EST 401 – Educational Statistics - 2013/2014

1. Which one of the following variables can be classified as unordered?
    a. Age of students
    b. School students attended
    c. Scores in quiz 1
    d. Weight of students
2. The Registrar of the University of Cape Coast wants a pictorial representation of the ages of students at the centre for continuing education. Which of the following graphs would be the most appropriate for him to use?
    a. Bar graph
    b. Histogram
    c. Line graph
    d. Ogive
3. Which one of the following variables would provide values in an interval scale?
    a. Gender of students in a Hall of residence.
    b. Number of students in the University of Cape Coast
    c. Scores in end-of-semester examination
    d. Types of cars University lecturers own
4. One weakness of the mean as a measure of location is that it
    a. Can be used for incomplete data
    b. Is affected by extreme scores
    c. Is sensitive to changes in scores
    d. Uses all the scores in a set of data
5. Pie charts can best be used to compare
    a. BECE results in 2012 in a school.
    b. Enrollment by gender in a school
    c. School attendance by teachers
    d. School fees paid by students
6. The upper quartile in the following distribution is   9   7   15   30   34   17   24   32
    a. 9
    b. 12
    c. 30
    d. 31
7. The following scores were available for 9 students in a statistics class.
    22 12 13 15 10 20 8 17 18
    The score for the 10th student was missing but it was known to be the highest score. What would be the median for the distribution?
    a. 15
    b. 16
    c. 17

    d. It cannot be determined

8. One weakness of line graphs as pictorial representation of data is that
   a. **Extreme values distort comparisons**
   b. They are difficult to construct
   c. They are effective with ratio scales
   d. Values can be read from the graph

9. The mode in a distribution was 40. One student whose score was missing had 40. The found score was added to the distribution. What would be the new mode?
   a. **40**
   b. 41
   c. 42
   d. More information is needed.

10. The mean score obtained by 50 students in a statistics quiz was 60. It was found later that a student who obtained 25 actually had 52. How would the discovery affect the mean?
    a. More information is needed.
    b. **New mean is greater than old mean.**
    c. Old mean is greater than new mean.
    d. There is no change in the old mean.

11. In an end-of-semester examination, a mean of 72 was obtained by a statistics class. The median score was 60. The performance of the class can be said to be
    a. Bimodal
    b. Normally distributed
    c. Skewed to the left
    d. **Skewed to the right**

12. Ten students obtained the following score in an examination.
    24   33   48   25   20   18   20   45   32   30
    Given that the scores are arranged in order above, find the value of

    $$\sum_{i=3}^{7} Xi$$

    a. 37
    b. 105
    c. **131**
    d. 188

13. The variance of a distribution is 36. What is the corresponding standard deviation of the distribution?
    a. 3
    b. **6**
    c. 36
    d. 1296

14. What is the range in the following distribution?

    11   18   36   14   -26   17   30   16   13   20

    a. 10
    b. 23
    c. 25
    d. 62

15. The first quartile (Q1) in a distribution of scores is 20. The third quartile (Q3) is 60. What is the value of the semi inter-quartile range?

    a. 10
    b. 20
    c. 30
    d. 40

16. Joseph's score in a statistics examination is 72. With this score he lies in the 70th percentile of his class. This information means that Joseph has done better than_____ of the students in his class.

    a. 2%
    b. 28%
    c. 30%
    d. 70%

17. The linear correlation coefficient between two variables is 0.8. this means that the coefficient of determination is

    a. 0.80
    b. 0.64
    c. 0.28
    d. 0.20

18. Adelaide obtained 65 in an examination. The mean for the class was 75, with a standard deviation of 5.0. what was Adelaide's Z-score?

    a. – 20
    b. -10
    c. 1.0
    d. 2.0

19. The phi coefficient is most appropriate as a measure of linear relationship between two variables when both variables are

    a. Continuous
    b. Nominal
    c. Ordinal
    d. true dichotomies

20. What area of the normal distribution is covered by $\bar{x} \pm 3s$?
   a. 0.4987
   b. 0.6826
   c. 0.9544
   d. 0.9974

SECTION B

1. You have just graduated from a College of Education. Explain to an untrained teacher, using appropriate examples, **five** reasons why it is important for teachers to study statistics in education.
   Reasons
   1. Using appropriate statistics to describe performance
   2. Understanding information from test scores
   3. Evaluating course grades
   4. Reading and understanding professional journals in education
   5. Carrying out research in education

