CERTIFIED PUBLIC ACCOUNTANT

FOUNDATION LEVEL 1 EXAMINATION

F1.1: BUSINESS MATHEMATICS AND QUANTITATIVE METHODS

DATE: THURSDAY 24, AUGUST 2023

MARKING GUIDE AND MODEL ANSWERS

QUESTION ONE

Marking Guide	Marks
a) Definition of semi-average method	1
Steps of computing the trend line values (1 mark each, 2 Max)	2
(Any other relevant discussion other than this award 3 marks maximum)	
Maximum Marks	**3**
b) Computation of the totals of seasonal index, seasonal effect and estimate sales (2 Marks for each total of each column, max 4)	**4**
c) Calculation of quantities Q0 and Q1 for A, B and C each (1 mark each, max 3)	3
Computation of P1Q0, P0Q0, Q1P1 and Q1Po (0.5 Marks value, max 8)	8
Computation of Fisher's quantity index (1 mark for formula, 1 mark calculation max 2)	2
Maximum Marks	**13**
Total	**20**

Model Answers

a) Definition of semi-average method

The semi-average method is used to estimate the slope and intercept of the trend line provided time-series is represented by a linear function.

Steps

1. In this method, the data are divided into two parts and their respectively arithmetic means are computed. The two arithmetic mean points are plotted corresponding to the midpoint of the class interval covered by the respective part and then these points are joined by straight line to get the required trend line.

2. The arithmetic mean of the first part is the intercept value, and the slope (change per unit time) is determined by the ratio of the difference in the arithmetic mean of the number of years between them to get a time-series of the form y=a+bx. The y equation should always be stated with reference to the year where x=0 and a description of the units of x and y.

b) Seasonal indexes are usually expressed as percentages. The total of all the seasonal indexes is 1200. The seasonal effect = (seasonal indexes)/ 100.

The yearly sales being FRW 24,000,000, the estimated monthly sales for a specified month:

$$\text{Estimated sales} = \frac{\text{Annual sales}}{12} \times \text{Seasonal effect}$$

$$= \frac{24{,}000{,}000}{12} \times \text{Seasonal effect}$$

$$= 2{,}000{,}000 \times \text{seasonal effect}$$

Month (1)	**Seasonal index (2)**	**Seasonal effect (3) = (2):100**	**Estimated sales (4) = (3) x 2,000,000**
January	75	0.75	1,500,000
February	80	0.80	1,600,000
March	98	0.98	1,960,000
April	128	1.28	2,560,000
May	137	1.37	2,740,000
June	119	1.19	2,380,000
July	102	1.02	2,040,000
August	104	1.04	2,080,000
September	100	1.00	2,000,000
October	102	1.02	2,040,000
November	82	0.82	1,640,000
December	73	0.73	1,460,000
Total	**1200**		**24,000,000**

c) The base year quantity q_0 and current year quantity q_1 for individual commodity can be calculated as follows:

$$\mathrm{q}_0(\text{for } 2021) = \frac{\text{Total value}}{\text{Price}} = \frac{50}{5} = 10; \frac{48}{4} = 6; \frac{18}{6} = 3$$

$$\mathrm{q}_1(\text{for } 2022) = \frac{\text{Total value}}{\text{Price}} = \frac{48}{4} = 12; \frac{49}{7} = 7; \frac{20}{5} = 4$$

The following table shows the calculations for Fisher's quantity index

Commodity	**2021**		**Quantity, 2022**		q_1p_0	p_0q_0	q_1p_1	q_0p_1
	p_0	q_0	p_1	q_1				
A	5	10	4	12	60	50	48	40
B	8	6	7	7	56	48	49	42
C	6	3	5	4	24	18	20	15
Total					**140**	**116**	**117**	**97**

$$Fisher'squantity\ index = \sqrt{\frac{\sum q_1p_0}{\sum q_0p_0} \times \frac{\sum q_1p_1}{\sum q_0p_1}} \times 100$$

$$= \sqrt{\frac{140}{116} \times \frac{117}{97}} \times 100 = 120.65$$

QUESTION TWO

Marking guide	**Marks**
a) Explanation of crashing and fast Tracking (2 Marks each, max 4)	**4**
b) i) Correct drawing of the network diagram (0.5 marks each activity, max 3.5)	3.5
Critical path identification	0.5
Maximum marks	**4**
ii) Calculation of normal costs	3
Normal duration	1
Maximum marks	**4**
iii) Draw a squared network and final crashed network (2 marks each, max 4)	4
Total project cost (0.5 marks for the total of each row, max 3.5)	3.5
Conclusion	0.5
Maximum marks	**8**
Total	**20**

