 21, 22, 28

$22 - 17 = 5$

(2) 21, 23, 36, 37, 44, 48, 50

$48 - 23 = 25$

(3) 10, 19, 22, 23, 23, 29, 44

$29 - 19 = 10$

(4) 42, 47, 49, 50, 52, 59, 60

$59 - 47 = 12$

put #'s into L1 (STAT) (EDIT)  
 (STAT) → (CALC) (1) (↓)

14) Calculate the range for the data (15, 20, 22, 25, 30)

$30 - 15 = 15$

15) Mark off the appropriate box with a check or X.

Descriptive:  Can't be measured  
 #'s:  Can be measured

a) Determine which of the following would represent qualitative or quantitative data

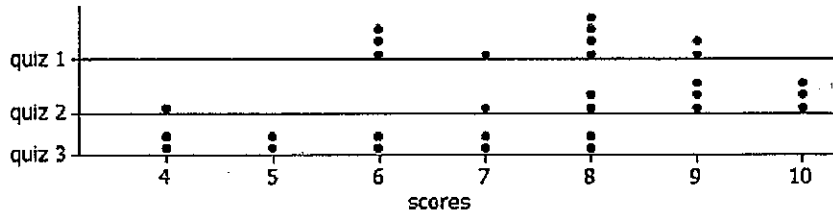
	Qualitative	Quantitative
SAT Scores		X
Nationality	X	
Color	X	
Cholesterol Level		X

b) Which group being surveyed is biased or unbiased when determining a school's favorite sport?

	Biased	Unbiased
Football Team	X	
Randomly selecting students from a lunch line		X
Every 3 <sup>rd</sup> girl that enters the building	X	
Surveying 5 random classes from each grade level		X

16) The scores of three quizzes are shown in the following data plot for a class of 10 students.

Each quiz has a maximum possible score of 10. Possible dot plots of the data are shown below.



a) Without performing any calculations, which quiz has the lowest standard deviation in the students' scores?

Justify your response.

Quiz #1 b/c it has the smallest spread of values.

b) If you were to calculate a measure of central tendency for Quiz 1, would you recommend using the mean or median? Explain your choice.

mean b/c it is not skewed (no outlier)

17) Which statement is true about the data set 4, 5, 6, 7, 8, 12?

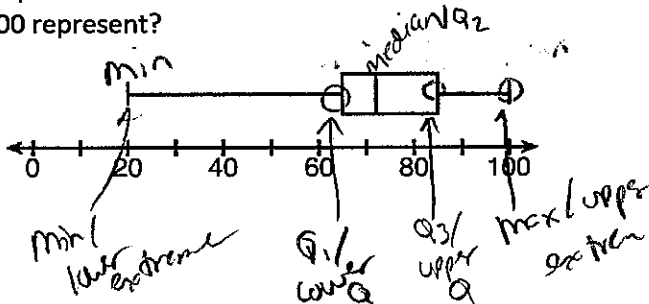
$$\frac{49}{7} = 7$$

- (1) mean = mode
- (2) mode = median
- (3) mean < median
- (4) mode > mean

mean = 7  
mode = 6  
median = 6

18) a. The box-and-whisker plot below represents the results of tests scores in a math class. What do the scores 65, 85, and 100 represent?

- (1)  $Q_1$ , median,  $Q_3$
- (2)  $Q_1$ ,  $Q_3$ , maximum
- (3) median,  $Q_1$ , maximum
- (4) minimum, median, maximum



b. What percent of the data do the test scores from 20 to 72 represent? Explain how you arrived at your answer.

50%. Two sections of the box plot & each section is 25%

19) Carrie has grades of 84, 65, and 76 on three social studies tests. What grade must she obtain on the next test to have an average of exactly 80 for the four tests? Show how you arrived at your answer algebraically.

let  $x =$  the 4<sup>th</sup> test score

$$\frac{84 + 65 + 76 + x}{4} = 80$$

$$225 + x = 320$$

$$x = 95$$

The fourth test score must be 95 to get an average of 80

$84 + 65 + 76 +$	$x$	$=$	$80$
$4$			
$320$		$=$	$80$
$4$			

20) Create a box plot (box and whiskers) for the following data: 85, 90, 100, 72, 88, 55, 73. What is the IQR?