2.
   i.   State three weaknesses of the mean as a measure of location
          Weaknesses
          1. It is influenced by extreme scores
          2. it is very sensitive to a change in the value of any score
          3. It cannot be computed if a score is missing and the sum of the scores or observations cannot be obtained.

   ii.   Explain, using examples, **three** properties of the mean.
          Properties
          1. The mean is influenced by every score or values that makes it up. If a score is changed, the values of the mean change.
          2. The mean is very sensitive to extreme scores which are called outliers.
          3. The mean is a function of the sum (or aggregate or total) of the scores. This implies that one cannot obtain the mean without knowing the sum of the scores. If one number is missing, the mean cannot be obtained.
          4. If the mean is subtracted from each individual score and the difference are summed, the result is 0
          5. If the same value is added to or subtracted from every number in a set of scores, the mean goes up or goes down by the value of the number
          6. If each score is multiplied or divided by the same value, the mean increases or decreases by the same value.

iii. In an end of semester examination, a mean of 60 and a standard deviation of 5 were obtained.

a. Obtain the coefficient of variation

$$CV = \frac{S}{x} \times 100$$

$$CV = \frac{5}{60} \times 100$$

$$CV = 8.3\%$$

b. What information do you derive from the value of the coefficient of variation concerning the nature of the class?

The class is **homogeneous** since the value of **CV is less than 33%**

c. State one use of the coefficient of variation in teaching and learning.

1. It is used to determine whether a group is homogeneous or heterogeneous. If the value of the CV is 33% or less, then the group is homogeneous, otherwise, it is heterogeneous.

2. It is used to compare variations within or between groups where there are different units of measurement.

3. it is used to compare variations within or between group where there are different means but with the same unit of measurement.

3.

i. What is a Z standard score?

It represents an individual raw score that has been transformed into a common standard based on a mean of 0 and a standard deviation of 1.

ii. State two differences between Z and T standard scores.

1. Z score has a mean of 0 while T score has a mean of 50
2. Z score has a standard deviation/variance of 1 while T score has a standard deviation of 10
3. Z score ranges from – 4 to +4 while T score ranges from 10 – 90
4. Z score could be positive or negative but T score is always positive.

iii. In an entrance examination a student obtained a raw score of 72. It is known that the student's Z score was 3.0. if the standard deviation for the group is 5, calculate the mean score for the group.

$$Z = \frac{X - \bar{x}}{S}$$

$$ZS = X - \bar{x}$$

$$\bar{x} = X - ZS$$

$\bar{x} = 72 - (3)(5)$
$\bar{x} = 72 - 15$
$\bar{x} = 57$

iv. Describe, with examples **three** ways by which the classroom teacher can improve teaching and learning by the use of Z scores.

Ways

1. It helps the teacher to know an individual's position in relation to the rest of the class
2. It enables the teacher to compare student's performances in different subjects to know the individual strengths and weaknesses
3. It helps the teacher to guide and counsel the student to choose the correct course for a future career and vocation

4. The table below shows the performance of 20 students in a course in two quizzes.

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Qiuz 1	11	17	12	15	8	15	16	10	17	12	17	15	12	14	13	15	20	20	12	9
Quiz 2	10	14	15	16	12	16	15	15	18	16	18	18	15	16	10	12	20	19	14	11

i. Obtain the Person product moment correlation coefficient

Student No.	Quiz 1(X)	Quiz 2 (Y)	X - $\bar{x}$	(X - $\bar{x})^2$	Y - $\bar{Y}$	(Y - $\bar{Y})^2$	(X - $\bar{x}$) (Y - $\bar{Y}$)
1	11	10	-3	9	-5	25	15
2	17	14	3	9	-1	1	-3
3	12	15	-2	4	0	0	0
4	15	16	1	1	1	1	1
5	8	12	-6	36	-3	9	18
6	15	16	1	1	1	1	1
7	16	15	2	4	0	0	0
8	10	15	-4	16	0	0	0
9	17	18	3	9	3	9	9
10	12	16	-2	4	1	1	-2
11	17	18	3	9	3	9	9
12	15	18	1	1	3	9	3
13	12	15	-2	4	0	0	0
14	14	16	0	0	1	1	0
15	13	10	-1	1	-5	25	5
16	15	12	1	1	-3	9	-3
17	20	20	6	36	5	25	30
18	20	19	6	36	4	16	24
19	12	14	-2	4	-1	1	2
20	9	11	-5	25	-4	16	20
Total	**280**	**300**		**210**		**158**	**129**

