

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BCom DEGREE EXAMINATION DECEMBER 2024
(First Semester)

Branch – COMMERCE (BUSINESS ANALYTICS)

STATISTICS FOR BUSINESS ANALYTICS

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer ALL questions

ALL questions carry EQUAL marks

(10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	Which of the following is a primary method of data collection? a) Textbook b) Survey c) Online article d) Lecture	K1	CO1
	2	The classification and tabulation of data falls under: a) Descriptive statistics b) Inferential statistics c) Qualitative data analysis d) Predictive analysis	K2	CO1
2	3	Which of the following is NOT a one-dimensional diagram? a) Bar chart b) Line chart c) Histogram d) Pie chart	K1	CO2
	4	What does a time series graph typically show? a) Comparison of multiple datasets b) Relationship between variables c) Distribution over time d) Distribution of frequencies	K2	CO2
3	5	Simple random sampling without replacement means: a) A sample item is not returned to the population after selection b) Each item has equal probability of being selected again c) Each group is randomly selected d) The population is divided into strata	K1	CO3
	6	In cluster sampling, clusters are formed based on: a) Random chance b) Demographics c) Probability d) Geographic location	K2	CO3
4	7	Which measure of central tendency is least affected by extreme values? a) Arithmetic mean b) Median c) Mode d) Geometric mean	K1	CO4
	8	The harmonic mean is best used in cases involving: a) Skewed distributions b) Rates and ratios c) Normal distributions d) Percentile calculations	K2	CO4
5	9	Quartile deviation measures: a) Total range of data b) Average deviation from the mean c) Spread of the middle 50% of data d) Standard deviation	K1	CO5
	10	Kurtosis refers to the: a) Skewness of data b) Flatness or peakedness of a distribution c) Variability in data d) Measure of central tendency	K2	CO5

Cont...

SECTION - B (35 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks

(5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO																	
1	11.a.	Describe the steps involved in collecting primary and secondary data. Provide examples of when each method would be appropriate.	K2	CO1																	
	(OR)																				
	11.b.	Define statistics and discuss its origin and development.																			
2	12.a.	Use the following data on the weight of products (in kg) to construct a histogram: Weight: 10, 15, 20, 25, 30 Frequency: 5, 10, 15, 8, 2	K3	CO2																	
	(OR)																				
	12.b.	Explain the construction of a time series graph.																			
3	13.a.	Describe the sampling errors	K2	CO3																	
	(OR)																				
	13.b.	Illustrate the cluster sampling method with a real-life example.																			
4	14.a.	Describe the Geometric mean and Harmonic mean with its limitations	K3	CO4																	
	(OR)																				
	14.b.	Compute the mode from the following: <table><tr><td>Size</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td></tr><tr><td>Frequency</td><td>20</td><td>22</td><td>25</td><td>29</td><td>30</td><td>18</td><td>14</td><td>13</td><td>12</td></tr></table>			Size	20	21	22	23	24	25	26	27	28	Frequency	20	22	25	29	30	18
Size	20	21	22	23	24	25	26	27	28												
Frequency	20	22	25	29	30	18	14	13	12												
5	15.a.	Discuss on the Skewness and Kurosis	K2	CO5																	
	(OR)																				
	15.b.	Describe the standard deviation and range with its limitations.																			

SECTION - C (30 Marks)

Answer ANY THREE questions

ALL questions carry EQUAL Marks

(3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO																		
1	16	Explain the key stages of statistical investigation.	K2	CO1																		
2	17	Discuss the two-dimensional diagrams with some examples	K4	CO 2																		
3	18	Elucidate the concept of Simple random sampling with and without replacement by examples.	K2	CO 3																		
4	19	Compute Mean, Q2 and Mode from the following distribution:	K3	CO 4																		
		<table><tr><td>Marks</td><td>10-20</td><td>20-30</td><td>30-40</td><td>40-50</td><td>50-60</td><td>60-70</td><td>70-80</td><td>80-90</td></tr><tr><td>No. of Students</td><td>14</td><td>17</td><td>25</td><td>21</td><td>16</td><td>19</td><td>14</td><td>13</td></tr></table>			Marks	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	No. of Students	14	17	25	21	16	19	14	13
		Marks			10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90										
No. of Students	14	17	25	21	16	19	14	13														
5	20	Compute variance, standard deviation and mean deviation from the following distribution:	K2	CO 5																		
		<table><tr><td>X</td><td>0-10</td><td>10-20</td><td>20-30</td><td>30-40</td><td>40-50</td><td>50-60</td><td>60-70</td><td>70-80</td></tr><tr><td>F</td><td>14</td><td>18</td><td>20</td><td>25</td><td>30</td><td>15</td><td>19</td><td>21</td></tr></table>			X	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	F	14	18	20	25	30	15	19	21
		X			0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80										
F	14	18	20	25	30	15	19	21														

