

Math2200 Exam 1, Feb 05, 2008

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

- 1) Variable 1: A person's gender
Variable 2: The verdict of a jury
Variable 3: The speed of a car in miles per hour
- A) Variable 1: Quantitative
Variable 2: Quantitative
Variable 3: Quantitative
- B) Variable 1: Categorical
Variable 2: Categorical
Variable 3: Quantitative
- C) Variable 1: Quantitative
Variable 2: Categorical
Variable 3: Quantitative
- D) Variable 1: Quantitative
Variable 2: Categorical
Variable 3: Categorical
- 1) _____

Provide an appropriate response.

- 2) A local park district is planning to build a recreation center. The park district conducted a poll to find out the types of physical activities the local population would be interested in. The poll was based on telephone responses from 1013 randomly selected adults. The table shows the percentages of people who expressed interest in various activities.
- 2) _____

Activity	Percent
Running/Walking	56
Weight Training	45
Biking	34
Aerobics	25
Swimming	16

- Is it reasonable to conclude that 59% expressed interest in either biking or aerobics?
- A) No, because the percentages in relative frequencies tables can never be added.
B) Yes, because the percentages can always be added in relative frequency tables.
C) No, because these categories overlap.
D) No, because the poll is not based on a large-enough number of individuals.
E) Yes, because these categories do not overlap.

- 3) A company held a blood pressure screening clinic for its employees. The results are summarized in the table below by age group and blood pressure level. 3) _____

		Age		
		Under 30	30-49	Over 50
Blood Pressure	Low	29	38	32
	Normal	46	86	89
	High	16	60	68

Find the marginal distribution of blood pressure level.

- A) 21% low, 48% normal, 31% high
 B) 32% low, 51% normal, 18% high
 C) 20% low, 40% normal, 41% high
 D) 27% low, 91% normal, 45% high
 E) 20% under 30, 40% between 30-49, 41% over 50
- 4) A company held a blood pressure screening clinic for its employees. The results are summarized in the table below by age group and blood pressure level. 4) _____

		Age		
		Under 30	30-49	Over 50
Blood Pressure	Low	26	40	38
	Normal	45	94	89
	High	23	53	73

Find the conditional distribution of blood pressure level for employees over 50.

- A) 22% low, 47% normal, 31% high
 B) 28% low, 48% normal, 24% high
 C) 8% low, 19% normal, 15% high
 D) 19% low, 45% normal, 37% high
 E) 21% low, 50% normal, 28% high
- 5) A survey of autos parked in student and staff lots at a large university classified the brands by country of origin, as seen in the table. 5) _____

		Driver	
		Student	Staff
Origin	American	107	91
	European	30	21
	Asian	63	57

What percent of the staff owned American cars?

- A) 46% B) 53.8% C) 53.5% D) 24.7% E) 116.7%

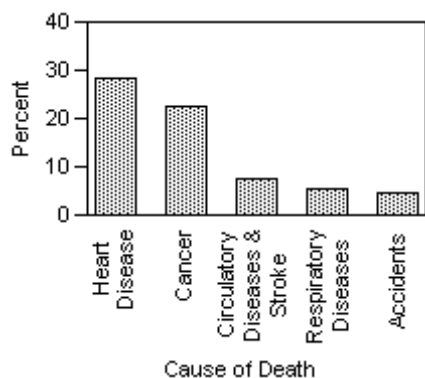
6) The Centers for Disease Control lists causes of death for individual states in 2002. The mortality data for one state is given.

6) _____

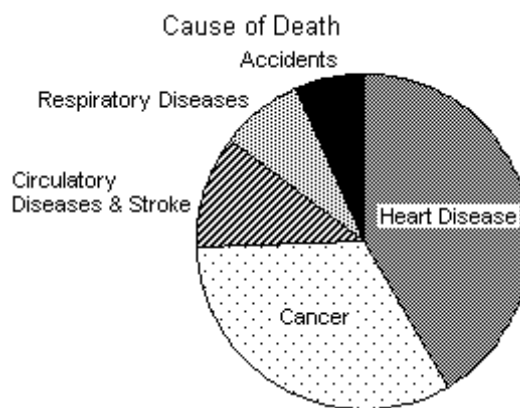
Cause of Death	Percent
Heart Disease	27.5
Cancer	21.9
Circulatory diseases and stroke	7.4
Respiratory diseases	5.9
Accidents	4.8

Which of the following displays is/are appropriate for these data? (More than one display may be appropriate.)

