

MATH 317 - Statistics Through Applications

- INST: i) Answer ALL Questions on this paper  
ii) every question should start on a fresh page

QUESTION 1: (15 marks)

The table below shows the final scores for 45 students in MATH 123. Use this data to complete the table below:

68	84	46	82	83	75	61	76	75
73	52	35	63	78	88	67	62	84
61	44	62	74	39	92	94	52	46
66	78	51	68	72	81	71	47	57
96	36	66	60	52	65	62	32	88

(a) Write a SAS programme that can perform the following operations (use PS=50 in the option statement);

- (i) read this data continuously (3 marks)
- (ii) produce the descriptive statistics , normal probability plot stem-leaf plot and a box plot. (3 marks)
- (iii) insert relevant remarks, titles and footnotes where necessary. (3 marks)

(b) Write a SAS programme that can draw a chart for the data above. Use seven classes (6 marks)

QUESTION 2: (20 marks)

(a) Data below shows the yield data in a maize variety trial;

Plot	Variety	yields
1	A	20
2	B	12
3	B	8
4	C	26
5	A	21
6	A	25
7	B	13
8	C	28
9	C	30

Write a SAS programme that can perform the following operations;

- (i) Transform yields by factor of 0.6 and create a new variable 'Tyields' (3 marks)
- (ii) Sort the data by varieties (3 marks)
- (iii) Print the sorted data (3 marks)
- (iv) Plot a graph of 'variety Vs yields' (3 marks)
- (v) Insert 2 titles in the programme (3 marks)

(b) Detect the mistakes and edit the programme;

```

/ height and gender /
OPTIONS PAGE SIZE=50 LINE SIZE=80 NO DATE
DATA=height;
INPUT GENDER HEIGHT @@;
DATA LINES
F 56
M 67
F 47
F 34
M 43
F 35;
TITLE GENDER VS HEIGHT STUDY;
PROC SORT DATA HEIGHT;
BY GENDER
PROC PRINT;
BY HEIGHT;
VAR HEIGHT;
PROC PLOT DATA=HEIGHT;
PLOT GENDER vs HEIGHT;
TITLE 'A PLOT OF GENDER vs HEIGHT';
FOOTNOTE3 'BY PETER MALANGI'
RUN;

```

(10 marks)

QUESTION 3: (20 marks)

(a) Write a SAS programme that will perform the Analysis of Variance (ANOVA) at  $\alpha = 0.01$  on the data set below;

*A comparative study of the yield potentials for four maize varieties*

Var 1	Var 2	Var 3	Var 4
18	23	10	25
16	19	12	29
21	28	8	24
16	22	11	26

(7 marks)

(b) State the hypothesis of the test.

(2 marks)

(c) Complete the ANOVA table below using the SAS OUTPUT;

SoV	df	SS	MS	F	p-value
Treatment					
Error					
Total					

(4 marks)

(d) Do you 'reject' or 'fail to reject' the null hypothesis ?

(2 marks)

- (e) Make a conclusion on the test. (2 marks)  
 (f) Conduct an LSD comparison between the mean yields. What can you conclude from this analysis. (2 marks)

Show appropriate titles and footnotes in the programme.

QUESTION 4: (15 marks)

(a) Using *IF*, *AND*, *THEN* and *ELSE*, write a complete SAS programme, including titles and footnotes, that you could use to grade the students in a class into the following categories. Insert the remarks 'pass' (score $\geq$ 40) or 'fail' (score $<$ 40) in the output;

Score (%)	Grade
< 40	E
40 – 49	D
50 – 59	C
60 – 74	B
$\geq$ 75	A

*Class scores (MATH 123)*

Name	Gender	Score (%)
Mwangi	m	68
Kiplono	f	79
Mutua	m	54
Kainda	m	46
Kamau	m	82
Muthoni	f	60
Wanjiku	f	31
Wacheke	f	58
Njeri	f	62
Maina	m	94
Wainaina	m	23
Muriuki	m	74
Njoki	f	42
Bulemi	m	73

(7 marks)

(b) For the data below, write a programme that will a linear regression analysis ( $y = a + bx$ );

*Examination scores*

Student	A	B	C	D	E	F	G	H	I	J
February	22	41	46	53	62	65	72	80	91	92
June	27	30	38	35	44	41	50	44	62	68

From the OUTPUT extract the values of *a*, *b* and *r*. Comment on the relationship between the February and June scores. (8 marks)