$\bar{x} = 14$ $\qquad\qquad\qquad\qquad$ $\bar{Y} = 15$

$r = \dfrac{\sum(X - \bar{x})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{x})^2 \ \sum(Y - \bar{Y})^2}}$

$r = \dfrac{129}{\sqrt{(210)(158)}}$

$r = \dfrac{129}{182.15} = 0.708$

    ii.   Interpret the correlation obtained

        There is a strong/high positive correlation

        A student with a high score in quiz 1 is more likely to perform well in quiz 2

    iii. Explain, using examples, two uses of correlation in education

        Uses

1. Selection for admission purposes
2. Placing students where they are likely to succeed in the future
3. Determination of reliability of standardized and classroom tests.
4. Aids in the provision of evidences for the validity of assessment instruments
5. Prediction of future performances
6. Research in education

s

SECTION B

1. You are a District Director of Education who took a course in Educational Statistics. Explain, using appropriate examples, five reasons why you encourage classroom teachers to study statistics in education.

    Reasons

1. Using appropriate statistics to describe performance
2. Understanding information from test scores
3. Evaluating course grades
4. Reading and understanding professional journals in education
5. Carrying out research in education

2.

   i.   State three strengths of the variance as a measure of dispersion

        Strengths

1. It uses every score in the distribution
2. It is used a lot for further statistical analysis
3. It is the most appropriate measure of variability for scores that are normally distributed

   ii.   Explain, using examples, three properties of the variance.

        Properties

1. The variance of the constant is zero
2. Variance is not resistant. It is affected by extreme scores.
3. Variance is independent of change of scale

iii. In an end-of-semester examination, a mean of 40 and a standard deviation of 4 were obtained.

a. Obtain the coefficient of variation

$CV = \frac{S}{x} \times 100$

$CV = \frac{4}{40} \times 100$

$CV = 10.0\%$

b. What information do you derive from the value of the coefficient of variation concerning the nature of the class?

The class is homogeneous since the value of CV is less than 33%

c. State one strength of the coefficient of variation as a measure of relative position.

Strengths

1. It is easy to compute
2. It is unitless and this makes it possible to compare variability for different distributions
3. It is based on every score in the distribution where the distribution is normal.
4. It is easy to interpret

3.

i. What is standard score?

A standard score indicates the number of standard deviation units an individual score is above or below the mean of each group.

It represents an individual score that has been transformed into a common standard using the mean and the standard deviation

ii. State two differences between Z and T standard scores.

1. Z score has a mean of 0 while T score has a mean of 50
2. Z score has a standard deviation / variance of 1 while T score has a standard deviation of 10
3. Z score ranges from -4 to +4 while T score ranges from 10 – 90
4. Z score could be positive or negative but T score is always positive.

iii. In an entrance examination a student obtained a raw score of 69. It is known that the mean for the group is 60 with a standard deviation of 3. Calculate the student's T score.

$T = 50 + 10Z$

$T = 50 + 10 \left(\frac{X - \bar{x}}{s}\right)$

$$T = 50 + 10\left(\frac{69-60}{3}\right)$$
$$T = 50 + 10\left(\frac{9}{3}\right)$$
$$T = 50 + 10(3)$$
$$T = 50 + 30$$
$$T = 80$$

iv. Describe with examples, three ways by which the classroom teacher can improve teaching and learning by use of T scores.

Ways
1. It helps the teacher to know an individual's position in relation to the rest of the class
2. It enables the teacher to compare student's performances in different subjects to know individual strengths and weaknesses.
3. It helps the teacher to guide and counsel the students to choose the correct course for a future career and vocation

4. The table below shows the performance of schools in the Jukwa district with respect to positions in the Inter-Schools Athletics Competition and the rank in terms of school size (i.e. number of pupils, the school with highest enrolment is 1st)

School	A	B	C	D	E	F	G	H	I	J
Position in Inter-Schools	10	8	4	1	3	9	2	6	5	7
Rank	10	5	4	1	2	8	7	3	6	9

i. Obtain the Spearman's correlation coefficient

School	R1	R2	R1 – R2 (d)	D2
A	10	10	0	0
B	8	5	3	9
C	4	4	0	0
D	1	1	0	0
E	3	2	1	1
F	9	8	1	1
G	2	7	-5	25
H	6	3	3	9
I	5	6	-1	1
J	7	9	-2	4
Total				**50**