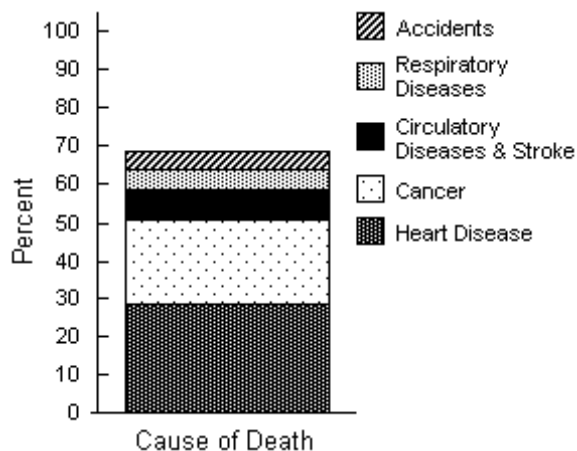
I



II



III



- A) I, II
- B) I, III
- C) I, II, III
- D) I
- E) None of these displays are appropriate.

- 7) An article in a magazine examined the graduate admissions process at a university for evidence of gender bias. The table below shows the number of applicants accepted to each of four graduate programs. 7) _____

Students Accepted

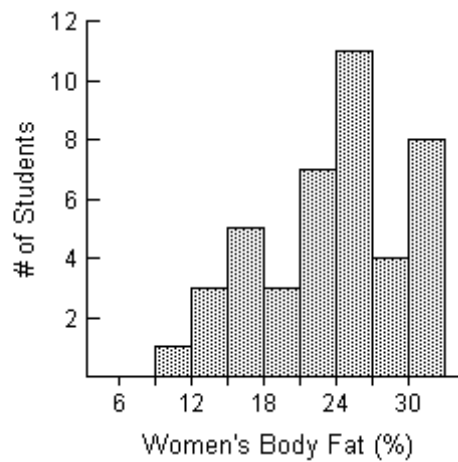
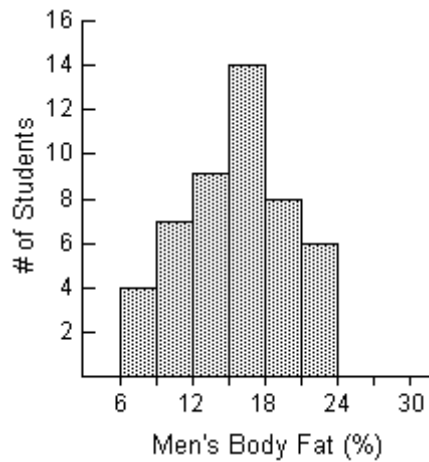
	Males accepted (of applicants)	Females accepted (of applicants)
Program 1	431 of 692	77 of 93
Program 2	268 of 414	27 of 36
Program 3	130 of 375	124 of 341
Program 4	14 of 233	17 of 205
Total	843 of 1714	245 of 675

After reading this article, a classmate claims that there is gender bias because the university clearly accepted more males than females. Do you agree with your classmate? Explain.

- A) Yes. Overall, 49.2% of males were accepted, compared to 36.3% of females were accepted. The overall percentages are always the most valid comparison.
- B) No. Overall, 49.2% of females were accepted, compared to 36.3% of males were accepted. Additionally, the comparison of the acceptance rate within each program shows that females had a higher acceptance rate for all four programs.
- C) No. Overall, 49.2% of males were accepted, compared to 36.3% of females were accepted. However, the comparison of the acceptance rate within each program shows that females had a higher acceptance rate for all four programs.
- D) Yes. Overall, 49.2% of males were accepted, compared to 36.3% of females were accepted. Additionally, the comparison of the acceptance rate within each program shows that males had a higher acceptance rate for all four programs.
- E) No. This is an example of Simpson's paradox and unfair averaging. Since the overall percentages and the individual percentages disagree on which gender was favored, the study is useless. Nothing can be determined from these data.

Solve the problem.