STAT → CALC → 1-Var Stats ↓

55, 65, 72, 73, 88, 90, 100

Min 55  
Q1 65  
Med 73  
Q3 90  
Max 100

Data ← Don't forget title

The IQR is  $Q_3 - Q_1$   
 $90 - 65 = 25$

21) Which situation should be analyzed using bivariate data?

Comparison of 2 variables

1) Ms. Saleem keeps a list of the amount of time her daughter spends on her social studies homework.

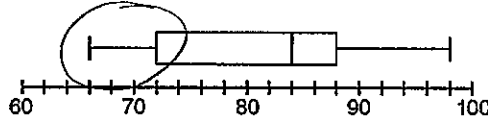
2) Mr. Benjamin tries to see if his students' shoe sizes are directly related to their heights.

3) Mr. DeStefan records his customers' best video game scores during the summer.

4) Mr. Chan keeps track of his daughter's algebra grades for the quarter.

Univariate = frequency

22) The box-and-whisker plot below represents the math test scores of 20 students.



What percentage of the test scores are less than 72?

1) 25

2) 50

3) 75

4) 100

Each section is 25%

23) Which data set describes a situation that could be classified as qualitative?

1) the elevations of the five highest mountains in the world

2) the ages of presidents at the time of their inauguration

3) the opinions of students regarding school lunches

4) the shoe sizes of players on the basketball team

descriptive  
(can't be measured)

24) The mean of five consecutive integers is 55. What are the integers?

$$\begin{aligned} X &= 1^{\text{st}} \text{ CI} \\ X+1 &= 2^{\text{nd}} \text{ CI} \\ X+2 &= 3^{\text{rd}} \text{ CI} \\ X+3 &= 4^{\text{th}} \text{ CI} \\ X+4 &= 5^{\text{th}} \text{ CI} \end{aligned}$$

$$\frac{X + X+1 + X+2 + X+3 + X+4}{5} = 55$$

$$\frac{5X + 10}{5} = 55$$

$$\begin{array}{r} 5X + 10 = 275 \\ -10 \quad -10 \\ \hline 5X = 265 \end{array}$$

$$\frac{5X}{5} = \frac{265}{5}$$

$$\begin{aligned} X &= 53 & X+2 &= 55 & X+4 &= 57 \\ X+1 &= 54 & X+3 &= 56 \end{aligned}$$

The 5 integers  
53, 54, 55, 56, 57

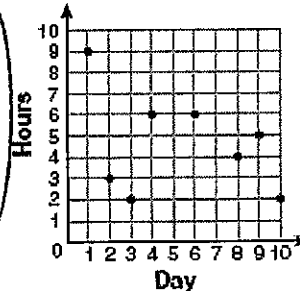
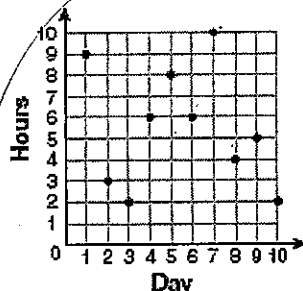
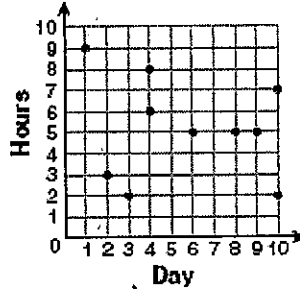
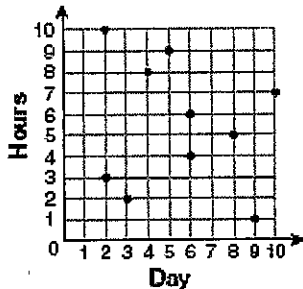
$$\frac{53 + 54 + 55 + 56 + 57}{5} = 55 \checkmark$$

25)

For 10 days, Romero kept a record of the number of hours he spent listening to music. The information is shown in the table below.

Day	1	2	3	4	5	6	7	8	9	10
Hours	9	3	2	6	8	6	10	4	5	2

Which scatter plot shows Romero's data graphically?





26) Twenty students were surveyed about the number of days they played outside in one week. The results of this survey are shown below.

{6, 5, 4, 5, 0, 7, 1, 5, 4, 4, 3, 2, 2, 3, 2, 4, 3, 4, 0, 7}

Use a Break if the Interval doesn't start at 0

(a) Complete the frequency table below for these data.