Model Answers

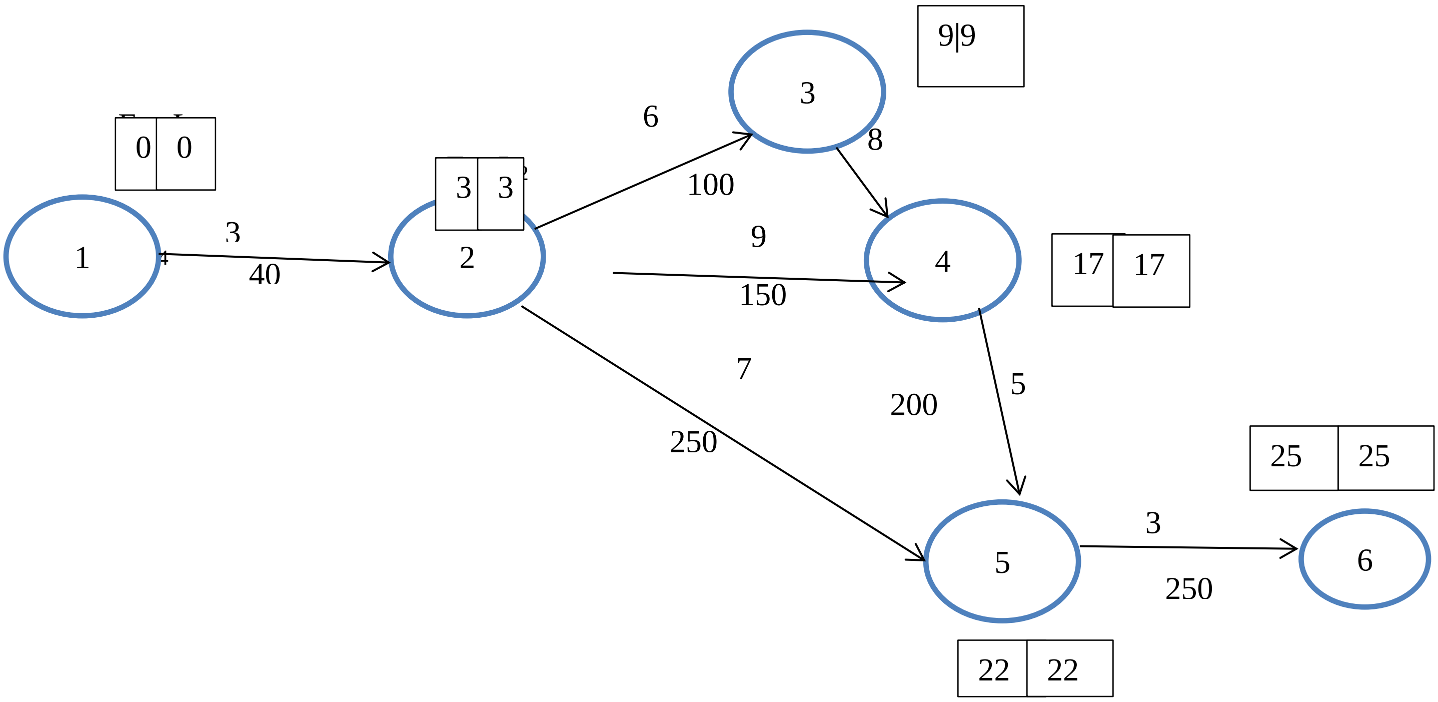
a) **Crashing** – Crashing involves either adding resources or increasing work hours (overtime, weekends) to shorten task duration. Shorter task durations typically result in higher task costs, so project teams must determine, prior to crashing, whether the total costs savings is enough to justify the higher costs. Crashing almost always requires cost increases because it usually necessitates new tasks. Crashing is a controversial technique because adding project resources can increase project complexity or risk and may ultimately have a negative impact on the schedule. Crashing does not involve reducing project scope or eliminating project tasks.