$P= 1 - \dfrac{6\sum d^2}{N(N^2-1)}$

$P= 1 - \dfrac{6(50)}{10(10^2-1)}$

$P= 1 - \dfrac{300}{10(99)}$

$P= 1 - \dfrac{300}{990}$

$P= 1 - 0.30$

$P= 0.70$

ii.   Explain, using examples, three uses of coefficient obtained
There is a strong/high negative correlation
A school with a large pupil population is more likely to perform well in the Inter-Schools Athletic Competition

iii.   Explain, using examples, three uses of correlation in education
Uses
1. Selection for admission purposes
2. Placing students where they are likely to succeed in the future
3. Determination of reliability of standardized and classroom tests
4. Aids in the provision of evidences for the validity of assessment instruments
5. Prediction of future performances
6. Research in education

Educational Statistics Quiz II – 2013/2014

1. Why should classroom teachers study Statistics in Education?
   a. **It helps them to better describe students' performance**
   b. Statistics helps them to study behaviour of pupils
   c. Teachers need statistics to write good less notes
   d. Teachers use it for promotion and certification

2. Which one of the following variables can be classified as ordered?
   a. Books bought for a class
   b. **Scores in an examination**
   c. Senior High School attended
   d. Types of mobile phones

3. An example of a variable that can best be classified as discrete is the
   a. **Goals scored in a football match**
   b. Height of pupils in Primary Six
   c. Speed in a 1500 metre race
   d. Time to complete a quiz

4. The Director of the Centre for Continuing Education, UCC wants a pictorial representation of the number of students from each region for the class of 2012 registered at the Centre. Which one of the following graphs would be most appropriate for him to use?
   a. Bar graph
   b. **Histogram**
   c. Line graph
   d. Ogive

5. Which one of the following variables would provide values in a ratio scale?
   a. Colour of books
   b. Days of the week
   c. **Number of cars in a college**
   d. Position in class test

6. One strength of the mean as a measure of location is that it
   a. Can be used for incomplete data

b. Is not affected by extreme scores
c. Is sensitive to changes in scores
d. Uses all the scores in a set of data

7. Line graphs can best be used to compare
   a. Ages of headmasters in a school district
   b. Attendance at one teachers' workshop
   c. BECE results from 2005 – 2010 in a school
   d. School fees paid by students in a school.

8. What is the first quartile in the following distribution? 9 25 32 28 40 11 30 3
   a. 9
   b. 10
   c. 28
   d. 36

9. In a class quiz, a mean of 60 was obtained with a median of 36. How would the performance of the class be described?
   a. Above average
   b. Negatively skewed
   c. Normal
   d. Positively skewed

10. The mode in a distribution was 48. One students whose score was missing had 48. The found score was added to the distribution. What would be the new mode?
    a. 45
    b. 46
    c. 48
    d. More information is needed

11. One strength of pie charts as pictorial representation of data is that
    a. Individual parts of the whole can be compared
    b. Predictions are possible for information not available
    c. They are useful for all scales of measurement
    d. Values from each component part are difficult to read.

Study the frequency distribution below and answer questions 12 – 15

Distribution of examination scores for Statistics students

Classes	Frequency
96 – 100	8
91 – 95	12
86 – 90	17
81 – 85	23
76 – 80	25
71 – 75	15
66 – 65	14
61 – 65	6
Total	120

12. What is the value of the mean score?
   a. 90.7
   b. 80.7
   c. 79.7
   d. 70.7

13. The cumulative relative frequency for the class, 71 – 75 is
   a. 0.125
   b. 0.29
   c. 0.17
   d. 0.24

14. The class which would probably contain the median in the above distribution is
   a. 71 – 75
   b. 76 – 80
   c. 81 – 85
   d. 86 – 90

15. The overall shape of the above distribution is that it is
   a. Bi-modal
   b. Negatively skewed
   c. Normal
   d. Positively skewed

16. Ten students obtained the following scores in an examination :

    48   32   25   45   20   33   30   18   24   20

    Given that the scores are arranged in the order above, find the value of

$$\sum_{i=2}^{8} x_i$$

   a. 28
   b. 215
   c. 227
   d. 251

17. The mean score obtained by 10 students in a statistics quiz out of a total of 25 was 20. It was observed that a student who obtained 5 actually had 20. How would the discovery affect the mean?
   a. More information is needed.
   b. New mean is greater than old mean
   c. Old mean is greater than the new mean
   d. There is no change in the old mean.

