UNIVERSITY OF SWAZILAND

MAIN EXAMINATION PAPER 2010

TITLE OF PAPER : DESCRIPTIVE STATISTICS

COURSE CODE : ST 132

TIME ALLOWED : TWO (2) HOURS

REQUIREMENTS : CALCULATOR

INSTRUCTIONS : ANSWER ANY FIVE (5) QUESTIONS

Question 1 [20 marks, 3+3+6+8]

(a) The summary statistics for two data sets are as follows:

	Sample size	Sample mean
X data	19	7.0
Y data	25	5.1

Compute the mean of the combined data sets.

(b) At a time of economic growth but political uncertainty, a sample of 40 economists (from the population of all economists) offers the following values for the growth of an economy in the next year:

1.3	3.8	4.1	2.6	2.4	2.2	3.4	5.1
1.8	2.7	3.1	2.3	3.7	2.5	4.1	4.7
2.2	1.9	3.6	2.8	4.3	3.1	4.2	4.6
3.4	3.9	2.9	1.9	3.3	5.9	5.4	3.3
4.5	5.2	3.1	2.5	3.3	3.4	4.4	5.2

(i) Construct a stem and leaf diagram of the data and comment on the shape of the distribution.

(ii) Using the stem-and-leaf diagram construct a frequency distribution with five (5) classes for these data and then estimate the median using both the raw and grouped data. Which of these two estimates would you use and why?

(c) Given the following information on hourly earnings of women and men in a certain firm:

Hourly earnings (E)	Women	Men
4.71 - 4.90	6	5
4.91 - 5.10	31	16
5.11 - 5.30	29	25
5.31 - 5.50	19	30
5.51 - 5.70	15	24

Find the coefficients of variation for each of the distributions of hourly earnings. Comment on the variability of the hourly earnings between women and men.

Question 2 [20 marks, 8+4+4+4]

(a) *The Wall Street Journal* Stock Market Data Bank reports the numbers of shares traded on the New York Stock Exchange in half-hourly intervals. Following are the combined numbers of shares traded (in millions of shares) at half-hourly intervals for three recent days.

Shares traded (in millions)	Number of half-hourly periods
1	1
2-4	3
1-9	12
10-14	17
15-19	8
20-24	3
25-29	2
30-34	1
5-9	8

(i) Calculate the coefficient of skewness.

(ii) Estimate the interquartile range.

(b) The consumer price index for medical services has had percentage changes 4.5 percent, 3.5 percent, 2.8 percent, and 3.2 percent for the years 2005-2008 respectively. Compute the mean percentage change in prices for medical services over this time period.

(c) A chef purchases SZL100 worth of ground sirloin at SZL2.50 per kilogram and SZL100 worth of ground beef at SZL1.00 per kilogram (SZL=Swazi Lilangeni). Determine the average cost per kilogram of the meat.

Question 3 [20 marks, 2+3+7+2+4+2]

The National Highway Association is studying the relationship between the number of bidders on a highway project and the winning (lowest) bid for the project to find out whether the number of bidders increases or decreases the amount of the winning bid.

Project	Number of bidders	Winning bid ($ millions)	Project	Number of bidders	Winning bid ($ millions)
1	9	5.1	9	6	8.3
2	9	6.0	10	6	8.0
3	3	9.7	11	4	8.8
4	10	6.8	12	7	7.4
5	5	7.7	13	7	8.6
6	10	5.5	14	7	8.1
7	7	8.3	15	6	7.8
8	11	5.5			

You can use these given results: $\sum x^2 = 837, \quad \sum y^2 = 856.52, \quad \sum xy = 757.60$

(a) Identify the dependent variable (y) and the independent variable (x).

(b) Draw a scatter plot on your answer scripts (no graph paper will be given). Can you conclude from the graph whether the number of bidders increases or decreases the amount of the winning bid?

(c) Determine the regression equation. Interpret the regression coefficients.

(d) Estimate the amount of the winning bid if there were seven bidders.

(e) Compute the correlation coefficient and interpret its value.

(f) Compute the coefficient of determination and interpret its value.

Question 4 [20 marks, 7+10+3]

A private game park owner is interested in forecasting the number of visitors (in hundreds) for 2011 using the following data:

	Quarter			
Year	**I**	**II**	**III**	**IV**
2007	86	62	28	94
2008	106	82	48	114
2009	140	120	82	154
2010	162	140	100	174

(a) Compute the trend.

(b) Deseasonalise the data.

(c) What do the deseazonalized data show about the number of visitors to the park.

Question 5 [20 marks, 8+3+3+2+4]

(a) Consider the basic food items in the following table, with their unit price and per capita annual consumption:

	Unit price (in Emalangeni)		Consumption	
Food Items	**2008**	**2009**	**2008**	**2009**
Milk (litres)	7.29	7.89	117	98
Bread (loaves)	4.25	4.45	56	64
Sugar (kg)	2.19	2.45	28	20
Maize meal (kg)	5.59	5.25	58	64

(i) Compute the Laspeyre's price and consumption indices and interpret them.

(ii) Which food item showed the largest price change from 2008 to 2009?

(iii) Which food item showed the largest consumption change from 2008 to 2009?

(b) The following table shows Consumer Price Index (CPI) for the period 2005 to 2009.

Year	**CPI**
2005	95
2006	100
2007	104
2008	110
2009	120

(i) Compute CPI using 2008 as base year.

(ii) Compute the average annual percentage change in CPI during 2005 and 2009.

Question 6 [20 marks, 4+4+4+4+4]

A police officer classifies a total of 150 reported crimes in 2009 by age (in years) of the criminal and whether the crime is violent or non-violent.

	Age (in years)		
Type of crime	**Under 20**	**20 to 40**	**Over 40**
Violent	27	41	14
Non-violent	12	34	22

You must define the respective event(s) in each case and must use one of the probability rules to compute the following probabilities:

(a) What is the probability of selecting a case to analyse and finding it involved a violent crime?

(b) What is the probability of selecting a case to analyse and finding the crime was committed by some one 40 or less than 40 years old?

(c) What is the probability of selecting a case that involved a violent crime or an offender less than 20 years old?

(d) Given that a violent crime is selected for analysis, what is the probability the crime was committed by a person under 20 years old?

(e) Two crimes are selected for review by a Judge. What is the probability that both are violent crime?