Fast Tracking – Fast tracking is a schedule compression technique in which project phases or activities usually conducted sequentially are performed in parallel to reduce duration. Care must be taken to ensure that parallel work does not create additional work or increase risk. Fast tracking frequently results in increased complexities in task dependencies, so additional project controls must be implemented to ensure ongoing and accurate insight into schedule performance.

b)

i) Network diagram

E represents Earliest Start Time
L represents Latest Start Time



Paths	Activity Duration					Total Duration
(1-2-3-4-5-6)	**3**	**6**	**8**	**5**	**3**	**25**
(1-2-4-5-6)	3	9		5	3	20
(1-2-5-6)	3	7			3	13

Critical path: 1 – 2 – 3 – 4 – 5 – 6 .

ii) Normal Project Duration: 25 days from the critical path
Normal cost = 360,000 + 1,400,000 + 2,000,000 + 1,000,000 + 400,000 + 1,600,000 + 500,000 = FRW 7,260,000
Indirect costs = FRW 150,000/day * 25 days (critical path) = FRW 3,750,000
Total Normal cost of the project is 7,260,000 + 3,750,000 = FRW 11,010,000

iii)Determine the optimum project duration, if the indirect cost is FRW150,000/day

Now draw a squared network as shown below. Choose the activities on the critical path to crash such that the present critical path continues to remain as (at least one of) the critical path. Also the cost of crashing/day shall be the least among available options at any stage.

Possible crashing

Paths	Activity Duration					Total Duration
Chrash 1-2 by 1 day						
(1-2-3-4-5-6)	**2**	**6**	**8**	**5**	**3**	**24**
(1-2-4-5-6)	2	9		5	3	19
(1-2-5-6)	2	7			3	12

Paths	Activity Duration					Total Duration
Chrash 3-4 by 4 day						
(1-2-3-4-5-6)	**2**	**6**	**4**	**5**	**3**	**20**
(1-2-4-5-6)	2	9		5	3	19
(1-2-5-6)	2	7			3	12

Paths	Activity Duration					Total Duration
Chrash 2-3 by 1 day						
(1-2-3-4-5-6)	**2**	**5**	**4**	**5**	**3**	**19**
(1-2-4-5-6)	2	9		5	3	19
(1-2-5-6)	2	7			3	12

Paths	Activity Duration					Total Duration
Chrash 4-5 by 2 day						
(1-2-3-4-5-6)	**2**	**5**	**4**	**3**	**3**	**17**
(1-2-4-5-6)	2	9		3	3	17
(1-2-5-6)	2	7			3	12

Paths	Activity Duration					Total Duration
Chrash 2-3 and 2-4 by 1 day						
(1-2-3-4-5-6)	**2**	**4**	**4**	**3**	**3**	**16**
(1-2-4-5-6)	2	8		3	3	16
(1-2-5-6)	2	7			3	12

Paths	Activity Duration					Total Duration
Chrash 5-6 by 1 day						
(1-2-3-4-5-6)	**2**	**4**	**4**	**3**	**2**	**15**
(1-2-4-5-6)	2	8		3	2	15
(1-2-5-6)	2	7			2	11

The following table is computed

Activity crashed	**Days saved (days)**	**Project duration (days)**	**Direct costs (FRW)**	**Overheads (FRW)**	**Total Project cost**
None	0	25	7260	3750	11,010
1-2	1	24	7300	3600	10,900
3-4	4	20	7500	3000	10,500
2-3	1	19	7600	2850	10,450
4-5	2	17	8000	2,550	10,550
2-3 & 2-4	1	16	8,250	2,400	10,650
5-6	1	15	8,500	2,250	10,750

The table shows that the project duration of 19 days is most economical and optimum

QUESTION THREE

Marking guide	**Marks**
a) Definition (1 Marks) and 3 components of a L.P (3 Marks each)	**4**
(Any other relevant explanation for linear programming award 4 marks accordingly)	
b) Finding the Linear Programming Problem for step 1	2
Computation of the basic feasible solution is optimal for step 2	2
Computation of the new basic feasible solution for step 3	2
Computation of the new basic feasible solution for step 4	2
Check if the basic solution is optimal	2
Maximum Marks	**10**
c) Correct table (3 Marks), basic solution (1 Mark) and corresponding cost (2 Marks)	**6**
Total Marks	**20**

Model Answers

a) **Definition**

Linear programming is the mathematical technique that deals with finding the best solution for the problems of maximizing or minimizing a linear function called the objective function subjected to constraints expressed by linear inequalities.