Z-Z-Z

END

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BCom DEGREE EXAMINATION DECEMBER 2024
(First Semester)

Branch – COMMERCE (BUSINESS ANALYTICS)

MATHEMATICAL TECHNIQUES FOR BUSINESS ANALYTICS

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer ALL questions
ALL questions carry EQUAL marks (10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	What percentage of 10 is 25? a) 400% b) 500% c) 250% d) 120%	K1	CO1
	2	In _____ series the common constant which multiplies any term to yield the following term is called the common ratio. a) arithmetic b) geometric c) harmonic d) hyperbolic	K2	CO1
2	3	A _____ matrix is a diagonal matrix in which all the diagonal elements are equal. a) skew symmetric b) scalar c) diagonal d) square	K1	CO2
	4	The rank of a matrix is the order of the largest square sub matrix whose determinant is not _____. a) - I b) I c) 0 d) -1	K2	CO2
3	5	The x- coordinate is also called as _____. a) slope b) plane c) ordinate d) abscissa	K1	CO3
	6	_____ is the slope intercept form. a) $y = mx + c$ b) $y = mx - c$ c) $y = c - mx$ d) $y = \frac{m}{x}$	K2	CO3
4	7	The differential coefficient of e^{2x-3} is _____. a) $2e^{2x-3}$ b) $2e^{2x-3} + 3$ c) $2e^{2x-3} - 3$ d) $(2x-3)e^{2x-3}$	K1	CO4
	8	Price elasticity of demand is _____. a) $-\frac{p}{x} \frac{dp}{dx}$ b) $\frac{p}{x} \frac{dx}{dp}$ c) $\frac{p}{x} \frac{dp}{dx}$ d) $-\frac{p}{x} \frac{dx}{dp}$	K2	CO4
5	9	$\int e^{2x} dx =$ _____. a) $-\frac{1}{2}e^{2x} + c$ b) $\frac{1}{2}e^{2x} + c$ c) $\frac{1}{2}e^{-2x} + c$ d) $\frac{1}{2x}e^{2x} + c$	K1	CO
	10	If $\int f(x)dx = F(x)$ then then $\int_a^b f(x)dx =$ _____. a) $F(b) + F(a)$ b) $F(a) - F(b)$ c) $F(b) - F(a)$ d) $F(b) * F(a)$	K2	CO5

SECTION - B (35 Marks)

Answer ALL questions
ALL questions carry EQUAL Marks (5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Two firms A and B have the same list price for an item. A offers discounts of 25% and 15%, B offers discounts of 20%, 10% and 10%. Which is the more advantageous to the purchaser?	K3	CO1
		(OR)		
	11.b.	What sum of money must be paid now to receive a freehold estate that will yield Rs.1200 a year, 10years hence if the rate of compound interest is 4% per annum?		

Cont...

Cont...