18. Educational research relies on Educational Statistics mainly for
   a. Direction for literature review
   b. Ideas for problem identification
   c. Tools for data analysis
   d. Tools for writing proposals

19. Which one of the following statements does not describe the construction of a frequency distribution table in education?
   a. Classes of equal intervals are used
   b. Number of classes is limited to 5 – 20
   c. Open-ended classes are avoided
   d. There can be a class with zero frequency

20. One limitation of the median as a measure of location is that it
   a. Can be used when data is incomplete
   b. Depends largely on extreme scores
   c. Is appropriate for skewed distributions
   d. Uses few values in a distribution.

Educational Statistics

1. Explain, using appropriate examples, five reasons why classroom teachers should study statistics in education.

   Reasons
   1. Using appropriate statistics to describe performance
   2. Understanding information from test scores
   3. Evaluating course grades
   4. Reading and understanding professional journals in education
   5. Carrying out research in education

2. The following score were obtained in a class quiz

14	42	19	20	42	21	48	22	26	47	27	42
28	30	45	31	42	31	35	47	34	22	35	42
36	38	50	38	40	47	40	41	42	20	42	45
40	46	47	38								

   i. Obtain a grouped equal interval frequency distribution table for the scores beginning with the interval 46 – 50.

Class	Class Mark	Tally	Frequency
46 – 50	48	⫫ //	7
41 – 45	43	⫫ ⫫	10
36 – 40	38	⫫ //	7
31 – 35	33	⫫	5
26 – 30	28	////	4
21 – 25	23	///	3
16 – 20	18	///	3
11 – 15	13	/	1
Total			40

ii.   Using the frequency distribution table, draw a frequency polygon to represent the scores.
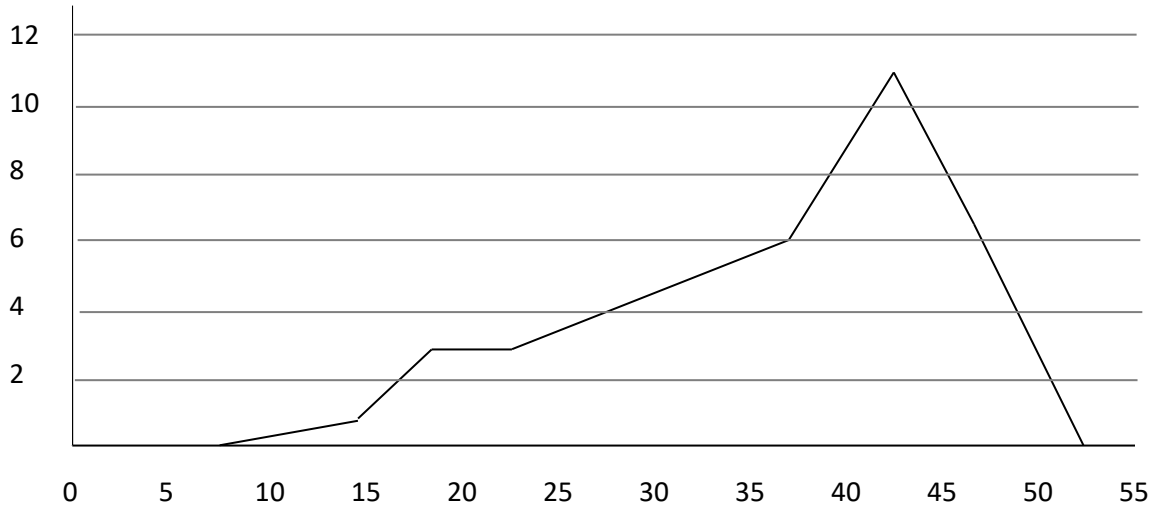
iii.   What information do you derive from the polygon concerning the performance of the students in the quiz?

The polygon is **negatively skewed** (skewed to the left). **Performance of the students is high.**

3.

i.   State three strengths of the standard deviation as a measure of dispersion.

Strengths
1. It uses every score in the distribution
2. It is used a lot for further statistical analysis.
3. It is the most appropriate measure of variability for scores that are normally distributed.
4. It is easy to interpret

ii.   Explain, using examples, three properties of the standard deviation.
Properties
1. The standard deviation of a set of scores that is constant is zero
2. Standard deviations are not resistant
3. Standard deviations are independent of change of origin.
4. Standard deviations are not independent of change of scale.

iii.   In an end of semester examination, a mean of 60 and a standard deviation of 8 were obtained

a. Obtain the coefficient of variation

$CV = \frac{S}{\bar{x}} \times 100$

$CV = \frac{8}{60} \times 100$

$CV = \frac{80}{6}$

$CV = 13.33\%$

b. What information do you derive from the value of the coefficient of variation concerning the nature of the class?

