UNIVERSITY OF ESWATINI

FINAL EXAMINATION PAPER 2018/2019

TITLE OF PAPER:	DESCRIPTIVE STATISTICS
COURSE CODE:	STA 131/ STA 132 / IDE 132
TIME ALLOCATED:	2 (TWO) HOURS
REQUIREMENTS:	CALCULATOR
INSTRUCTION:	ANSWER ALL PARTS OF SECTION A AND ANY 1 (ONE) QUESTION OF YOUR CHOICE FROM SECTION B. SECTION A IS **COMPULSORY AND THE ANSWER SHEET IS ATTACHED AT THE END OF THE QUESTION PAPER.** ALL QUESTIONS CARRY THE MARKS AS INDICATED WITHIN THE PARENTHESIS.

THIS PAPER IS NOT TO BE OPENED UNTIL PERMISSION HAS BEEN GRANTED BY THE INVIGILATOR.

SECTION A : Compulsory Questions *(Multiple Choice Questions)*

1. The number of floors of each building in a particular city is recorded. Make a stem-and-leaf plot of the data with the *Key*: 1 | 1 *means* 11

Number of Floors																			
24	38	46	16	28	31	12	27	32	49	33	12	38	34	48	22	36	29	47	41
49	30	21	17	40	13	32	15	31	21	13	16	43	33	30	25	28	29	13	11

a. **Number of Floors**

Stems	Leaves
1	1 2 3 5 6 7
2	1 2 4 5 7 8 9
3	0 1 2 3 4 6 8
4	0 1 3 6 7 8 9

c. **Number of Floors**

Stems	Leaves
1	11 12 12 13 13 13 15 16 16 17
2	21 21 22 24 25 27 28 28 29 29
3	30 30 31 31 32 32 33 33 34 36 38
4	40 41 43 46 47 48 49 49

b. **Number of Floors**

Stems	Leaves
1	1 2 2 3 3 3 5 6 6 7
2	1 1 2 4 5 7 8 8 9 9
3	0 0 1 1 2 2 3 3 4 6 8 8
4	0 1 3 6 7 8 9 9

d. **Number of Floors**

Stems	Leaves
1	11 12 13 15 16 17
2	21 22 24 25 27 28 29
3	30 31 32 33 34 36 38
4	30 31 33 36 37 38 39

2. The table shows the scores that 20 students received on a test. Use the data to make a frequency table showing the students scores.

Students Test Scores				
75	82	83	90	77
93	85	88	72	74
81	88	76	79	83
86	94	78	82	80

a.

Range	70 – 79	80 – 89	90 – 99
Frequency	7	10	3

b.

Range	70 – 79	80 – 89	90 – 99
Frequency	8	10	4

c.

Range	70 – 79	80 – 89	90 – 99
Frequency	7	11	2

d.

Range	70 – 79	80 – 89	90 – 99
Frequency	7	11	3

3. The relative frequency of a class is computed by
   a. dividing the midpoint of the class by the sample size
   b. dividing the frequency of the class by the midpoint
   c. dividing the sample size by the frequency of the class
   d. dividing the frequency of the class by the sample size

4. A researcher is gathering data from four geographical areas designated: South = 1; North = 2; East = 3; West = 4. The designated geographical regions represent
   a. qualitative data
   b. quantitative data
   c. label data
   d. either quantitative or qualitative data

5. The total number of data items with a value less than the upper limit for the class is given by the
   a. frequency distribution
   b. relative frequency distribution
   c. cumulative frequency distribution
   d. cumulative relative frequency distribution

Exhibit 2-1

The numbers of hours worked (per week) by 400 statistics students are shown below.

Number of hours	**Frequency**
0 - 9	20
10 - 19	80
20 - 29	200
30 - 39	100

6. Refer to Exhibit 2-1. The number of students working 19 hours or less
   a. is 80
   b. is 100
   c. is 180
   d. is 300

7. Refer to Exhibit 2-1. The percentage of students working 19 hours or less is
   a. 20%
   b. 25%
   c. 75%
   d. 80%

8. Refer to Exhibit 2-1. The cumulative percent frequency for the class of 30 - 39 is
   a. 100%

b. 75%
c. 50%
d. 25%

9. Refer to Exhibit 2-1. If a cumulative frequency distribution is developed for the above data, the last class will have a cumulative frequency of
   a. 100
   b. 1
   c. 30 - 39
   d. 400

10. Refer to Exhibit 2-1. The number of students who work 19 hours or less is
   a. 80
   b. 100
   c. 200
   d. 400

11. In constructing a frequency distribution, as the number of classes are decreased, the class width
   a. decreases
   b. remains unchanged
   c. increases
   d. can increase or decrease depending on the data values

12. In a cumulative frequency distribution, the last class will always have a cumulative frequency equal to
   a. one
   b. 100%
   c. the total number of elements in the data set

13. In a cumulative percent frequency distribution, the last class will have a cumulative percent frequency equal to
   a. one
   b. 100
   c. the total number of elements in the data set

Exhibit 2-2

A survey of 800 college seniors resulted in the following cross tabulation regarding their undergraduate major and whether or not they plan to go to graduate school.

