POKHARA UNIVERSITY

Level: Bachelor

Semester: Spring

Year: 2021

Programme: BBA/BI/TT/BCIS

Full Marks: 100

Course: Data Analysis and Modeling

Pass Marks: 45

Time: 3 hrs.

Candidates are required to answer in their own words as far as practicable. The figures in the margin indicate full marks.

Section "A"
Very Short Answer Questions

Attempt all the questions. [10×2]

1. Interpret the slope of the regression model $\hat{Y} = 35-10X$. What is the estimated value of Y when X=1.5.

2. The Coefficient of correlation between variables y and x is -0.86. Calculate and interpret the value of the coefficient of determination.

3. Complete the following ANOVA summary table which was obtained form a multiple regression model with four independent variables:

Sources	Degree of freedom (d.f.)	Sum of square (SS)	Mean square (MS)	F
Regression	?	58	?	
Error	?	?	?	
Total	22	160	?	

4. What is the cost of living index number? What does it measure?

5. Describe the importance of the time series analysis in business decision making.

6. Find the active and inactive constraints in the following LPP;

    $Maximize\ Z = 2x + 4y$

    $Subject\ to\ x + y \le 14, 3x + 2y \ge 30, 2x + y \le 18, x \ge 0, y \ge 0.$

    $where\ \max z = 36\ at\ x = 6\ and\ y = 6$

7. What do you understand by balanced and unbalanced transportation problems?

8. Convert the following profit matrix into opportunity loss matrix.

Job	P	Q	R	S
A	23	11	16	15
B	18	12	12	13
C	19	15	17	14
D	16	13	11	17

9. From the given information calculate forecast value for the month April using Exponential smoothing constant 0.1 and forecast for January was 195 units.

Month	Demand	Forecast
January	200 units	195 units
February	110 units	?
March	300 units	?
April	-	?

10. Draw network diagram for the following data:

Activity	A	B	C	D	E	F	G
Predecessors	-	-	A	B	B	CD	EF

Section "B"
Descriptive Answer Questions

Attempt **any six** questions. [6×10]

11. A farm administers a test to sales trainees before they go into the field. The management of the farm is interested in determining the relationship between the test scores and the sales made by the trainees at the end of one year in the field. The following data were collected for 10 sales personnel who have been in the field:

Sales Person Number	Test Score(X)	Number of units sold(Y)
1	2.6	95
2	3.7	140
3	2.4	85
4	4.5	180
5	2.6	100
6	5.0	195
7	2.8	115
8	3.0	136
9	4.0	175
10	3.4	150

Calculation shows that:

$$\sum X = 34 \quad \sum Y = 1371 \quad \sum X^2 = 122.62 \quad \sum Y^2 = 201121 \quad \sum XY = 4954$$

a) Find the correlation coefficient between the test score and number of units sold, examine if this linear relationship is significant at the 5% level of significance.

b) Compute the 90% prediction interval for the number of units sold of salesperson 9.

c) Find coefficient of determination and interpret its meaning.

12. From the following data show that Fisher's index number is an ideal index number.

Commodities	Base Year (2021)		Current Year (2022)	
	Price	Quantity	Price	Quantity
A	4	8	8	6
B	5	10	16	5
C	8	14	10	10
D	4	19	4	13

13. Suppose you are financial analyst of ABC Corporation whose working capital requirement over the last seven years are presented in following table.

Year	2021	2020	2019	2018	2017	2016	2015
Working Capital (in mil)	2.2	2.1	2.4	2.6	3	3.4	3.8

a) Find the linear estimating equation that best describes the data.
b) Calculate the percent of trend for these data

c) Calculate the relative cyclical residual for these data.

d) In which year does the largest fluctuation from trend occur and is same for both methods?

14. A manufacturer of ski clothing makes ski pants and ski jackets. The profit on a pair of ski pants is $2.00 and on a jacket is $1.50. Both pants and jackets require the work of sewing operators and cutters. There are 60 minutes of sewing operator time and 48 minutes of cutter time available. It takes 8 minutes to sew one pair of ski pants and 4 minutes to sew one jacket. Cutters take 4 minutes on pants and 8 minutes on a jacket. Find the maximum profit and the amount of pants and jackets to maximize the profit.

15. Solve the following transportation problem for minimum cost by taking initial feasible solution by Vogel's Approximation Method. The entries in the matrix indicate the cost in rupees of transporting a unit from a particular source to a particular destination.

Origin	Destination				Availability
	D_1	D_2	D_3	D_4	
O_1	10	8	11	7	20
O_2	9	12	14	6	40
O_3	8	9	12	10	35
Requirement	16	18	31	30	95

16. a) The following table shows the number of visitors who tourist area over a 10-year period.

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
No. of visitors	28	27	33	25	34	33	35	30	33	35

   i.   Forecast the number of visitors in 2008 by using Naïve model. Use the first 6 period as a warm up sample.

   ii. Also compute the MSE and MAD for the forecasting sample.

   b) There are four machine on which to do three jobs. The cost (in Rs.) of each job on each machine is given in the following table.

Job	Machine			
	W	X	Y	Z
A	18	24	28	32
B	8	13	17	19
C	10	15	11	22

   Assign the job to each machine to minimize costs.

17. Following table lists the activity of a project along with their time estimates.

Activity	Predecessor	Most likely (t_m)	Optimistic (t_o)	Pessimistic (t_p)
A	-	5	4	6
B	-	12	8	16
C	A	5	4	12
D	B	3	1	5
E	D,A	2	2	2
F	B	6	4	8
G	C,E,F	14	10	18
H	G	20	18	34

The scheduled completion date for this project is 62 days.

a) Draw the network diagram and compute the expected duration of the project.

b) Find the probability that the project will be finished within the scheduled date.

c) Find the probability that the project will be completed at least 6 days prior to the expected time.

<div align="center">

Section "C"
Case Analysis

</div>

18. *Read the case situation given below and answer the questions that follow:* [20]

A manager selected a representative sample of 24 monthly customer bills taken from several recent heating seasons. The manager considers kilowatt hours per month (Y) as a liner function of square feet heated space (X_1), an index of roof insulation quality (X_2), PRESENCE/ABSENCE of insulated windows (X_3), mean temperature (X_4) and heat pump/electric forced air (X_5). **The SPSS output is as following**

	Unstandarized Coefficients(β)	Std.error	T	Sig.
Intercept	6356.07	838.701	7.58	0.0000
X_1	0.56038	0.15811	3.54	0.0023
X_2	-31.2077	8.95905	-3.48	0.0027
X_3	-327.503	149.169		
X_4	-113.895	16.2604	-7.00	0.0000
X_5	-621.485	147.828	-4.20	0.0005

ANOVA TABLE

Sources	Degree of freedom (d.f.)	Sum of square (SS)	Mean square (MS)	F
Regression	5			
Residual		2166000		
Total	23	14370000		

a) Fit a multiple regression equation.

b) Test which of the independent variable makes significant contribution to the model.

c) Compute the standard error of the estimate (S_e).

d) Compute the r^2 and interpret its meaning

e) Obtain confidence interval estimate of coefficient of X_3.

f) Given that $X_1 = 129$, $X_2 = 18$, $X_3 = 5$, $X_4 = 3$, $X_5 = 1129$, estimate the value of Y.

g) Set up the null and alternative hypothesis, carry out F-test and interpret your result.