2	12.a.	Verify $(AB)C = A(BC)$ for $A = \begin{pmatrix} 2 & 1 \\ 4 & -1 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 0 & -1 \end{pmatrix}$ $C = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 4 & 1 \\ -1 & 3 & 2 \end{pmatrix}$	K3	CO2
	(OR)			
	12.b.	Find the inverse of $\begin{pmatrix} 1 & 2 & -1 \\ 3 & 8 & 2 \\ 4 & 9 & -1 \end{pmatrix}$		
3	13.a.	i) Find the equation of the line whose slope is $3/2$ which cuts off 3 units along OY. ii) As the number of units manufactured increases from 2000 to 3000, the total cost of production increases from Rs.11000 to Rs 15000. Find the relation between the cost(y) and the number of units made (x), if the relationship is linear.	K3	CO3
	(OR)			
	13.b.	Find the point of intersection of the lines $x+3y-5=0$, $x-2y+5=0$ and the equation of the line passing through the point (2, -3) having the slope $-5/7$.		
4	14.a.	Compute $\frac{d(x^n)}{dx}$ and $\frac{d(e^x)}{dx}$ from the first principle.	K3	CO4
	(OR)			
	14.b.	The total cost function y for x units is given by $y = \frac{3x(x+7)}{x+5} + 5$. Show that the marginal cost decreases continuously as the output increases.		
5	15.a.	Evaluate $\int \frac{x^3 - x + 4}{x^2}$ with respect to x.	K3	CO5
	(OR)			
	15.b.	The marginal cost function for producing x units is $y=23+16x-3x^2$ and the total cost for producing one unit is Rs.40. Obtain the total cost function and the average cost function.		

SECTION -C (30 Marks)

Answer ANY THREE questions

ALL questions carry EQUAL Marks

(3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO																				
1	16	A man has Rs.1000 to his credit in his account on 1 st January. The bank allows interest at 4% per annum which is calculated on 30 th June and 31 st December and charge $\frac{1}{4}$ % commission for collecting and paying cheques. At the end of each month the bank collect cheques to the value Rs.400 for account holder and paid out a total of Rs.300. Find the sum in his account on 30 th June.	K3	CO1																				
2	17	Consider an economy of 2 industries P and Q, where the data in millions of Rs. is given below. <table border="1"> <tr> <td></td><td>P</td><td>Q</td><td>FINAL DEMAND</td><td>TOTAL OUTPUT</td></tr> <tr> <td>P</td><td>14</td><td>6</td><td>8</td><td>28</td></tr> <tr> <td>PRODUCER</td><td></td><td></td><td></td><td></td></tr> <tr> <td>Q</td><td>7</td><td>18</td><td>11</td><td>36</td></tr> </table>		P	Q	FINAL DEMAND	TOTAL OUTPUT	P	14	6	8	28	PRODUCER					Q	7	18	11	36	K3	CO2
	P	Q	FINAL DEMAND	TOTAL OUTPUT																				
P	14	6	8	28																				
PRODUCER																								
Q	7	18	11	36																				
3	18	i) Sam buys a radio making a payment of Rs.200 at the time of purchase with the agreement that he will pay at the rate of Rs.15 for the next 20 months. Find the relationship between the amount (y) he has paid and the number of months (x) since he bought the radio. ii) When the price is Rs.50, 60 cameras of a particular type are available and when the price is Rs.80,140 cameras of the same type are available in the market. Determine the supply curve.	K3	CO3																				
4	19	If $y = ae^{mx} + be^{-mx}$ show that $\frac{d^2 y}{dx^2} = m^2 y$	K3	CO4																				
5	20	Solve $\int (\log x)^3$ with respect to x.	K3	CO5																				

P Z-Z-Z

END

**PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)**

**BCom DEGREE EXAMINATION DECEMBER 2024
(Second Semester)**

Branch – **COMMERCE (BUSINESS ANALYTICS)**

APPLIED BUSINESS STATISTICS - I

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** marks

(10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	The sum of the probabilities of all the events of an experiment are a) 0 b) 0.5 c) 1 d) 2	K1	CO1
	2	If A and B are two events, the probability of occurrence of either A or B is given as a) $P(A B)$ b) $P(A \cup B)$ c) $P(A \cap B)$ d) $P(A)P(B)$	K2	CO1
2	3	A _____ random variable assumes all the values in the given range a) discrete b) continuous c) both a & b d) neither a nor b	K1	CO2
	4	The distribution whose mean and variance are equal is a) Binomial b) Poisson c) Normal d) Gamma	K2	CO2
3	5	Large sample theory is applicable when: a) $n > 30$ b) $n < 30$ c) $n = 30$ d) $n \leq 30$	K1	CO3
	6	A hypothesis that completely specifies the distribution is _____ hypothesis a) simple b) composite c) null d) alternative	K2	CO3
4	7	Analysis of variance technique is applied for testing the homogeneity of a) proportions b) means c) variances d) standard deviations	K1	CO4
	8	Goodness of test of a distribution is tested by a) t-test b) F-test c) chi square test d) Z-test	K2	CO4
5	9	Distribution-free tests are also known as a) parametric tests b) non parametric tests c) z-tests d) t-tests	K1	CO5
	10	When testing for randomness, we can use a) Mann-Whitney U test b) Sign test c) Runs test d) t-test	K2	CO5