Interval	Tally	Frequency
0-1		3
2-3		6
4-5		8
6-7		+ 3

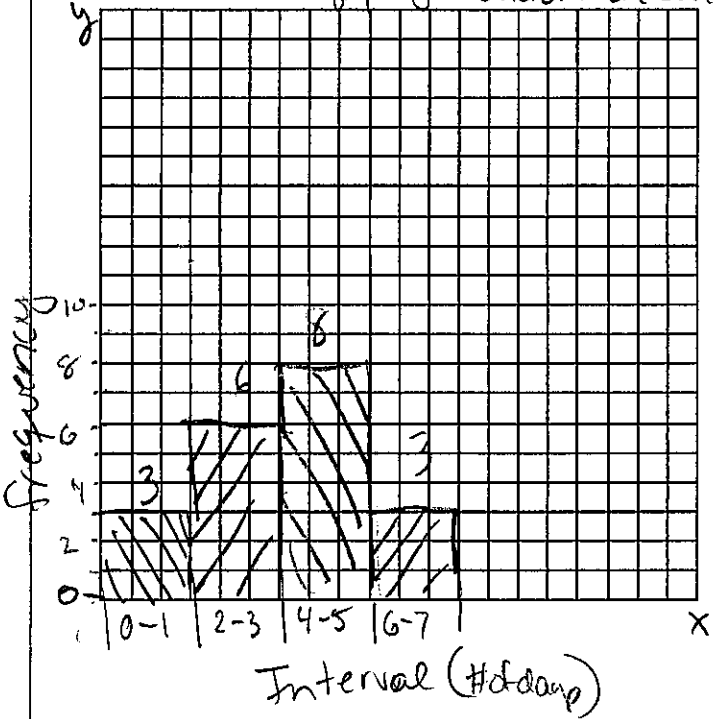
total freq = 20 ✓

(b) Complete the cumulative frequency table below using these data.

Interval	Cumulative Frequency
0-1	3
0-3	9
0-5	17
0-7	20 ✓

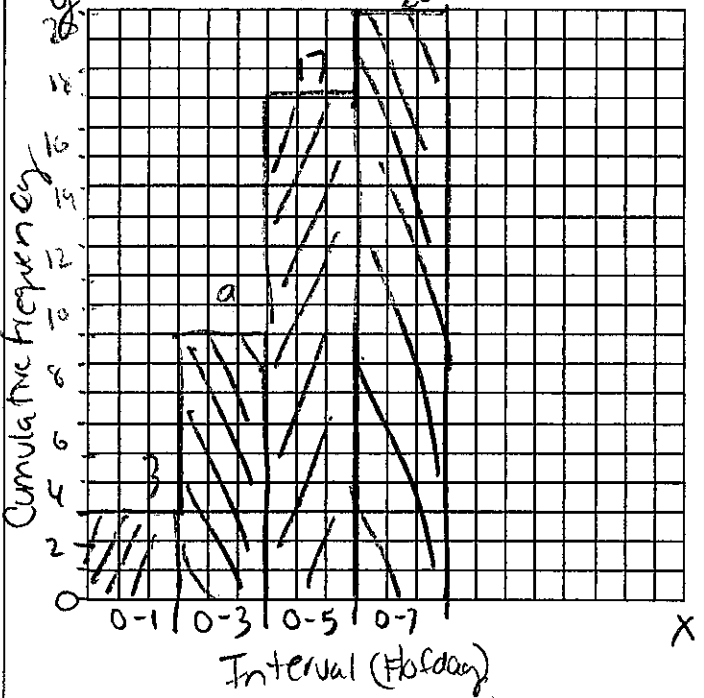
Create a frequency histogram

# of days played outside in one week



Create a cumulative frequency histogram

# of days played outside in one week

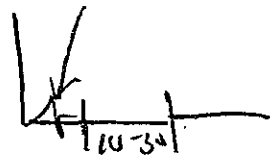


c) which interval contains the 40th percentile?

$$20 \times 40\% = 8$$

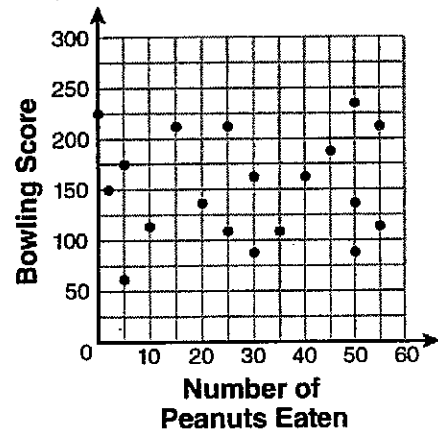
8th person in [2-3]

Use a Break if the Interval doesn't start at 0



27) The scatter plot below represents the relationship between the number of peanuts a student eats and the student's bowling score. Which conclusion about the scatter plot is valid?