The class is **homogenous** since the value of **CV is less than 33%.**

c. State one weakness of the coefficient of variation as a measure of relative position

Weaknesses
1. It is affected by extreme values
2. It cannot be used when the mean is negative, zero or near zero
3. It is sensitive to a change in the value of any score in the distribution
4. It cannot be computed for a variable that is normally distributed if missing data is reported.

4. The frequency distribution below summarizes the performance of students in an end-of-semester examination

Scores	Frequency
91 – 100	8
81 – 90	12
71 – 80	25
61 – 70	30
51 – 60	17
41 – 50	7
31 – 40	1

i.   Obtain the mean score

Scores	Class Mark (x)	Frequency (f)	fx
91 – 100	95.5	8	764.0
81 – 90	85.5	12	1026.0
71 – 80	75.5	25	1887.5
61 – 70	65.5	30	1965.0
51 – 60	55.5	17	943.5
41 – 50	45.5	7	318.5
31 – 40	35.5	1	35.5
Total			6940

$$\bar{x} = \frac{\Sigma fx}{\Sigma f}$$

$$\bar{x} = \frac{6940}{100} = 69.4$$

Coding Method

Scores	Class Mark (x)	Frequency (f)	Code (d)	fd
91 – 100	95.5	8	3	24
81 – 90	85.5	12	2	24
71 – 80	75.5	25	1	25
61 – 70	65.5	30	0	0
51 – 60	55.5	17	-1	-17
41 – 50	45.5	7	-2	-14
31 – 40	35.5	1	-3	-3
Total				39

$$\bar{x} = AM + \left(\frac{\Sigma fd}{\Sigma f}\right)$$

$$\bar{x} = 65.5 + \left(\frac{39}{100}\right) 10$$

$$\bar{x} = 65.5 + 3.9$$

$$\bar{x} = 69.4$$

ii.   Obtain the median score

Scores	Class Mark (x)	Frequency (f)	*Cum. Freq*
91 – 100	95.5	8	100
81 – 90	85.5	12	92
71 – 80	75.5	25	80
61 – 70	65.5	30	55
51 – 60	55.5	17	25
41 – 50	45.5	7	8
31 – 40	35.5	1	1
Total			

$$\text{Mdn} = L1 + \left(\frac{\frac{N}{2} - Cf}{fmdn} \right) i$$

$$\text{Mdn} = 60.5 + \left(\frac{50 - 25}{30} \right) 10$$

$$\text{Mdn} = 60.5 + \left(\frac{25}{30} \right) 10$$

$$\text{Mdn} = 60.5 + 8.3$$

$$\text{Mdn} = 68.8$$

iii.   What information do you derive from the mean and median scores concerning the overall performance of the students in the test?

The values of the mean and median are close so the overall performance in normal

5.
i. State four features of the normal distribution

Features
1. It is bell-shaped
2. It is uni-modal
3. It is symmetrical about the mean.
4. It is asymptotic
5. The total area under the curve is 1
6. The mode, mean and median are equal

ii. In an entrance examination taken by 1200 students, the results were normally distributed with mean 56 and standard deviation of 8.

a. What is the probability that a student selected at random from the class obtains a score less than 48?

$P(x < 48) = P(Z < \frac{48-56}{8})$
$P(x < 48) = P(Z < \text{-}1)$

$\text{-}1 \quad 0$
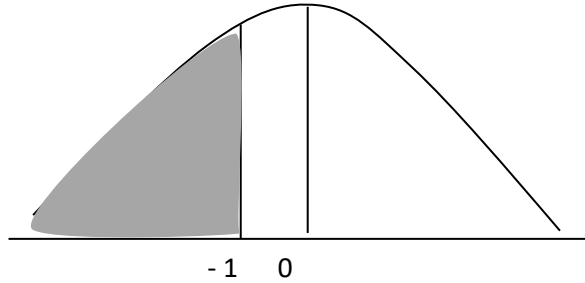
$P(x < 48) = 0.5000 - 0.3413$
$P(x < 48) = 0.1587$

b. What percentage of students obtained scores more than 72?
$P(x > 72) = P(Z > \frac{72-56}{8})$
$P(x > 72) = P(Z > \frac{16}{8})$
$P(x > 72) = P(Z > 2)$
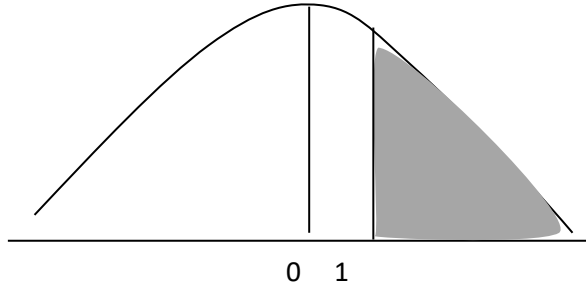