SECTION - B (35 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** Marks

(5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Prove the addition theorem on probability.	K2	CO1
		(OR)		
	11.b.	Explain the following terms: Axiomatic approach to probability and conditional probability.		
2	12.a.	Identify the differences between discrete and continuous random variables.	K3	CO2
		(OR)		
	12.b.	The probability that a part time student of a college will graduate is 0.6. Solve and find the probability that out of 6 students (a) none (b) one and (c) atleast one will graduate.		

Cont...

3	13.a.	Differentiate the following pairs of concepts: (i) Null hypothesis and Alternative hypothesis (ii) Type I and Type II errors	K3	CO3											
	(OR)														
	13.b.	A sample of 400 male students is found to have a mean height of 171.38 cm. can it be reasonably regarded as a sample from a large population with mean height 171.17cm and standard deviation 3.3 cm.													
4	14.a.	Describe the steps involved in testing the significance of two variances.	K4	CO4											
	(OR)														
	14.b.	A sample of 790 units of a manufactured product was classified according to the quality of the product and the production shift is given below: <table><tr><th rowspan="2">Production shift</th><th colspan="2">Quality</th></tr><tr><th>Good</th><th>Bad</th></tr><tr><td>Day Shift</td><td>30</td><td>160</td></tr><tr><td>Night Shift</td><td>140</td><td>460</td></tr></table> Using Chi square test, find whether depends on shift?			Production shift	Quality		Good	Bad	Day Shift	30	160	Night Shift	140	460
Production shift	Quality														
	Good	Bad													
Day Shift	30	160													
Night Shift	140	460													
5	15.a.	Describe how to perform sign test for two samples.	K4	CO5											
	(OR)														
	15.b.	Analyze the following sequence of time intervals to determine if it can be considered a random sequence of numbers using the run test. 27, 55, 57, 57, 43, 55, 51, 37, 41, 34, 40, 56, 49, 52, 14, 56, 48, 33, 49, 38.													

SECTION -C (30 Marks)

Answer ANY THREE questions

ALL questions carry EQUAL Marks (3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO
1	16	A company has two plants to manufacture scooters. Plant I manufactures 80% of the scooters and plant II manufactures 20%. At plant I 85% scooters and at plant II 65% of scooters are rated as standard quality. What is the probability that a scooter selected at random (if it is known that the scooter is of standard quality) came from (i) plant I, (ii) plant II?	K4	CO1
2	17	For the following density function $f(x) = Cx^2(1-x)$, $0 \leq x \leq 1$. Find C, mean and variance.	K4	CO2
3	18	1000 articles from a factory A are examined and found to have 3% defectives. 1500 similar articles from a second factory B are found to have only 2% defectives. Can it reasonably be concluded that the product of the first factory is inferior to the second?	K4	CO3
4	19	Describe the procedure of performing an one way analysis of variance with an example	K4	CO4
5	20	Describe the steps and circumstances in which the Mann-Whitney U test is conducted.	K4	CO5