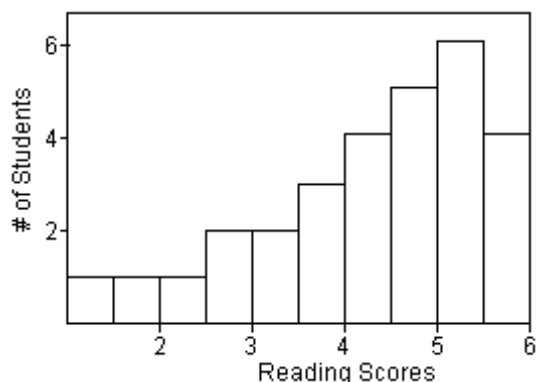
- 8) The histograms display the body fat percentages of 42 female students and 48 male students taking a college health course. For which of the variables depicted in the histograms would you be most satisfied to summarize the center with a mean? Explain. 8) _____



- A) The histogram of Women's Body Fat is most nearly symmetric, is not strongly skewed and shows no outliers. That makes it the best candidate of summarizing with a mean.
- B) The histogram of Men's Body Fat is most nearly symmetric, is not strongly skewed and shows no outliers. That makes it the best candidate of summarizing with a mean.
- C) The histogram of Men's Body Fat is skewed on the left. That makes it the best candidate of summarizing with a mean.
- D) The histogram of Women's Body Fat is skewed on the left. That makes it the best candidate of summarizing with a mean.
- E) The histogram of Women's Body Fat shows no outliers. That makes it the best candidate of summarizing with a mean.

9) Shown below are the histogram and summary statistics for the reading scores of 29 fifth graders.

9) _____



Count	Mean	Median	StdDev	Min	Max	Q1	Q3
29	4.2	4.6	1.3	1.3	5.9	3.5	5.4

Which measures of center and spread would you use for this distribution?

- A) Median and IQR, because the data is skewed to the left.
- B) Mean and standard deviation, because the data is skewed to the left.
- C) Mean and IQR, because the data is skewed to the left.
- D) Mean and standard deviation, because the data is symmetric.
- E) Median and standard deviation, because the data is skewed to the left.

Find the mean and median of the data.

10) Here are the number of hours that Bill has exercised each week since he started keeping records.

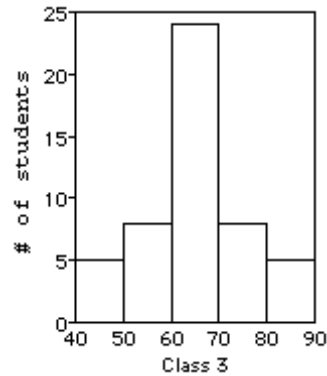
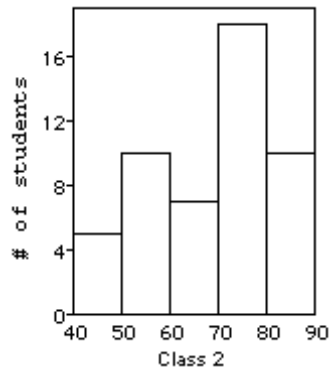
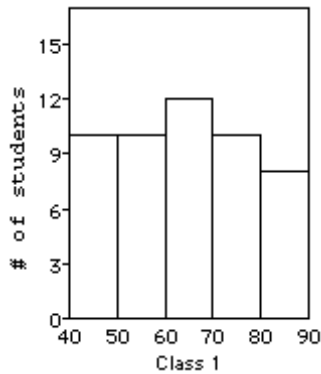
10) _____

8.5 6.5 7.1 8.7 6.9 8.5
 8.6 7.1 7.8 8.5 8.7 7.9
 8.6 6.5 6.5 8.8 6.9 8.6

Round your answer to the nearest tenth.

- A) mean = 7.4 hours, median = 7.8 hours
- B) mean = 9.3 hours, median = 8.5 hours
- C) mean = 8.3 hours, median = 8.7 hours
- D) mean = 8.0 hours, median = 8.6 hours
- E) mean = 7.8 hours, median = 8.2 hours

Three statistics classes (50 students each) took the same test. Shown below are histograms of the scores for the classes. Use the histograms to answer the question.



11) Which class had the highest mean score?

- A) Class 2
- B) Class 1 and class 3
- C) Class 1
- D) Class 3
- E) None, because the classes had the same mean.

11) _____

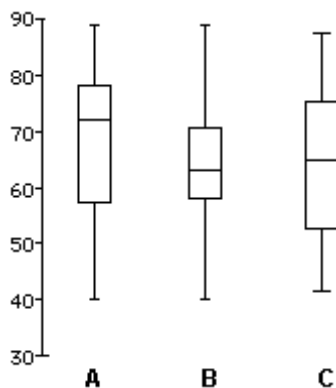
12) For class 2, compare the mean and the median.