The word "Programming" is used here in the sense of "planning"; the necessary relationship to computer programming was incidental to the choice of name.

A *Linear Program* (LP) is a problem that can be expressed as follows (the so-called Standard Form or model): Minimize or maximize; $Z = C_o + C_1X_1 + \ldots\ldots + C_nX_n$ (Objective function)

subject to $a_{i1}x_1 + a_{i2}x_2 + \ldots\ldots + a_{in}x_n \geq b_i$ (Constraints) C is referred to as objective function coefficients, $x_1, x_2, \ldots\ldots x_n$ are decision variables

Components

Linear programming model is comprised of three components

1. The decision variables x, y or z on which to base to make a decision.
2. The objective function where the objective is either to maximize profits, revenues and contribution or to minimize costs, time and other resources.
3. The constraints which are restrictions or limitations expressed as system of linear inequalities

The decision variables in this problem are the number of Type I and II boxes to be built. They are denoted by x_1 and x_2 respectively. Since the goal is to maximize revenues and the revenues are a function of the number of boxes of each type sold, we can represent the objective function as Max Revenue $=120{,}000X_1+160{,}000X_2$

One of the constraints in this problem is availability of different types of wood. Therefore, based on the number of boxes produced, the sum of the total wood requirement must be less than or equal to the available amount of wood for each type. We can represent this type of constraint with three inequalities referring to maple, cherry and walnut respectively as follows:

$$2x_1 \leq 10$$
$$3x_2 \leq 11$$
$$x_1 + x_2 \leq 5$$

In addition, there are the non-negativity constraints which ensure that our solution does not have negative number of boxes. These constraints are shown as

$$x_1, x_2 \geq 0$$

Step 1:
Convert the linear program into standard form.
The linear program in standard form is

$$
\begin{array}{llllllll}
\max z = & 120x_1 & +160x_2 & & & & & \\
s.t. & 2x_1 & & +s_1 & & & =10 \\
 & & +3x_2 & & +s_2 & & =11 \\
 & x_1 & +x_2 & & & +s_3 & =5 \\
 & & & & x_1, x_2, s_1, s_2, s_3 & & \geq 0
\end{array}
$$

Step 2:
Determine if the basic feasible solution is optimal.
At this step we create the tableau for this basic feasible solution reproduced as Table 1.

Table 1. The initial tableau in proper form

Basic	Z	x_1	x_2	s_1	s_2	s_3	RHS	Ratio
Z	1	-120	-160	0	0	0	0	
s_1	0	2	0	1	0	0	10	None
s_2	0	0	3	0	1	0	11	$11/3$
s_3	0	1	1	0	0	1	5	5

Step 3:
Use elementary row operations to solve for the new basic feasible solution. Return to Step 2
The new basic feasible solution is shown in Table 2

Table 2. The tableau for the new basic feasible solution in the first iteration

Basic	Z	x_1	x_2	s_1	s_2	s_3	RHS	Ratio
Z	1	-120	0	0	$160/3$	0	$1760/3$	
s_1	0	2	0	1	0	0	10	5
x_2	0	0	1	0	$1/3$	0	$11/3$	None
s_3	0	1	0	0	$-1/3$	1	$4/3$	$4/3$

Step 4:
Use elementary row operations to solve for the new basic feasible solution. The new basic feasible solution is shown in Table 3.

Table 3. The tableau for the basic feasible solution in the second iteration.

Basic	Z	x_1	x_2	s_1	s_2	s_3	RHS	Ratio
Z	1	0	0	0	$40/3$	120	$2240/3$	
s_1	0	0	0	1	$2/3$	-2	$22/3$	
x_2	0	0	1	0	$1/3$	0	$11/3$	
x_1	0	1	0	0	$-1/3$	1	$4/3$	

Step 5:

Determine if the basic feasible solution is optimal.