Z-Z-Z

END

**PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)**

**BCom DEGREE EXAMINATION DECEMBER 2024
(Third Semester)**

Branch – **COMMERCE (BUSINESS ANALYTICS)**

NoSQL-MongoDB

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** marks

(10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	Which type of consistency is commonly associated with NoSQL databases? a) Strong consistency b) Immediate consistency c) Eventual consistency d) Transactional consistency	K1	CO1
	2	What type of data model does MongoDB use? a) Relational model b) Document-based model c) Key-value pair model d) Columnar model	K2	CO1
2	3	What is the primary purpose of the XLRD library in Python? a) To read XML files b) To read Excel files (XLS) c) To parse CSV files d) To process JSON data	K1	CO2
	4	Which of the following is NOT a valid CRUD operation in MongoDB? a) Create b) Read c) Update d) Transfer	K2	CO2
3	5	Which of the following is an example of a field query in MongoDB? a) Finding documents where a specific field exists b) Returning only specific fields from documents c) Sorting documents by field value d) Limiting the number of documents returned	K1	CO3
	6	Which MongoDB method is used to limit the number of documents returned in a query? a) .sort() b) .find() c) .limit() d) .aggregate()	K2	CO3
4	7	Which of the following is a valid way to create a user in MongoDB? a) db.createUser({}) b) db.createAdmin({}) c) mongo.createUser() d) db.addUser({})	K1	CO4
	8	Which of the following tools is used to back up MongoDB databases? a) mongobackup b) mongodump c) mongoexport d) mongosave	K2	CO4
5	9	Which of the following MongoDB operators is used to perform text searches in collections? a) \$regex b) \$text c) \$group d) \$match	K1	CO5
	10	Which MongoDB feature is most suitable for processing and analyzing large text datasets? a) Indexes b) Aggregation Pipeline c) MapReduce d) Replication	K2	CO5

Cont...

SECTION - B (35 Marks)Answer **ALL** questions**ALL** questions carry **EQUAL** Marks (5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Explain how the CAP theorem affects database design.	K2	CO1
		(OR)		
	11.b.	Compare the performance of MongoDB and RDBMS in large-scale applications.		
2	12.a.	Build a script to parse a CSV file and extract specific data columns.	K3	CO2
		(OR)		
	12.b.	Identify how XLRD helps in reading Excel files.		
3	13.a.	Illustrate how sorting and limiting can improve the efficiency of query results.	K2	CO3
		(OR)		
	13.b.	Explain out the main stages of an aggregation pipeline.		
4	14.a.	Analyze about data replication in MongoDB.	K4	CO4
		(OR)		
	14.b.	Compare the performance of a query with and without an index on a large dataset.		
5	15.a.	List out the purpose of the MapReduce function in MongoDB.	K4	CO5
		(OR)		
	15.b.	Classify the common use cases for text processing using MapReduce in MongoDB.		

SECTION -C (30 Marks)Answer **ANY THREE** questions**ALL** questions carry **EQUAL** Marks (3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO
1	16	Compare the benefits and challenges of sharing versus replication in MongoDB server configuration.	K4	CO1
2	17	Analyze basic CRUD operations on a MongoDB collection using the Mongo shell.	K4	CO2
3	18	Evaluate the advantages of using an aggregation pipeline over traditional queries in MongoDB.	K5	CO3
4	19	Design a complex database schema in MongoDB that handles multiple types of relationships (one-to-one, one-to-many, and many-to-many).	K5	CO4
5	20	Discuss what MongoDB feature is used for full-text search and text indexing.	K6	CO5

Z-Z-Z

END

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)
BCom DEGREE EXAMINATION DECEMBER 2024
(Third Semester)

Branch – **COMMERCE (BUSINESS ANALYTICS)**

ECONOMETRICS

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** marks

(10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	In a regression analysis, what is another name for the variable that helps predict the outcome? (a) Explained Variable (b) Response Variable (c) Dependent Variable (d) Explanatory Variable	K1	CO1
	2	Which of the following is a limitation of econometric analysis? (a) The assumption of linearity (b) The availability of data (c) The use of statistical methods (d) The focus on economic phenomena	K2	CO1
2	3	Which of the following is a common test for detecting multicollinearity? (a) Durbin-Watson test (b) Variance Inflation Factor (VIF) (c) Breusch-Pagan test (d) Chow test	K1	CO2
	4	A bivariate econometric model involves (a) One dependent variable and one independent variable (b) One dependent variable and multiple independent variables (c) Multiple dependent variables and one independent variable (d) Multiple dependent variables independent variables	K2	CO2
3	5	Which term describes when the spread of data points around a regression line is inconsistent? (a) Evenly scattered (b) Unevenly scattered (c) Heteroscedastic (d) Homoscedastic	K1	CO2
	6	A dummy variable is: (a) A continuous variable that takes on values between 0 and 1 (b) A categorical variable that takes on only two values, 0 and 1 (c) A dependent variable in a regression model (d) An independent variable in a regression model	K2	CO2
4	7	Input-output analysis is primarily used to: (a) Analyze the demand for goods and services (b) Analyze the supply of goods and services (c) Analyze the interdependencies between industries (d) Analyze the impact of government policies	K1	CO3
	8	A closed input-output model differs from an open model by: (a) Including foreign trade (b) Excluding intermediate goods (c) Assuming a fixed level of technology (d) Government Policies	K2	CO3
5	9	Which of the following best describes autocorrelation in time series data? (a) The correlation between a variable and its own past values (b) The correlation between two different variables over time (c) The variance of a variable over time (d) The mean of a variable over time	K1	CO3
	10	How can the concept of autocorrelation be tested in time series data? (a) Using the Durbin-Watson test (b) Using the Breusch-Pagan test (c) Using the F-test (d) Using the t-test	K2	CO3