Undergraduate Major Graduate School	Business	Engineering	Others	Total
Yes	70	84	126	280
No	182	208	130	520
Total	252	292	256	800

14. Refer to Exhibit 2-2. What percentage of the students does not plan to go to graduate school?
   a. 280
   b. 520
   c. 65
   d. 32

15. Refer to Exhibit 2-2. Of those students who are majoring in business, what percentage plans to go to graduate school?
    a. 27.78
    b. 8.75
    c. 70
    d. 72.22

16. The 75th percentile is referred to as the
    a. first quartile
    b. second quartile
    c. third quartile
    d. fourth quartile
    e. None of the above answers is correct.

17. The difference between the largest and the smallest data values is the
    a. variance
    b. interquartile range
    c. range
    d. coefficient of variation
    e. None of the above answers is correct.

18. The first quartile
    a. contains at least one third of the data elements
    b. is the same as the 25th percentile
    c. is the same as the 50th percentile
    d. is the same as the 75th percentile
    e. None of the above answers is correct.

19. Which of the following is not a measure of central location?
    a. mean
    b. median
    c. variance
    d. mode
    e. None of the above answers is correct.

20. The weights (in pounds) of a sample of 36 individuals were recorded and the following statistics were calculated.
    mean = 160 range = 60 mode = 165 variance = 324 median = 170
    The coefficient of variation equals
    a. 0.1125%
    b. 11.25%
    c. 203.12%
    d. 0.20312%
    e. None of the above answers is correct.

21. The standard deviation of a sample of 100 observations equals 64. The variance of the sample equals
    a. 8
    b. 10
    c. 6400
    d. 4,096

e. None of the above answers is correct.

Exhibit 3-3
A researcher has collected the following sample data. The mean of the sample is 5.

3 5 12 3 2

22. Refer to Exhibit 3-3. The variance is
a. 80
b. 4.062
c. 13.2
d. 16.5
e. None of the above answers is correct.

23. Refer to Exhibit 3-3. The standard deviation is
a. 8.944
b. 4.062
c. 13.2
d. 16.5
e. None of the above answers is correct.

24. Refer to Exhibit 3-3. The coefficient of variation is
a. 72.66%
b. 81.24%
c. 264%
d. 330%
e. None of the above answers is correct.

25. Refer to Exhibit 3-3. The range is
a. 1
b. 2
c. 10
d. 12
e. None of the above answers is correct.

26. Refer to Exhibit 3-3. The interquartile range is
a. 1
b. 2
c. 10
d. 12
e. None of the above answers is correct.

27. Which of the following symbols represents the standard deviation of the population?
a. σ^2
b. σ
c. μ
d. $\bar{x}$
e. N

28. Which of the following symbols represents the mean of the population?
a. σ^2
b. σ

c. μ
d. $\bar{x}$
e. N

29. Which of the following symbols represents the mean of the sample?
a. σ^2
b. σ
c. μ
d. $\bar{x}$
e. N

30. Which of the following symbols represents the size of the sample
a. σ^2
b. σ
c. N
d. $\bar{x}$
e. n

31. If two events (both with probability greater than 0) are mutually exclusive, then:
a. They also must be independent.
b. They also could be independent.
c. They cannot be independent.

32. If two events (both with probability greater than 0) are mutually exclusive, then:
a. They also must be complements.
b. They also could be complements.
c. They cannot be complements.

33. Suppose that the probability of event A is 0.2 and the probability of event B is 0.4. Also, suppose that the two events are independent. Then P(A|B) is:
a. P(A) = 0.2
b. P(A)/P(B) = 0.2/0.4 = ½
c. P(A) × P(B) = (0.2)(0.4) = 0.08
d. None of the above.