- A) Median is lower than mean.
- B) Mean is equal to median.
- C) Median is higher than mean.
- D) Mean is higher than median.
- E) No comparison possible

12) _____

13) Match each class with the corresponding boxplot below.

13) _____



- A) Class 1 is B, Class 2 is A, Class 3 is C
- B) Class 1 is C, Class 2 is A, Class 3 is B
- C) Class 1 is A, Class 2 is B, Class 3 is C
- D) Class 1 is C, Class 2 is B, Class 3 is A
- E) Class 1 is B, Class 2 is C, Class 3 is A

Describe the distribution (shape, center, spread, unusual features).

14) The stem-and-leaf plot displays how populations of the 50 states and Washington, D.C., in millions of people, according to the 2000 census.

14) _____

State Populations (1 | 2 means 12 million)

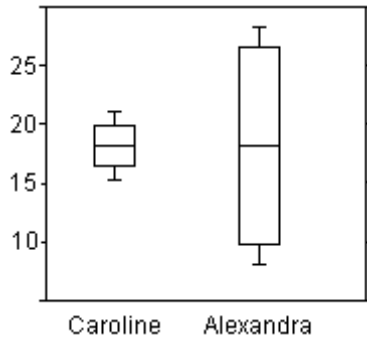
```
3 | 4
2 |
2 | 1
1 | 69
1 | 0122
0 | 5555666667888
0 | 111111111111112222233333344444
```

- A) The distribution of populations of the states and Washington, D. C., is unimodal and skewed to the right. The median population is 4 million. One state is an outlier, with a population of 34 million.
- B) The distribution of populations of the states and Washington, D. C., is unimodal and skewed to the right. The median population is 10 million. One state is an outlier, with a population of 34 million.
- C) The distribution of populations of the states and Washington, D. C., is bimodal and skewed to the left. The median population is 4 million. One state is an outlier, with a population of 34 million.
- D) The distribution of populations of the states and Washington, D. C., is unimodal and skewed to the left. The median population is 4 million. One state is an outlier, with a population of 34 million.
- E) The distribution of populations of the states and Washington, D. C., is unimodal and skewed to the left. The median population is 10 million. One state is an outlier, with a population of 34 million.

Solve the problem.

- 15) Here are boxplots of the points scored during the first 10 games of the basketball season for both Caroline and Alexandra. Summarize the similarities and differences in their performance so far.

15) _____



- A) Both girls have a median score of about 18 points per game. Alexandra is much more consistent, because her IQR is about 15 points, while Caroline's is over 3.
- B) Both girls have a median score of about 18 points per game. Caroline is much more consistent, because her IQR is about 4 points, while Alexandra's is over 15.
- C) The girls have a different average score per game. Caroline is much more consistent, because her IQR is about 4 points, while Alexandra's is over 15.
- D) Both girls have a median score of about 18 points per game. Caroline is much more consistent, because her IQR is about 6 points, while Alexandra's is over 20.
- E) The girls have a different average score per game, but the same median score of about 18 points per game. Their IQR are different, but this does not give anymore information on the girls' performance.

Use summary statistics to answer the question.

- 16) A local ice cream shop hand scoops each of its ice cream cones. The cones vary in weight from 4.3 ounces to 7.7 ounces with a mean of 6.45 ounces and a standard deviation of 1.2 ounces. The quartiles and median weights are 5.2, 8.7, and 7.2 ounces.

16) _____

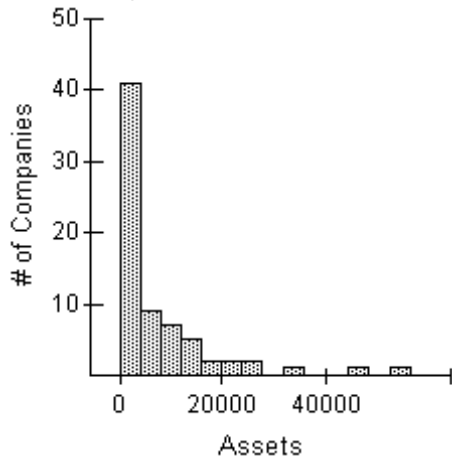
Is the distribution symmetric, skewed to the left, or skewed to the right? Explain.