Since there are no negative coefficients in row 0, we have reached the optimal solution where the objective function value is $2240/3$ and

$s_1 = 22/3$

$x_2 = 11/3$

$x_1 = 4/3$

$s_2 = s_3 = 0$

Note that all these values can be read from the tableau shown in Table 5. This solution also corresponds to the extreme point B in Figure 4 which was also determined to be optimal using the graphical solution approach.

Finally, the woodworker should build $4/3$ Type I boxes and $11/3$ Type II boxes to maximize his revenue to FRW 586,667.

b) The allocation is shown in the following tableau

Factories	Retail Agency					Capacity
	1	2	3	4	5	
A	1(50)	9	13	36	51	50(50)
B	24	12	16	20	1	100
C	14	33	1	23	26	150
Requirement	100(50)	60	50	50	40	300

Factories	Retail Agency					Capacity
	1	2	3	4	5	
B	24(50)	12	16	20	1	100(50)
C	14	33	1	23	26	150
Requirement	50(50)	60	50	50	40	250

Factories	Retail Agency				Capacity
	2	3	4	5	
B	12(50)	16	20	1	50(50)
C	33	1	23	26	150
Requirement	60 (50)	50	50	40	200

Factories	Retail Agency				Capacity
	2	3	4	5	
C	33(10)	1(50)	23(50)	26(40)	150(150)
Requirement	10(10)	50(50)	50(50)	40(40)	200

The starting **basic solutio**n is given as

$x_{11} = 50$, $x_{21} = 50$, $x_{22} = 50$ $x_{32} = 10$, $x_{33} = 50$, $x_{34} = 50$, $x_{35} = 40$

The corresponding transportation cost is

Schedule	**Transportation cost per unit**	**Quantity**	**Amount**
A-1	1	50	50
B-1	24	50	1200
B-2	12	50	600
C-2	33	10	330
C-3	1	50	50
C-4	23	50	1150
C-5	26	40	1040
Total		300	4420

QUESTION FOUR
Marking guide

Marking guide	**Marks**
a) Assumptions for building the model (1 Mark each, max 3)	**3**
b) Maximin strategy (0.5 marks each correct row in the table, max 2)	2
Minimax strategy (0.5 marks each correct row in the table, max 2.5)	2.5
Identification of the saddle point	0.5
Interpretation	2
Maximum marks	**7**
c) i) Computation of payoff values	1
Computation of daily expectations	1
Maximum marks	**2**
ii) Conditional payoff table (0.5 marks each correct column, max 3.5)	3.5
Expected payoffs and EMV table (0.5 marks each correct column, max 3.5)	3.5
Conclusion	1
Maximum marks	**8**
Total	**20**

Model Answers

a) Assumptions for building a model of two-person zero sum game:

i) Each player has available to him a finite number of possible courses of action. Sometimes the set of courses of action may be the same for each player. Or, certain courses of action may be available to both players while each player may have certain specific courses of action which are not available to the other player.

ii) Player A attempts to maximize gains to himself. Player B tries to minimize losses to himself.

iii)The decisions of both players are made individually prior to the play with no communication between them.

iv)The decisions are made and announced simultaneously so that neither player has an advantage resulting from direct knowledge of the other player's decision.

v) Both players know the possible payoffs of themselves and their opponents.

B) Maximin Strategy

First consider the minimum of each row and the pick the maximum out of minimum values

Row	**Minimum value**
A	-3
B	-1
C	2
D	12

Maximum of {-3, -1, 2, 12} = 12

Minimax Strategy

Next consider the maximum of each column and then choose the minimum of those maximum values**.**

Column	**Maximum value**
X	15
Y	14
Z	18
W	12
V	20

Minimum of {15, 14, 18, 12, 20} = 12

Identification of the saddle point

We see that the maximum of row minima = the minimum of the column maxima. So, the game has a saddle point. The common value is 12. Therefore, the value V of the game =12.

Interpretation:

In the long run, the following best strategies will be identified by the two players:
The best strategy for player Isaac is strategy d.
The best strategy for player Butera is strategy W.