34. Two dice are thrown simultaneously. What is the probability of getting two numbers whose product is even?
a. ½
b. ¾
c. 3/8
d. 5/16

35. In a lottery, there are 10 prizes and 25 blanks. A lottery is drawn at random. What is the probability of getting a prize?
a. 1/10
b. 2/5
c. 2/7
d. 5/7

36. Two dice are tossed. The probability that the total score is a prime number is:
a. 1/6
b. 5/12

c. ½
d. 7/9

37. Two cards are drawn together from a pack of 52 cards. The probability that one is a spade and one is a heart, is:
    a. 3/20
    b. 29/34
    c. 47/100
    d. 13/102

38. One card is drawn at random from a pack of 52 cards. What is the probability that the card drawn is a face card (Jack, Queen and King only)
    a. 1/13
    b. 3/13
    c. ¼
    d. 9/52

39. A numerical measure of linear association between two variables is the
    a. variance
    b. covariance
    c. standard deviation
    d. coefficient of variation
    e. None of the above answers is correct.

40. Positive values of covariance indicate
    a. a positive variance of the x values
    b. a positive variance of the y values
    c. the standard deviation is positive
    d. positive relation between the independent and the dependent variables
    e. None of the above answers is correct.

41. The coefficient of correlation ranges between
    a. 0 and 1
    b. -1 and +1
    c. minus infinity and plus infinity
    d. 1 and 100
    e. None of the above answers is correct.

This scenario applies to Questions 42 and 43:
A study was done to compare the lung capacity of coal miners to the lung capacity of farm workers. The researcher studied 200 workers of each type. Other factors that might affect lung capacity are smoking habits and exercise habits. The smoking habits of the two worker types are similar, but the coal miners generally exercise less than the farm workers.

42. Which of the following is the explanatory variable in this study?
    a. Exercise
    b. Lung capacity
    c. Smoking or not
    d. Occupation

43. Which of the following is a confounding variable in this study?
    a. Exercise

b. Lung capacity
c. Smoking or not
d. Occupation

44. The correlation coefficient is used to determine:
    a. A specific value of the y-variable given a specific value of the x-variable
    b. A specific value of the x-variable given a specific value of the y-variable
    c. The strength of the relationship between the x and y variables
    d. None of these

45. If there is a very strong correlation between two variables then the correlation coefficient must be
    a. any value larger than 1
    b. much smaller than 0, if the correlation is negative
    c. much larger than 0, regardless of whether the correlation is negative or positive
    d. None of these alternatives is correct.

46. The relationship between number of beers consumed (x) and blood alcohol content (y) was studied in 16 male college students by using least squares regression. The following regression equation was obtained from this study: $Y = -0.0127 + 0.0180x$. The above equation implies that:
    a. each beer consumed increases blood alcohol by 1.27%
    b. on average it takes 1.8 beers to increase blood alcohol content by 1%
    c. each beer consumed increases blood alcohol by an average of amount of 1.8%
    d. each beer consumed increases blood alcohol by exactly 0.018

47. If the correlation coefficient is 0.8, the percentage of variation in the response variable explained by the variation in the explanatory variable is
    a. 0.80%
    b. 80%
    c. 0.64%
    d. 64%

Use the following story to answer questions 48-52

A variety of summary statistics were collected for a small sample (10) of bivariate data, where the dependent variable was y and an independent variable was x.
$\sum X = 90$, $\sum Y = 170$, $n = 10$, $SSE = 505.98$, $Sxy = 466$, $Sxx = 234$, $Syy = 1434$

48. Compute the sample correlation coefficient:
    a. 0.8045
    b. -0.8045
    c. 0
    d. 1

49. The least squares estimate of β_1 equals
    a. 0.923
    b. 1.991
    c. -1.991
    d. -0.923

50. The least squares estimate of β_0 equals
    a. 0.923

b. 1.991
c. -1.991
d. -0.923

51. The coefficient of determination equals
a. 0.6471
b . -0.6471
c. 0
d. 1

52. The point estimate of y when x = 0.55 is
a. 0.17205
b. 2.018
c. 1.0905
d. -2.018
e. -0.

53. The strength (degree) of the correlation between a set of independent variables X and a dependent variable Y is measured by
a. Coefficient of Correlation
b. Coefficient of Determination
c. Standard error of estimate
d. All of the above

54. An instructor gives the same y vs x data as given below to four students.

x	1	10	20	30	40
y	1	100	400	600	1200

They each come up with four different answers for the straight line regression model. Only one is correct. The correct model is
a. $y = 60x - 1200$
b. $y = 30x - 200$
c. $y = -139.43 + 29.684x$
d. $y = 1 + 22.782x$

55. An overall upward or downward pattern in an annual time series would be contained in which component of the times series
a. trend
b. cyclical
c. irregular
d. seasonal

56. When using exponentially weighted moving average for the purposes of forecasting rather than smoothing, the smoothed value for the period t becomes for the period
a. t-1
b. t
c. t+1
d. none of the above

57. The following table contains the number of complaints received in a department store for the first 6 months of last year.