$0 \quad 2$
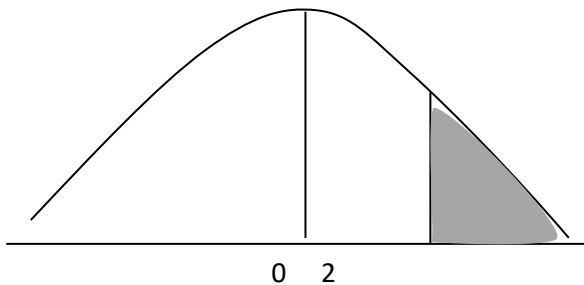
$P(X > 72) = 0.500 - 0.4772$
$P(X > 72) = 0.0228$
$P(X > 72) = 2.28\%$

c. If the pass mark is 64, approximately how many students passed?

$P(X > 64) = P(Z > \frac{64-56}{8})$

$P(X > 64) = P(Z > \frac{8}{8})$

$P(x > 68) = P(Z > 1)$

0   1

$P(x > 64) = 0.5000 - 0.3413$

$P(x > 64) = 0.1587 \times 1200$

$P(x > 64) = 190.44$

Number of students passed is approximately 191

6. The table below shows the performance of schools in the Pensa district with respect to positions in the BECE and the rank in terms of schools size (i.e. number of pupils)

School	A	B	C	D	E	F	G	H	I	J
Position in BECE	4	10	3	9	2	6	1	7	8	5
Rank in school size	5	1	6	2	10	3	9	7	3	4

i.   Obtain the Spearman's Correlation coefficient

School	R1	R2	$R_1 - R_2$ (d)	d^2
A	4	5	-1	1
B	10	1	9	81
C	3	6	-3	9
D	9	2	7	49
E	2	10	-8	64
F	6	3	3	9
G	1	9	-8	64
H	7	7	0	0
I	8	3	5	25
J	5	4	1	1
Total				303

$$\rho = 1 - \frac{6 \sum d^2}{N(N^2-1)}$$

$$\rho = 1 - \frac{6(303)}{10(10^2-1)}$$

$$\rho = 1 - \frac{1818}{10(99)}$$

$$\rho = 1 - 1.836$$

$$\rho = -0.84$$

ii.   Interpret the correlation coefficient obtained

There is a **strong/high negative correlation**. A school with small pupil population is more likely to perform well in the BECE

iii.   Explain, using examples, three uses of correlation in education.
1. Selection for admission purposes
2. Placing students where they are likely to succeed in the future
3. Determination of reliability of standardized and classroom tests
4. Aids in the provision of evidences for the validity of assessment instruments
5. Prediction of future performances
6. Research in education

QUIZ TWO

1. The standard deviation of a distribution is 8. What is the corresponding variance of the distribution?
    A. 80.0
    B. 64.0
    C. 4.2
    D. 2.8
2. What is the range in the following distribution?
        -4 18 22 13 6 12 20 24 15 8
    A. 28
    B. 26
    C. 20
    D. 18
3. A percentile range is a measure of the
    A. Variation of a distribution.
    B. Variability of a distribution.
    C. **Relative position in a distribution.**
    D. Location of a distribution.
4. The first quartile (Q1) in a distribution of a score is 15. The third quartile (Q3) is 45. What is the Value of the quartile deviation?
    A. 30
    B. 25
    C. 15
    D. 10
5. T-scores are used in educational statistics as measures of
    A. Variability
    B. **Relative position.**
    C. Dispersion.
    D. Central tendency.
6. Which one of the following situation is a use of correlation in education?
    A. Promoting students from one class to other
    B. **Predicting the future performance of a student**
    C. Determining the future performance of a student
    D. Awarding proficiency certificates of grades.
7. The upper quartile in the following distribution is
    A. 42
    B. 32
    C. 19
    D. 16

8. George's score in a Statistics examination is 56. With this score he lies in the 90th percentile of this class. The information means that George has **done worse** than_____ of the students in his class
   A. 90%
   B. 56%
   C. 44%
   D. **10%**