C) Given that FRW 90 000 is the fixed cost and FRW 200 is variable cost. The payoff values with 4 cars at the disposal of decision –maker is calculated as under:

Number of cars Demanded:	0	1	2	3	4
Payoff Values (‘000’):	0-(90x4)	200-(90x4)	400-(90x4)	600-(90x4)	800-(90x4)
With 4 cars	= -360	= -160	= 40	= 240	= 440

Thus, the daily expectation is obtained by multiplying the payoff values with the given corresponding probabilities of demand:
Daily expectation = (-360) (0.1) +(-160) (0.2) +(40) (0.3) +(240) (0.2) +(440) (0.2) =FRW 80,000

ii) Cars to buy

The conditional payoffs and for each course of action is shown in the following table

Demand of cars	**Probability**	**Conditional payoff due to the decision to purchase cars (course of action (000’FRW))**				
		0	**1**	**2**	**3**	**4**
0	0.1	0	-90	-180	-270	-360
1	0.2	0	110	20	-70	-160
2	0.3	0	110	220	130	40
3	0.2	0	110	220	330	240

4	0.2	0	110	220	330	440

The expected payoffs and EMV is captured in the following table

Demand of cars	**Probability**	**Conditional payoff due to the decision to purchase cars (course of action (000'FRW))**				
		0	**1**	**2**	**3**	**4**
0	0.1	0	-9	-18	-27	-36
1	0.2	0	22	4	-14	-32
2	0.3	0	33	66	39	12
3	0.2	0	22	44	66	48
4	0.2	0	22	44	66	88
EMV				140	130	80

Since the EMV of FRW140,000 for the course of action 2 is the highest, the company should buy 2 cars.

QUESTION FIVE

Marking Guide	**Marks**
a) i) Explanation of mutually exclusive events	1
ii) Explanation of impossible events	1
Maximum marks	**2**
b) Probability for passing Business Mathematics (1 mark for each calculation)	**3**
c) Probability of one firm being selected (1 mark for each calculation)	**5**
d) i) Probability of 2 or less calls per minute (1 Mark for each step calculated)	2
ii) Probability of up to 4 calls per minute (1 Mark for each step calculated)	2
iii) Probability of more than 4 calls (1 Mark for each step calculated)	1
Maximum marks	**5**
e) Computation of summation of x*P(x)	1
Computation of summation of x^2*P(x)	1
Computation of expected value E(x)	1
Computation of variance	1
Computation of standard deviation	1
Maximum marks	**5**
Total	**20**

Model Answers

a) i) Mutually exclusive events: They are events which cannot happen at the same time on a single trial. Example: Suppose a coin is tossed then either head or tail occurs. If head occurs, then tail will not occur or vice-versa.

ii) Impossible events: These are events which cannot occur at all. They are elements which are not in the sample space. Example: If a dice is rolled, the appearance of 7 is an impossible event.

b) Let A be event that the student passes in Law test and B the event that the student passes in Business Mathematics test. We are again given

$$P(A) = \frac{2}{3}, P(A \cap B) = \frac{14}{25}, and\ P(A \cup B) = \frac{4}{5}$$

We want P(B)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\frac{4}{5} = \frac{2}{3} + P(B) - \frac{14}{25}$$

$$P(B) = \frac{4}{5} + \frac{14}{25} - \frac{2}{3} = \frac{60 + 42 - 50}{75} = \frac{52}{75} = 0.69$$

c) Let A and B be the event that an MBA will be selected in firm X and will be rejected in firm Y, respectively. Then given that

$$P(A) = 0.7, P(\bar{A}) = 1 - 0.7 = 0.3$$

$$P(B) = 0.5, P(\bar{B}) = 1 - 0.5 = 0.5\ and\ P(\bar{A} \cup \bar{B}\ ) = 0.6$$

Since $P(A \cap B) = 1 - P(\bar{A} \cup \bar{B}) = 1 - 0.6 = 0.4$

Therefore, probability that he will be selected by one of the firms is given by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= 0.7 + 0.5 - 0.4 = 0.8$$

Thus, the probability of an MBA being selected by one of the firm is 0.8.