Cont...

SECTION - B (35 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks (5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Explain the scope and Limitation of Econometric approach.	K2	CO1
		(OR)		
	11.b.	Interpret the coefficients in a two-variable linear regression model.	K2	
2	12.a.	Differentiate bivariate modeling from multivariate modeling.	K3	CO2
		(OR)		
	12.b.	Identify statistical tests for detecting multicollinearity.	K3	
3	13.a.	Construct the Log Modelling in Econometrics.	K3	CO2
		(OR)		
	13.b.	Identify the use of Dummy Variables in Seasonal Analysis.	K3	
4	14.a.	Examine the assumptions of input-output analysis.	K4	CO3
		(OR)		
	14.b.	We assume that in a village there is a farmer, carpenter, and a tailor, who provide the three essential goods: food, shelter, and clothing. Suppose the farmer himself consumes 40% of the food he produces, and gives 40% to the carpenter, and 20% to the tailor. Thirty percent of the carpenter's production is consumed by himself, 40% by the farmer, and 30% by the carpenter. Fifty percent of the tailor's production is used by himself, 30% by the farmer, and 20% by the tailor. Analyze the matrix that describes this closed model.		
5	15.a.	Analyze the steps involved in the Box-Jenkins approach.	K4	CO3
		(OR)		
	15.b.	List out the Assumptions of Auto Correlation.	K4	

SECTION - C (30 Marks)

Answer ANY THREE questions

ALL questions carry EQUAL Marks (3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO																
1	16	Illustrate the Errors of Measurement and its testing Method.	K4	CO1																
2	17	Examine the Causes of Multicollinearity.	K4	CO2																
3	18	Analyze the Tests for detecting Heteroskedasticity.	K4	CO2																
4	19	Let us assume that whatever the farmer produces, 20% is used by him, 15% by the carpenter, 10% by the tailor, and the consumer uses the other 40 billion dollars' worth of the food. Ten percent of the carpenter's production is used by him, 25% by the farmer, 5% by the tailor, and 50 billion dollars' worth by the consumer. Fifteen percent of the clothing is used by the tailor, 10% by the farmer, 5% by the carpenter, and the remaining 60 billion dollars' worth by the consumer. Evaluate the internal consumption in the following table, and express the demand as the matrix D. <table><tr><td></td><td>F produces</td><td>C produces</td><td>T produces</td></tr><tr><td>F uses</td><td>.20</td><td>.25</td><td>.10</td></tr><tr><td>C uses</td><td>.15</td><td>.10</td><td>.05</td></tr><tr><td>T uses</td><td>.10</td><td>.05</td><td>.15</td></tr></table> The consumer demand for each industry in billions of dollars is given below. 40 D=50 60		F produces	C produces	T produces	F uses	.20	.25	.10	C uses	.15	.10	.05	T uses	.10	.05	.15	K4	CO3
	F produces	C produces	T produces																	
F uses	.20	.25	.10																	
C uses	.15	.10	.05																	
T uses	.10	.05	.15																	
5	20	Categorize the various Tests for Auto Correlation.	K4	CO3																

Z-Z-Z

END

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BCom DEGREE EXAMINATION DECEMBER 2024
(Fourth Semester)

Branch – COMMERCE (BUSINESS ANALYTICS)

R PROGRAMMING

Time: Three Hours

Maximum: 50 Marks

SECTION-A (5 Marks)