9. The linear correlation coefficient between two variables is 0.64. This is means that the coefficient of determination is
   A. 0.89
   B. 0.80
   C. 0.64
   D. **0.41**

10. Alice obtained 80 in an examination. The means for the class was 65, with a standard deviation of 5.0. What was Alice's Z-score?
    A. 3.5
    B. **3.0**
    C. - 2.5
    D.- 3.0

11. When a distribution is normal's the <u>most appropriate</u> measure of variability is the
    A. Third quartile
    B. **Standard deviation.**
    C. Quartile deviation
    D. First quartile

12. In a statistics class, a mean of 30 and a standard deviation of 9 were obtained in a quiz. The instructor later added 5 to each score. What is the new standard deviation of the scores?
    A. 81
    B. 45
    C. 9
    D. **5**

13. The phi coefficient <u>is most appropriate</u> as a measure of relationship between two variables are
    A. **True dichotomies**.
    B. Ordinal
    C. Nominal.
    D. Continuous

14. Fine the variance for the following set of scores.

    14 22 18 25 15 8 20 26 25 17

    A. 298

    B. 29.8

    C. 17.3

    D. 5.5

15. What area of the normal distribution is covered by $X \pm 2s$?

    A. **0.9544**

    B. 0.6826

    C. 0.4772

    D. 0.3413

16. A teacher wishes to find the relationship between the Quiz 2 score and the age of the students in his class. What is the most appropriate measure of relationship to use?

    A. **Person's product moment correlation coefficient**

    B. Phi coefficient.

    C. Point-biserial correlation coefficient

    D. Spearman's rank correlation coefficient

17. Given that a distribution had a chi-square $(\chi 2)$ value of 100 for a group of 120 students, what is the value of phi (φ) coefficient

    A. 0.6

    B. 0.7

    C. 0.8

    D. **0.9**

18. Gadzella, Baloglu and Stephen (2002) did a study in the USA and found that Education Psychology scores predicated Grade Point Average (GPA) better than WGCTA. Which variable is the criterion?

    A. WGCTA

    B. USA

    C. GPA

    D. Educational Psychology

Use information from the scatter plot to answer questions **19 and 20.**

High school GPA and College Entrance Exam Scores
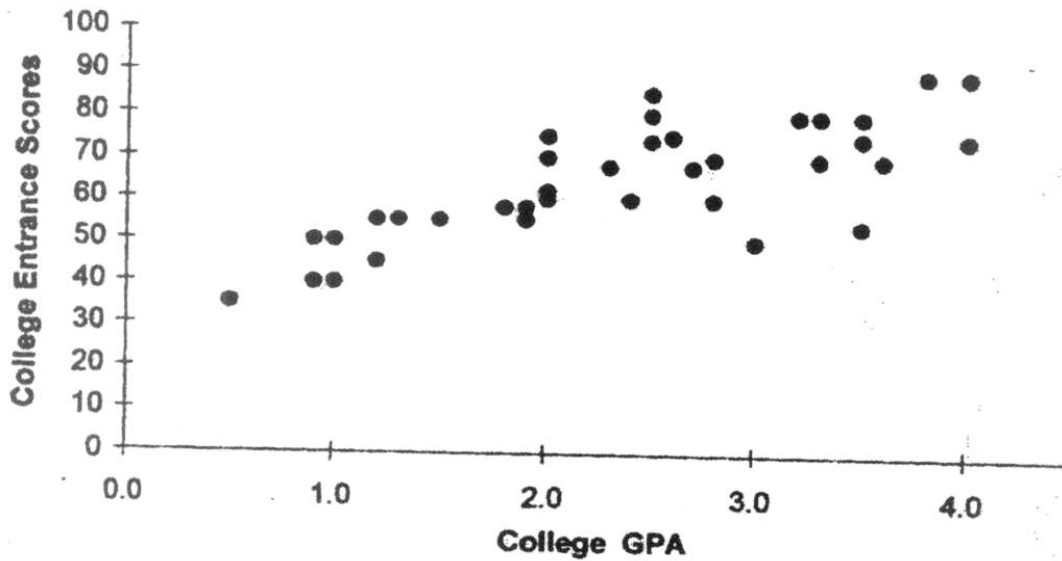
19. The scatter plots shows that students with low college entrance examination scores are more likely to have

    A. Zero College GPA
    B. Perfect College GPA
    C. **Low College GPA**
    D. High College GPA

20. The correlation coiefficient for the relationship between the college entrance examination and college GPA is approximately.
    A. -0.8
    B. -0.3
    C. **0.3**
    D. 0.8