d) Given that the average number of calls per minute = m =4

Thus,

$$P(x) = \frac{e^{-m}m^x}{x!} = \frac{e^{-4}4^x}{x!}$$

i) Probability of 2 or less call per minute

$$P(X \leq 2) = P(0) + P(1) + P(2)$$

$$= \frac{e^{-4}4^0}{0!} + \frac{e^{-4}4^1}{1!} + \frac{e^{-4}4^2}{2!}$$

= 0.0183x13 = 0.24

ii) Probability up to 4 calls

$$P(X \leq 4) = P(0) + P(1) + P(2) + P(3) + P(4)$$

$$= 0.24 + P(3) + P(4) = 0.63$$

iii)Probability of more than 4 calls

$$P(X > 4) = 1 - P(X \leq 4)$$

= 1- 0.63 = 0.37

e) The necessary calculations required to calculate expected and standard deviation of a random variable, let say x are shown in the following table:

Number of cars per household (x)	**Relative frequency [P(x)]**	$\boldsymbol{x \times P(x)}$	$\boldsymbol{x^2 \times P(x)}$
0	0.10	0.00	0.00
1	0.30	0.30	0.30
2	0.40	0.80	1.60
3	0.12	0.36	1.08
4	0.06	0.24	0.96
5	0.02	0.10	0.50
Total	1.00	**1.80**	**4.44**

Expected value, $E(x) = \sum xP(x) = 1.80$. This value indicates that there are on an average 1.8 cars per household.

Variance $\sigma^2 = \sum x^2 P(x) - [E(x)]^2 = 4.44 - (1.80)^2 = 4.44 - 3.24 = 1.20$

Standard deviation $\sigma = \sqrt{\sigma^2} = \sqrt{1.20} = 1.095$ cars (1 car).

QUESTION SIX

Marking Guide	**Marks**
a) Advantages of stratified sampling (1 Mark each advantage)	**3**
b) Computation of log $_{10}$ x (0.5 Marks each correct log)	2
Computation of ω log $_{10}$ x (0.5 Marks each correct log)	2
Computation of the two totals in the table (0.5 Marks each)	1
Formula for weighted geometric mean	0.5
Calculation of weighted geometric mean (0.5 Marks each step calculated)	1.5
Maximum marks	**7**
c) i) Total wage bill paid for x	2
Total wage bill paid for y	2
Conclusion	1
Maximum marks	**5**
ii) Calculation of combined mean	2
Computation of d_1 and d_2 (0.5 Marks each)	1
Computation of combined standard deviation (0.5 Marks each step)	2
Maximum marks	**5**
Total	**20**

Model Answers

a) Advantages of stratified sampling: stratification serves many useful purposes. The principal among them are:

- Precision of the estimate is increased.
- By virtue of stratification, supervision of field work is more convenient and simpler.
- Stratified sampling also ensures a better cross-section of population than that possible under unstratified sampling.
- Stratification makes it possible to use different sampling designs in different strata.
- Selection and identification of units is simple. They can easily be enumerated.
- Non sampling errors are very minimised.

b) Let weight of fourth number be ω. Then the weighted geometric mean of four numbers can be calculated as shown in below table

Number (x)	Weight of Each Number (ω)	$log_{10}x$	$\omega \log_{10} x$
8	5	0.9031	4.5155
25	3	1.3979	4.1937
17	4	1.2304	4.9216
30	ω	1.4771	1.4771 ω
Total	12+ ω		13.6308+1.4771 ω

Thus, the weighted Geometric Mean (G.M) is

$$\log\{G.M(\omega)\} = \left[\left(\frac{1}{\sum\omega}\right)\sum\omega\log x\right]$$

$$\log(15.3) = [\left(\frac{1}{12+\omega}\right)(13.6308 + 1.4771\omega)]$$

$$1.1847(12+\omega) = 13.6308+1.4771\,\omega$$

$$14.2164+1.1847\omega = 13.6308+1.4771\,\omega$$

$$0.5856 = 0.2924\,\omega$$

$$\omega = \frac{0.5856}{0.2924} = 2 \text{ (approx..)}$$

The weight of the fourth number is 2.

c) i) Comparing the total wages to find out which organization X or Y pays larger number of monthly wages.