Answer ALL questions

ALL questions carry EQUAL marks

(5 x 1 = 5)

- 1 What is R primarily used for?
(i) Web development (ii) Statistical analysis
(iii) Graphic design (iv) Game development
- 2 Which function is used to check for missing values in R?
(i) is.na() (ii) is.null()
(iii) is.nan() (iv) is.empty()
- 3 What is a list in R?
(i) A collection of vectors of the same type (ii) A collection of objects of different types
(iii) A matrix with named rows and columns (iv) A type of vector
- 4 Which of the following is a valid for loop syntax in R?
(i) for (i in 1:10) { print(i) } (ii) for i in 1:10 { print(i) }
(iii) for (i : 1 to 10) { print(i) } (iv) for (i = 1; i <= 10; i++) { print(i) }
- 5 What does the pch argument control in the plot() function?
(i) The color of the points (ii) The type of plot
(iii) The shape of the points (iv) The size of the plot

SECTION - B (15 Marks)

Answer ALL Questions

ALL Questions Carry EQUAL Marks

(5 x 3 = 15)

- 6 a List the different ways to create a vector in R. Provide examples for each.
OR
b Describe what a data frame is in R. How does it differ from a vector?
- 7 a Explain about 2-dimensional array in R.
OR
b Discuss how to multiply two matrices in R.
- 8 a Explain the purpose of the attach() and deattach() functions.
OR
b How can you combine two data frames in R? Describe the difference between rbind() and cbind() with examples.

Cont...

- 9 a Discuss the structure of if and else if statements in R. Provide suitable examples.
OR
b With suitable examples explain the purpose of the switch() function in R.
- 10 a Explain how to create a histogram in R.
OR
b Describe how to add a legend to a plot in R.

SECTION -C (30 Marks)Answer any **Three** questions**ALL** questions carry **EQUAL** Marks

(3 x 10 = 30)

- 11 Explain the R environment and its primary components. How do you interact with R?
- 12 Define an object in R. What are the different types of objects available in R? Provide examples for each type.
- 13 Explain the scan() function in R. Provide an example for reading data from a text file.
- 14 Discuss about for, while and repeat loops with suitable examples.
- 15 Explain the purpose of the plot() function in R. Provide an example for creating a basic scatter plot with labeled axes.

Z-Z-Z

END

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BCom DEGREE EXAMINATION DECEMBER 2024
(Fifth Semester)

Branch – COMMERCE (BUSINESS ANALYTICS)

APPLIED BUSINESS STATISTICS - III

Time: Three Hours

Maximum: 50 Marks

SECTION-A (5 Marks)

Answer ALL questions

ALL questions carry EQUAL marks (5 x 1 = 5)

1. Which method is commonly used to estimate the parameters of a regression model?
 - a) Maximum likelihood
 - b) Least squares
 - c) Bayesian estimation
 - d) Ordinary differential equations
2. What is extrapolation?
 - a) Estimating values outside the range of a dataset
 - b) Estimating values within the range of a dataset
 - c) Finding the median of a dataset
 - d) Calculating averages
3. What is the primary purpose of statistical quality control?
 - a) To improve product design
 - b) To ensure products meet quality standards
 - c) To reduce production costs
 - d) To increase sales
4. ISO 9001:2000 is primarily focused on which aspect?
 - a) Environmental management
 - b) Quality management systems
 - c) Financial management
 - d) Supply chain management
5. In reliability engineering, what is the goal of using hazard rates?
 - a) To increase production speed
 - b) To reduce operational costs
 - c) To enhance understanding of failure processes
 - d) To predict market trends

SECTION - B (15 Marks)

Answer ALL Questions

ALL Questions Carry EQUAL Marks (5 x 3 = 15)

- 6 a If you have a dataset with an R-squared value of 0.75, how much of the variability in the dependent variable is unexplained?
- OR
- b From the following data, show the regression equations X on Y and Y on X by normal equations.

X	6	2	10	4	8
Y	9	11	5	8	7

Cont...