- A) Skewed to the right, mean higher than median.
- B) Skewed to the left, mean higher than median.
- C) Skewed to the left, mean lower than median.
- D) Symmetric, mean lower than median.
- E) Skewed to the right, mean lower than median.

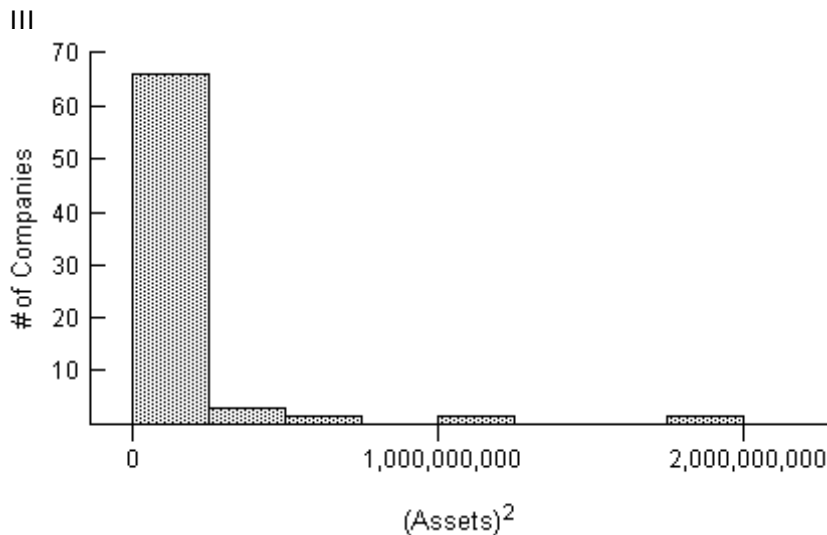
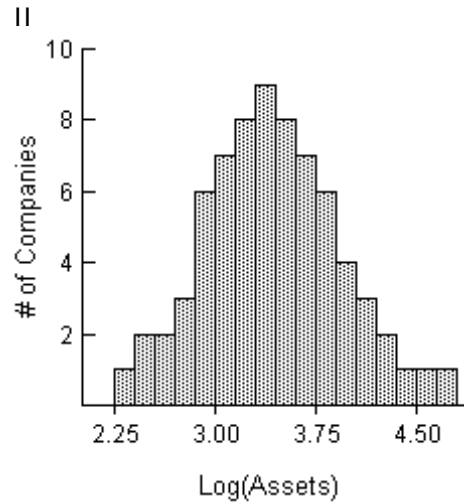
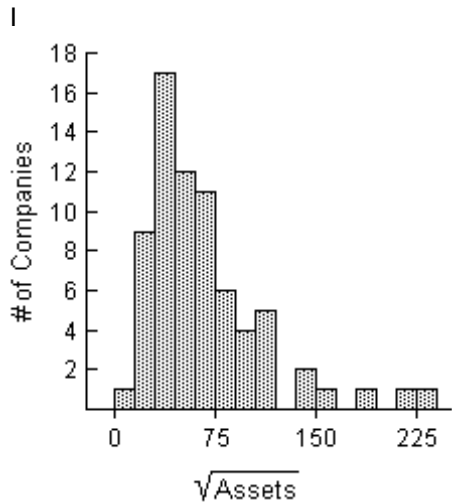
Provide an appropriate response.

17) Here is a histogram of the assets (in millions of dollars) of 71 companies.

17) _____



Which of the following is the most appropriate re-expression of these data? Explain.



- A) I and II are equally appropriate, because re-expression using logs or square roots yields the same results.
- B) II, because the distribution is nearly symmetric.
- C) I, because the distribution has a greater spread.
- D) III, because the distribution more closely resembles the original histogram.
- E) I, because the distribution is nearly symmetric.

Solve the problem.

18) A basketball coach kept stats for his team in free throw percentage and steals (among others). At the last game, Erin's free throw percentage was 79% and she had 4 steals. The team averaged 90% from the free throw line with a standard deviation of 6 and they averaged 7 steals with a standard deviation of 4. In which category did Erin do better compared with her team? Explain.

18) _____

A) Free throw percentage. 79% free throw average is $-\frac{11}{6}$ standard deviations from the mean

while 4 steals is $-\frac{3}{4}$ standard deviations from the mean.