- (1) There is almost no relationship between eating peanuts and bowling score.
- (2) Students who eat more peanuts have higher bowling scores.
- (3) Students who eat more peanuts have lower bowling scores.
- (4) No bowlers eat peanuts.



28) Jenna has bought a new hybrid car. Each week for a year, she has noted the fuel efficiency (in miles per gallon) of her car. The results for 7 weeks are shown below:

45 44 43 44 45 44 43 → put #s into L1 (STAT) 1:Edit

a) Determine if this data set is a population or a sample

Sample, only 7 weeks out of the total year

b) Find the standard deviation to the nearest hundredth.

Stat → CALC  
1:Var-Stats  
Sx = .82

c) Find the variance to the nearest hundredth

Vars 5:Statistics  
3:Sx  
(Sx)<sup>2</sup> = .67

29) The data below shows the comparison of the price of a particular product with the number of units sold.

If the least-squares line is  $y = 2x + 53$ , determine the residuals for the last two data sets and complete the table accordingly. Show all work below.

*LOBF*  
*line of best fit*  
*DO a Calc*

\* put  $y = 2x + 53$  into  $y$ ,  
 + then go to table

Price Per Unit (Dollars) $X$	Units Sold Actual $y_1$	Predicted Units Sold $y_2$	Actual - Predicted Residual $y_1 - y_2$
32	100	117	-17
34	144	121	23
36	130	125	5
38	118	129	-11

Get these #'s from the table

30) The table shows math scores from Smith's math class:

Interval	Frequency	Cumulative Frequency
61-70	2	61-70   2
71-80	3	61-80   5
81-90	8	61-90   13
91-100	+ 7	61-100   20 ✓

T.F 20

total frequency

\* you MUST start counting from the lowest interval

1st Q / lower Quartile = 25%  
 2nd Q / Median = 50%  
 3rd Q / upper Quartile = 75%

a) Complete the table. ✓

b) What is the modal interval?

Highest frequency interval  
81-90

c) What is the interval containing the median?

T.F. 50%  
 $20 \cdot 50\% = 10$   
 10th person in 81-90

d) What interval contains the upper quartile?

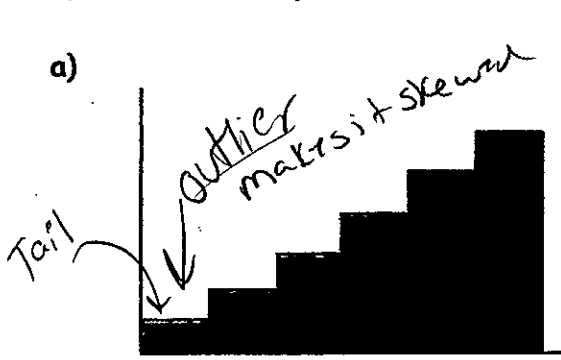
T.F. 75%  
 $20 \cdot 75\% = 15$   
 15th person in 91-100

use the regular frequency

e) What percent of the students scored 81-100?

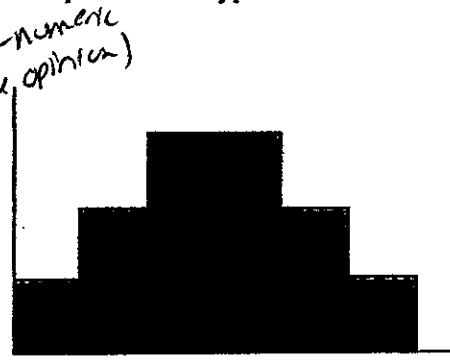
$8 + 7 = 15$   
 $\frac{15}{20} = .75$   
75%

31) For the following graphs, a) describe the distribution as symmetrical, skewed, uniform, or U-shaped. b) state whether you would use the mean, median, or neither to represent a typical value.



a) skewed left

b) median (B/c the outlier affects the answer)