- 7 a Identify the Newton's methods of forward and backward differences.
OR
b Apply the Lagrange interpolation method.
- 8 a A quality control manager uses an R-chart and collects the following ranges from five samples: 2, 3, 4, 2, 5. Test for the average range?
OR
b A company produces 1,000 items and finds 10 defective. Compare the defect rate?
- 9 a Classify the benefits of implementing TQM.
OR
b Categorize the challenges do organizations face when implementing TQM?
- 10 a Criticize the scope of reliability engineering.
OR
b Explain the concept of hazard rate and its importance in reliability.

SECTION -C (30 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks (5 x 6 = 30)

- 11 a. Summarize the lines of two regression for the following data and calculate the coefficient of correlation. Also obtain the value of Y when X = 10.

X	1	2	3	4	5
Y	10	20	15	25	30

OR

- b. From the following information, show the expected value of Y when X = 10.

	X	Y
Mean	7.6	14.80
Standard deviation	3.6	2.5
Co-efficient of correlation is	0.99	

- 12 a. Choose the advantages of using interpolation methods.
OR
b. Select the role of the Lagrange basis polynomials in interpolation.
- 13 a. A factory records the following daily output (in units): 100, 105, 98, 110, and 95. Calculate the variance.
OR
b. A quality control inspector finds that 5 out of 100 sampled products are defective. Test for the p-value and the defect rate?
- 14 a. Analyse the future trends in TQM.
OR
b. Categorize the ISO 9001:2000 quality management standard.
- 15 a. Explain the significance of reliability standards and certifications.
OR
b. Choose the methods used for reliability testing.

Z-Z-Z

END

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BCom DEGREE EXAMINATION DECEMBER 2024
(Fifth Semester)

Branch – COMMERCE (BUSINESS ANALYTICS)

**MAJOR ELECTIVE COURSE – I: DATA MINING AND BUSINESS
INTELLIGENCE**

Time: Three Hours

Maximum: 50 Marks

SECTION-A (5 Marks)

Answer ALL questions

ALL questions carry EQUAL marks

(5 x 1 = 5)

1. In today's perspective, how has Business Intelligence (BI) evolved?
 - a) BI now focuses solely on producing static reports
 - b) BI has shifted towards real-time analytics and self-service tools for business users
 - c) BI has become obsolete with the introduction of block chain technology
 - d) BI now exclusively uses on-premise servers
2. The operation of "Slice and Dice" in OLAP refers to:
 - a) Viewing data from different perspectives by selecting specific dimensions or attributes
 - b) Performing complex calculations on data
 - c) Rotating data in different axes to change its order
 - d) Transforming data from one format to another
3. What is the main objective of data transformation?
 - a) To improve data visualization
 - b) To reduce the size of the database
 - c) To make the data compatible for mining processes
 - d) To remove noisy data
4. In Incremental Association Rule Mining, which of the following is true?
 - a) The entire dataset is reprocessed when new data is added.
 - b) New rules are generated only for the newly added data.
 - c) Existing rules are deleted when new data is added.
 - d) It is a classification technique.
5. Big Data refers to:
 - a) Large volumes of data that can be analyzed for insights
 - b) High-speed data transmission networks
 - c) Financial data transactions
 - d) Cloud storage solutions

SECTION - B (15 Marks)

Answer ALL Questions

ALL Questions Carry EQUAL Marks

(5 x 3 = 15)

6. a) Narrate the features of building blocks.
(OR)
b) Classify the main components of a data warehouse.
7. a) Compare OLTP and OLAP system.
(OR)
b) Bring out the Design the data warehouse architecture.

Cont...

8. a) Sketch the major issues in DM.
(OR)
b) Analyse why do we need data preprocessing step in data mining.
- 9 a) Explain the concept of Association Rule Mining and its application in Market Basket Analysis.
(OR)
b) Describe the mining clustering Techniques.
- 10 a) Summarize the role of data mining in fraud detection.
(OR)
b) Examine the use of data mining techniques in the banking and finance sectors.

SECTION -C (30 Marks)Answer any **Three** questions**ALL** questions carry **EQUAL** Marks

(3 x 10 = 30)

11. Identify the significant trends in data warehousing.
12. Examine in detailed about Data Warehouse architecture with neat diagram.
13. Elucidate the classification of DM System.
14. Discuss various clustering techniques used in data mining.
15. Describe the application of data mining in the telecommunications Industry.

Z-Z-Z

END