B) Steals. 4 steals is $-\frac{11}{6}$ standard deviations from the mean while 79% free throw average is

$-\frac{3}{4}$ standard deviations from the mean.

C) Steals. 4 steals is $-\frac{3}{4}$ standard deviations from the mean while 79% free throw average is

$-\frac{11}{6}$ standard deviations from the mean.

D) Free throw percentage. 79% free throw average is $-\frac{3}{4}$ standard deviations from the mean

while 4 steals is $-\frac{11}{6}$ standard deviations from the mean.

E) One can't compare the two categories, they are too different.

19) The volumes of soda in quart soda bottles can be described by a Normal model with a mean of 32.3 oz and a standard deviation of 1.2 oz. What percentage of bottles can we expect to have a volume less than 32 oz?

19) _____

A) 47.15%

B) 40.13%

C) 38.21%

D) 9.87%

E) 59.87%

Solve the problem. Round to the nearest tenth.

20) Based on the Normal model for car speeds on an old town highway $N(77, 9.1)$, what is the cutoff value for the highest 15% of the speeds?

20) _____

A) about 11.6 mph

B) about 86.5 mph

C) about 65.5 mph

D) about 63.1 mph

E) about 67.5 mph

Solve the problem. Round to the nearest hundredth.

21) After increased patrol, cars on an old town highway travel at speeds averaging 54 mph. If 10% of vehicles travel above 70 mph, what approximate standard deviation could represent this model (assuming a Normal model is appropriate)?

21) _____

A) 35.94

B) -12.50

C) 7

D) 5.4

E) 12.50

Find the correlation.

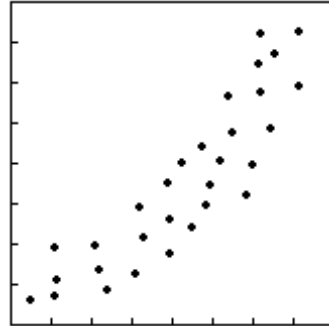
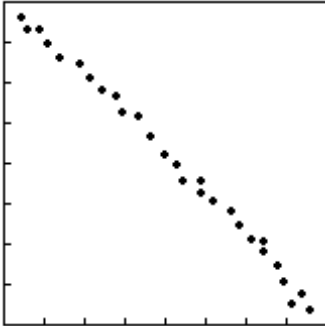
- 22) A scientist collects data to predict the wheat yield (in bushels per acre) based on rainfall (in inches). The results are recorded in the table below. 22) _____

Rainfall (in.)	Wheat Yield (bushels per acre)
11.9	61.8
8	26.4
11.6	51.2
18.4	79.1
9	41.5
10.9	42.9
15.1	71
13.5	54.6

- A) 0.941 B) 0.942 C) 0.932 D) 0.952 E) 0.943

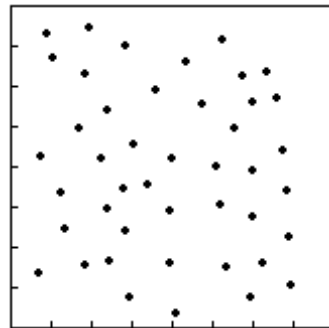
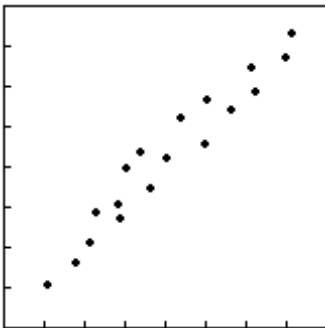
Several scatterplots are given with calculated correlations. Which is which?

- 23) 1) _____ 2) _____ 3) _____ 4) _____



3)

4)



a) -0.993, b) -0.047, c) 0.774, d) 0.937

- A) 1a, 2c, 3b, 4d
 B) 1c, 2d, 3b, 4a
 C) 1a, 2c, 3d, 4b
 D) 1b, 2d, 3c, 4d
 E) 1d, 2b, 3a, 4c

Solve the problem.

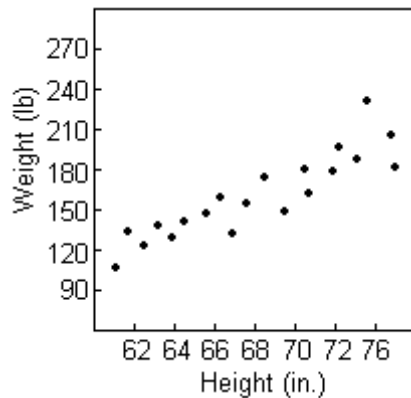
24) A science instructor assigns a group of students to investigate the relationship between the pH of the water of a river and its water's hardness (measured in grains). Some students wrote these conclusions: "there was a very strong correlation of $0.902 \text{ grains}^{-1}$ between pH of the water and water's hardness." Is the calculation of the correlation appropriate?