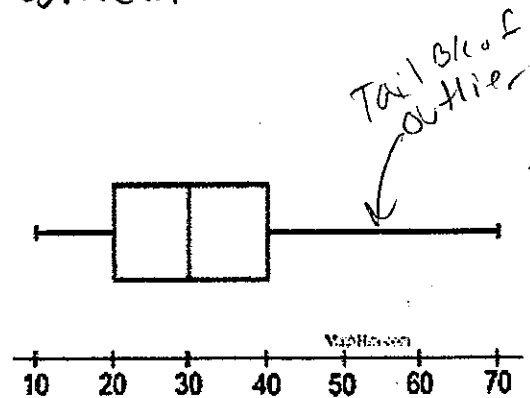
a) symmetrical (Bell Curve)

b) mean

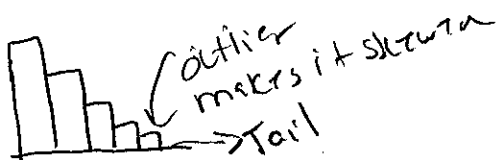
32) The box plot shown at the right is described as being a \_\_\_\_\_.

Choose:

- a) symmetric distribution
- b) skewed left distribution
- c) skewed right distribution



More Examples !!



a) skewed right

b) median



a) symmetrical

↓  
U-shaped

b) mean



a) uniform / symmetric

b) mean

33) A radio station did a survey to determine what kind of music to play by taking a sample of middle school, high school, and college students. They were asked which of the three different types of music they prefer on the radio: hip-hop, alternative, or classic rock. The results are summarized in the table below.

\* Must have  
"Total" row +  
column to answer

	Hip-Hop	Alternative	Classic Rock	Total
Middle School	28	18	4	50
High School	22	22	6	50
College	16	20	14	50
Total	66	60	24	150

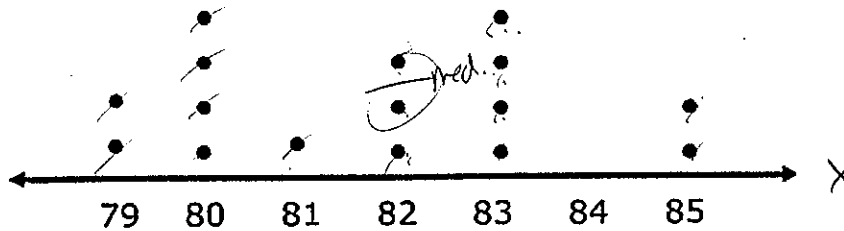
Cond. rel. freq.

What percent of college students prefer classic rock? → Num: what you want to know

Denom  
who is being surveyed

$$\frac{14}{50} = .28 = 28\%$$

34) The resting pulse rates were recorded for 16 boys in gym class before they exercised. The dot plot here shows the results.



a) What is the range (spread) of the pulse rates?

Max - Min  $85 - 79 = 6$

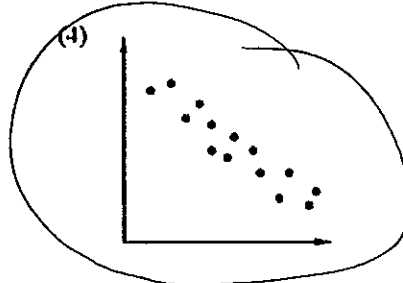
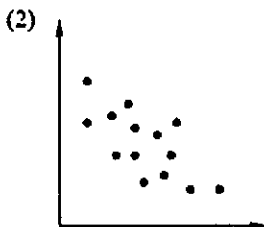
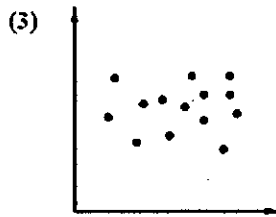
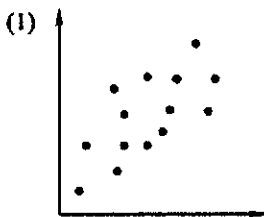
b) What is the mode of the pulse rates?

highest frequency 80 & 83

c) What is the median of the pulse rates?

middle #  $\frac{16}{2} = 8$  8th pos. [82] or put #'s in calc

35) Which of the following scatter plots would have a correlation coefficient closest to -1?



$r = -1$   
↑  
strong  
Neg.  
Close to Negative  
straight line.

STAT 1! Edit  
+ go to STAT  
→ CALC  
1: 1-Var Stats