Total wage bill paid monthly by X and Y is

$$X: n_1 \times \bar{x}_1 = 550 \times 5000 = FRW\ \ 2{,}750{,}000$$

$$Y: n_2 \times \bar{x}_2 = 650 \times 4500 = FRW\ \ 2{,}925{,}000$$

Organization Y pays a larger amount as monthly wages as compared to organization X.

ii) Calculation of combined variation,

First calculating the combined mean as follows:

Combined mean; $\bar{x}_{12} = \frac{n_1\,\bar{x}_2 + n_2\,\bar{x}_2}{n_1+n_2} = \frac{2{,}750{,}000+2{,}925{,}000}{1200} = FRW\ 4{,}729.166$

Compute $d_1 = \bar{x}_{12} - \bar{x}_1 = 4{,}729.166 - 5{,}000 = -270.834$

$d_2 = \bar{x}_{12} - \bar{x}_2 = 4{,}729.166 - 4{,}500 = 229.166$

Then computed standard deviation of the two organizations combined

$$\sigma_{12} = \sqrt{\frac{n_1(\sigma_1^2+d_1^2)+n_2(\sigma_2^2+d_2^2)}{n_1+n_2}}$$

$$= \sqrt{\frac{550\ (900+73{,}351.05\ )+650(1600+52{,}517.05\ )}{550+650}}$$

$$= \sqrt{\frac{40{,}838{,}077.5+\ 35{,}176{,}082.5}{1200}}$$

$$= \sqrt{63345.13}\ = 251.68$$

QUESTION SEVEN

Marking Guide	**Marks**
a) Definition of scatter diagram	1
Explanation about how it shows the relationship between 2 variables	1
Maximum marks	**2**
b) i) Computation of the totals in the table from column 4 to 8 (1 Mark each)	5
Computation of the totals of x and y in the table (0.5 Marks each)	1
Formula for regression line Y on X	1
Computation of mean for Y and X (0.5 Marks for formula and 0.5 for Calculation for each)	2
Calculation of regression (1 Mark each step calculated)	2
Prediction	1
Maximum marks	**12**
ii) Calculation of b_{xy}	1
Calculation of the regression line X on Y	2
Prediction	1
Maximum marks	**4**
iii) The Karl Pearson's coefficient of correlation (1 mark for each step)	**2**
Total	**20**

Model Answers

a) The scatter diagram also known as dottogram, is a graph of plotted points where each point represents the value of x and y as a coordinate. It portrays the relationship between these two variables graphically. By looking at the scatter of the various points on the chart, it is possible to determine the extent of the association between these two variables.

b) Regression

Column 1	**Column 2**	**Column 3**	**Column 4**	**Column 5**	**Column 6**	**Column 7**	**Column 8**
SN	**X**	**Y**	**X-66**	**Y-67**	**dx dy**	**dx^2**	**dy^2**
			dx	dy			
1	63	66	-3	-1	3	9	1
2	65	68	-1	1	-1	1	1
3	66	65	0	-2	0	0	4
4	67	67	1	0	0	1	0
5	67	69	1	2	2	1	4
6	68	70	2	3	6	4	9
Total	**396**	**405**	**0**	**3**	**10**	**16**	**19**

i) Regression equation of Y on X: $Y - \bar{Y} = b_{yx}(X - \bar{X})$

$$\bar{X} = \frac{\sum X}{N} = \frac{396}{6} = 66$$

$$\bar{Y} = \frac{\sum Y}{N} = \frac{405}{6} = 67.5$$

$$b_{yx} = \frac{\sum xy}{\sum x^2} = \frac{10}{16} = 0.625$$

Substituting the values in the equation,

$$Y_c - 67.5 = 0.625(X - 66) \Rightarrow Y_c = 0.625X + 26.25$$

Hence if the height of the father is 70 inches or X=70, the height of the son would be

$$Y_c = 0.625(70) + 26.25 = 70$$

ii) Regression equation of X on Y: $X - \bar{X} = b_{xy}(Y - \bar{Y})$

Where $b_{xy} = \frac{N\sum dxdy - \sum dx.\sum dy}{Ndy^2 - (\sum dy)^2} = \frac{60}{105} = 0.571$

Substituting the values in the equation,

$$X_c - 66 = 0.571(Y - 67.5) \Rightarrow X_c = 0.571Y + 27.458$$

For Y=65, $X_c = 0.571(65) + 27.458 = 64.573$

iii)The Karl Pearson's coefficient of correlation (r)

$$r = \sqrt{b_{xy} \times b_{yx}} = \sqrt{0.571 \times 0.625} = \sqrt{0.356875}$$

$$r = \pm 0.597$$

END OF MARKING GUIDE AND MODEL ANSWERS