UNIVERSITY OF SWAZILAND

FINAL EXAMINATION PAPER 2017/2018

TITLE OF PAPER:	DESCRIPTIVE STATISTICS
COURSE CODE:	STA 131/ STA 132 IDE
TIME ALLOCATED:	2 (TWO) HOURS
REQUIREMENTS:	GRAPH PAPER AND CALCULATOR
INSTRUCTION:	ANSWER ANY THREE (3) QUESTIONS. THE QUESTIONS CARRY THE MARKS AS INDICATED WITHIN THE PARENTHESIS

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

QUESTION ONE [5+8+4+3]

The marks of 40 students in a mathematics test are as follows:

49 73 58 40 61 45 59 61 51 43 79 55 77 52 54 45 59 53 55 38 34 72
46 68 42 65 65 64 29 48 39 67 28 56 66 48 42 56 51 53

i. Draw a stem-and-leaf diagram for the above data.
ii. Fill up the frequency distribution table for the marks of the 40 students below and draw a histogram.

Marks	Class boundary	Frequency
20-29		
30-39		
40-49		
50-59		
60-69		
70-79		

iii. Find the median for the
   a. raw data and
   b. grouped data using the frequency distribution table.
iv. Justify why the values in (iii.) are different.

QUESTION TWO [2+3+2+2+2+2+2+3+2]

a) A custom inspector examined the luggage of 20 persons who were entering Swaziland and recorded the time it took to inspect them. The following summary quantities (in minutes) were obtained:

$$\sum X = 49.3 \qquad \sum X^2 = 144.43$$

Calculate

i. the sample mean
ii. standard deviation of the inspection time

b) For further investigations, the station manager asked for the raw data (inspection times) of the 20 visitors and that data is as follows: 3.0, 2.0, 3.1, 2.0, 0.5, 5.0, 2.0, 1.1, 2.0, 1.9, 2.9, 1.8, 2.8, 2.8, 3.2, 2.7, 1.6, 1.6, 2.4 and 4.9 minutes.

Find the

i. mode
ii. Q_1
iii. Q_2
iv. Q_3
v. Interquartile range
vi. Represent your findings (ii, iii, iv) on a box and whisker diagram
vii. Comment on the skewness of the inspection times.

QUESTION THREE [5+2+3+4+3+3]

The table below shows data on the number of visitors to Swaziland in a month, x (1000s), and the amount of money they spent, y (E millions), for each of the eight months.

Number of visitors x (1000s)	**2450**	**2480**	**2540**	**2420**	**2350**	**2290**	**2400**	**2460**
Amount of money spent y (E millions)	1370	1350	1400	1330	1270	1210	1330	1350

i. Calculate and interpret r and r^2.
ii. Give a reason to support fitting a regression model of the form $y=a+bx$ to this data.
iii. Find the value of b and interpret it.
iv. Determine the equation of the regression line of y on x.
v. Use your answer to part iv. to estimate the amount of money spent when the number of visitors to Swaziland in a month was 2 500 000.
vi. Comment on the reliability of your estimate in part v. Give a reason for your answer.

QUESTION FOUR **[1+2+2+2+2+3+8]**

SECTION A

Let A and B be events in a sample space S, with P(A) = 0.40, P(B) = 0.25 and $P(A \cap B) = 0.10$. Find each of the following.

i. $P(\overline{A})$
ii. $P(A \cup B)$
iii. $P(A \cap \overline{B})$
iv. $P(\overline{A} \cap B)$
v. P(B|A)
vi. Are A and B independent?

SECTION B

Select the appropriate word from below that correctly completes each of the following statements

i. A list of the entire population from which a sample is to be drawn is called the _____
ii. A sample that selects every nth item from a list is called a _____ sample.
iii. _____ sampling attempts to adequately represent differing groups and population.
iv. _____ sampling is usefully employed when the population is scattered over a wide geographical area.

stratified, sampling frame, cluster, systematic

QUESTION FIVE [13+7]

A major amusement park has the following number of visitors each quarter from 2001 through 2005:

	Number of Visitors (thousands)			
Year	1	2	3	4
2001	155	231	270	105
2002	182	255	351	294
2003	160	250	280	279
2004	210	310	356	353
2005	225	325	348	368

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.

QUESTION SIX **[5+2+5+8]**

The data in the table below relate to a car wash facility. The facility offered two types of wash (quick wash and full wash) in 2007. In 2008 the super wash was introduced as an additional service. Note that it is impossible to calculate a price relative for the super wash (if 2007 is the base period) because the 2007 price is undefined.

Service	*2007 Sales (E'000)*	*2008 Sales (E'000)*	*Price relative (2008 based on 2007)*
Quick wash	1.0	1.2	105.2
Full wash	0.8	0.7	102.7
Super wash		0.5	(Undefined)

i. Calculate the Paasche price index for 2008 using 2007 as the base period.
ii. The undefined price relative for the super wash service complicates the calculation of a Laspeyres price index for 2008 using 2007 as the base period. State how you would overcome this problem.
iii. Calculate the Laspeyres price index for 2008 using 2007 as the base period.
iv. Calculate the Paasche volume index for 2008 using 2007 as the base period.

END OF EXAMINATION