24) _____

- A) Yes: correlation is less than 1.
- B) No: there is little or no association.
- C) No: pH and hardness of water have different units.
- D) Yes: the pH and the hardness of the water are data collected from the same river.
- E) No: correlation has no units.

25) Data collected from students in Statistics classes included their heights (in inches) and weights (in pounds). For the students' heights and weights, the correlation is 0.653. Suppose the variable weight is recorded in kilograms rather than in pounds. What will be the correlation?

25) _____



- A) 0.653
- B) -0.653
- C) 1.438 in./kg
- D) 0.653 in./kg
- E) 0.653 kg/in.

Fill in the missing information.

26)	\bar{x}	s_x	\bar{y}	s_y	r	$\hat{y} = b_0 + b_1x$
	10	5	?	?	-0.8	$\hat{y} = 200 - 2x$

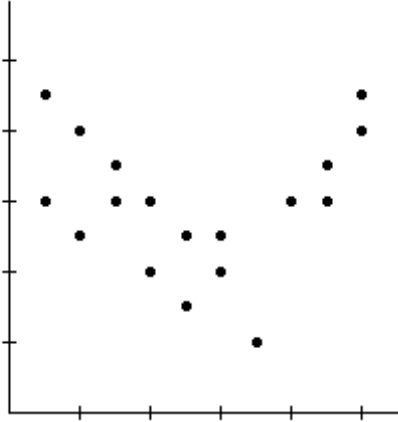
26) _____

- A) $\bar{y} = 190; s_y = 0.32$
- B) $\bar{y} = 180; s_y = 12.50$
- C) $\bar{y} = 210; s_y = 6$
- D) $\bar{y} = 20; s_y = 2.50$
- E) $\bar{y} = 220; s_y = 12.50$

Tell what the residual plot indicates about the appropriateness of the linear model that was fit to the data.

27)

27) _____



- A) Model is not appropriate. The relationship is nonlinear.
- B) Model may not be appropriate. The spread is changing.
- C) Model is appropriate.

Answer the question appropriately.

28) A golf ball is dropped from 15 different heights (in inches) and the height of the bounce is recorded (in inches.) The regression analysis gives the model $\hat{\text{bounce}} = 0.4 + 0.72 \text{ drop}$. A golf ball dropped from 64 inches bounced 1 inch less than expected. How high did it bounce?

28) _____

- A) 86.94 inches
- B) 47.48 inches
- C) 45.08 inches
- D) 66.12 inches
- E) 45.48 inches

Use the given data to find the equation of the regression line. Round to 3 significant digits, if necessary.

29) Managers rate employees according to job performance and attitude. The results for several randomly selected employees are given below.

29) _____

Attitude	59	63	65	69	58	77	76	69	70	64
Performance	72	67	78	82	75	87	92	83	87	78

- A) $\hat{\text{Performance}} = 92.3 - 0.669 \text{ Attitude}$
- B) $\hat{\text{Performance}} = 2.81 + 1.35 \text{ Attitude}$
- C) $\hat{\text{Performance}} = 11.7 + 1.02 \text{ Attitude}$
- D) $\hat{\text{Performance}} = 100.3 - 0.453 \text{ Attitude}$
- E) $\hat{\text{Performance}} = -47.3 + 2.02 \text{ Attitude}$

30) A biology student does a study to investigate the association between the amount of sunlight and the number of roses on a rosebush in one summer. (The R^2 value is 58%) He claims that the amount of sunlight determines 58% of the number of roses on a rosebush in one summer. Explain what is wrong with the interpretation. Assume calculations are done correctly.

30) _____

- A) There is nothing wrong with the interpretation.
- B) The amount of sunlight will increase the number of roses 58% of the time.
- C) The R^2 has to be greater than 90% to make a statement like this.
- D) The amount of sunlight accounts for 58% of the variation in the number of roses. It does not determine the number of roses.
- E) The amount of variation in sunlight changes 58% of the time. This tells us nothing about the number of roses.