Month	Complaints
Jan	36
Feb	45
Mar	81
Apr	90
May	108
Jun	144

If a 3-term moving average is used to smooth this series, what would be the second calculated term?

a. 36
b. 40.5
c. 54
d. 72

58. A composite price index where the prices of the items in the composite are weighted by their relative importance is known as the

a. none of the below
b. price relative
c. CPI
d. weighted aggregate price index

59. A weighted aggregate price index where the weight for each item is its current-period quantity is called the

a. Aggregate index
b. Consumer Price Index
c. Laspeyres Index
d. Paasche Index

60. An index that is designed to measure changes in quantities over time is known as the

a. Quantity index
b. Time index
c. None of the above
d. Paasche index

SECTION B : Choose Any One Question *(Either Question One or Question Two)*

QUESTION ONE [3+4+5+6+1+1]

A statistics lecturer conducted a study to investigate the relationship between performance of his students on exams and their anxiety. Ten students from his class were selected for the experiment. Just prior to taking the final exam, the 10 students were given an anxiety questionnaire. Here are the final exam and anxiety scores for the 10 students:

Anxiety	28	41	35	39	31	42	50	46	45	37
Final Exam	82	58	63	89	92	64	55	70	51	72

i. Which factor is the dependent variable and the independent variable and why?
ii. Construct a scatter plot of the paired scores and describe the relationship shown in the graph.
iii. Assuming the relationship is linear, compute the coefficient of correlation (r)
iv. Construct a regression equation for predicting the dependent variable.
v. Predict the final exam mark of a student whose anxiety score is 52%.
vi. Predict the anxiety score of a student given that they got 78% in the Final exam.

QUESTION TWO [13+7]

The table below gives quarterly data of food consumer price indices for Swaziland 2013-2017:

	Quarter			
Year	1	2	3	4
2013	98.5	98.9	101.4	107.3
2014	106.5	109.9	110.2	112.4
2015	113.5	117.6	119.0	124.2
2016	128.0	129.4	130.7	134.3
2017	137.3	140.5	142.2	143.8

i. Construct the four-quarter centred moving average for these data and determine the percentages of the moving average for the quarters.
ii. Determine the seasonal indexes for the quarters and de-seasonalize the original time series.

END OF EXAMINATION

UNIVERSITY OF SWAZILAND

FINAL EXAMINATION ANSWER SHEET

STUDENT ID.....

PROGRAM.....YEAR.....

INSTRUCTIONS

1. Cross out all your correct answers. For example 1. 

2. In the event that more than one answer is crossed out that is deemed automatically wrong.

1. A B C D E	17. A B C D E
2. A B C D E	18. A B C D E
3. A B C D E	19. A B C D E
4. A B C D E	20. A B C D E
5. A B C D E	21. A B C D E
6. A B C D E	22. A B C D E
7. A B C D E	23. A B C D E
8. A B C D E	24. A B C D E
9. A B C D E	25. A B C D E
10. A B C D E	26. A B C D E
11. A B C D E	27. A B C D E
12. A B C D E	28. A B C D E
13. A B C D E	29. A B C D E
14. A B C D E	30. A B C D E
15. A B C D E	31. A B C D E
16. A B C D E	32. A B C D E

33. (A) (B) (C) (D) (E)	47. (A) (B) (C) (D) (E)
34. (A) (B) (C) (D) (E)	48. (A) (B) (C) (D) (E)
35. (A) (B) (C) (D) (E)	49. (A) (B) (C) (D) (E)
36. (A) (B) (C) (D) (E)	50. (A) (B) (C) (D) (E)
37. (A) (B) (C) (D) (E)	51. (A) (B) (C) (D) (E)
38. (A) (B) (C) (D) (E)	52. (A) (B) (C) (D) (E)
39. (A) (B) (C) (D) (E)	53. (A) (B) (C) (D) (E)
40. (A) (B) (C) (D) (E)	54. (A) (B) (C) (D) (E)
41. (A) (B) (C) (D) (E)	55. (A) (B) (C) (D) (E)
42. (A) (B) (C) (D) (E)	56. (A) (B) (C) (D) (E)
43. (A) (B) (C) (D) (E)	57. (A) (B) (C) (D) (E)
44. (A) (B) (C) (D) (E)	58. (A) (B) (C) (D) (E)
45. (A) (B) (C) (D) (E)	59. (A) (B) (C) (D) (E)
46. (A) (B) (C) (D) (E)	60. (A) (B) (C) (D) (E)