CHUKA   **UNIVERSITY**

UNIVERSITY EXAMINATIONS

FIRST YEAR EXAMINATIN FOR THE AWARD OF DEGREE OF BACHELOR OF SCIENCE IN MATHEMATICS

MATH 142:   EXPLORATORY DATA ANALYSIS

STREAMS: B.Sc,   **TIME: 2 HOURS**

DAY/DATE: THURSDAY 7/12/2017   **8.30 A.M - 10.30 A.M.**

INSTRUCTIONS:
- **Answer Question ONE (Compulsory) and any other TWO Questions.**

QUESTION ONE [30 MARKS]
(a) Define the following terms as used in Statistics.

   (i) Population

   (ii) Sample

   (iii) Variable

   (iv) Quantitative variable

   (v) Qualitative variable

   (vi) Census   [6 Marks]

(b) The following is the data on weight of 50 cartons of cooking salt are given below

41	64	53	43	76	47	86	55	66	46
63	31	35	36	13	63	72	29	56	40
19	50	80	25	61	56	26	69	83	57
52	17	57	44	23	42	38	33	46	45
22	45	38	65	78	58	55	32	52	48

(i) Create a suitable grouped frequency distribution. (start with 11-20 class)   [5 Marks]

(ii) Using (i) draw a histogram and frequency polygon on the same axes.   [5 Marks]

(c) Suppose a hotel has three departments ABC fro where sales are made and the annual records of the net profit of the department for three consecutive years are presented blow.

Years	Net Profit in the department (m)		
	A	B	C
2014	3.2	3.4	2.8
2015	2.8	3.0	2.6
2016	4.0	3.2	3.6

Required:
Present the above data on a component bar chart.   [6 Marks]

(d) The weight of 30 cartons of cooking oil were recorded as follows:

320	234	335	312	324	312	300	348	325	351
340	314	315	340	308	301	316	326	350	310
326	352	327	303	350	342	328	327	328	341

Required:
Plot a stem and leaf display.   [8 Marks]

QUESTION TWO [20 MARKS]
(a) State 4 roles of statistics in Science.   [4 Marks]

(b) Discuss primary and secondary data in terms of;
   (i) Definition   [4 Marks]
   (ii) Source of data   [6 Marks]
   (iii)Illustrate with examples.   [6 Marks]

QUESTION THREE [20 MARKS]
(a) Outline 7 basic steps or stages of survey method of data collection.   [7 Marks]

(b) The table below shows the frequency distribution of ages of employees in a firm.

Age (in years)	Frequency
18-20	2
21-23	5
24-26	13
27-29	25
30-32	22
33-35	8
36-38	6
39-41	3
42-44	5
45-47	3
48-50	3
51-53	2
54-56	2
57-59	1

Required:
Prepare the cumulative frequency and the cumulative relative frequency for the ages of the employees. [4 Marks]

Determine
  (i) Mode [3 Marks]
  (ii) Median [3 Marks]
  (iii)Mean [3 Marks]

QUESTION FOUR [20 MARKS]
(a) The following tables shows the distribution of marks score by a class of 40 students in promotion examination

Marks	20-29	30-39	40-49	50-59	60-69	70-79	80-89
No. of students	6	5	7	10	5	4	3

Calculate
  (i) Quartile deviation [5 Marks]
  (ii) Standard deviation [5 Marks]

(b) The demand and prices of a bag of a hybrid 100 bags of maize in different regions of the country is as shown below

Price(X)	56	60	62	65	70	80	90
Demand (Y)	138	148	150	156	153	160	173

Calculate Pearson product moment correlation coefficient. [10 Marks]

QUESTION FIVE [20 MARKS]
(a) Outline 4 applications of Chi-Square distribution. [4 Marks]

(b) Suppose a product of four groups XYZW have a proportion of 9:3:3:1 respectively. In an experiment among 3200 units, the number in the 4 groups were 1764, 626, 574, and 236. Does the experimental result support the theoretical assumption? Alpha =5%. [6 Marks]

(c) The producer tables a random sample from persons, advertising a pre-reviewing show of the new move and obtained the result in the table below

		Age group (years)			
	$< 20\ Years$	20-39	40-50	60>	Total
Liked the movie	320	80	110	200	710
Disliked the movie	50	15	70	60	195
Indifference	30	5	20	40	95
Total	400	100	200	300	1000

Use Chi-Square test to arrive at the conclusion. [Tabe alpha=5%] [10 Marks]
....