

BUSINESS ANALYTICS
(CSEN 4242)

Time Allotted : 2½ hrs

Full Marks : 60

Figures out of the right margin indicate full marks.

*Candidates are required to answer Group A and
any 4 (four) from Group B to E, taking one from each group.*

Candidates are required to give answer in their own words as far as practicable.

Group – A

1. Answer any twelve:

12 × 1 = 12

Choose the correct alternative for the following

- (i) If the assumed hypothesis is tested for rejection considering it to be true is called?
(a) Null Hypothesis (b) Statistical Hypothesis
(c) Simple Hypothesis (d) Composite Hypothesis
- (ii) A statement whose validity is tested on the basis of a sample is called?
(a) Null Hypothesis (b) Statistical Hypothesis
(c) Simple Hypothesis (d) Composite Hypothesis
- (iii) A hypothesis which defines the population distribution is called?
(a) Null Hypothesis (b) Statistical Hypothesis
(c) Simple Hypothesis (d) Composite Hypothesis
- (iv) Which of the following are types of correlation?
(a) Positive and Negative (b) Simple, Partial and Multiple
(c) Linear and Nonlinear (d) All of the above
- (v) Which of the following is true for the coefficient of correlation?
(a) The coefficient of correlation is not dependent on the change of scale
(b) The coefficient of correlation is not dependent on the change of origin
(c) The coefficient of correlation is not dependent on both the change of scale and change of origin
(d) None of the above
- (vi) Which of the following statements is true for correlation analysis?
(a) It is a bivariate analysis (b) It is a multivariate analysis
(c) It is a univariate analysis (d) Both (a) and (c)
- (vii) Which of the following statements is true about the null hypothesis?
(a) Any wrong decision related to the null hypothesis results in two types of errors
(b) Any wrong decision related to the null hypothesis results in one type of an error
(c) Any wrong decision related to the null hypothesis results in four types of errors
(d) Any wrong decision related to the null hypothesis results in three types of errors

- (viii) Which of the following statements is true about the type two error?
 (a) Type two error means to accept an incorrect hypothesis
 (b) Type two error means to reject an incorrect hypothesis
 (c) Type two error means to accept a correct hypothesis
 (d) Type two error means to reject a correct hypothesis
- (ix) By 2025, the volume of digital data will increase to
 (a) TB (b) YB (c) ZB (d) EB Advertisement
- (x) Data Analysis is a process of
 (a) inspecting data (b) cleaning data
 (c) transforming data (d) All of Above

Fill in the blanks with the correct word

- (xi) _____ as a result of data accessibility, data latency, data availability, or limits on bandwidth in relation to the size of inputs
- (xii) For a standard normal variate, the value of Standard Deviation is _____.
- (xiii) Normal Distribution is also known as _____.
- (xiv) In Normal distribution, the highest value of ordinate occurs at _____.
- (xv) The shape of the normal curve depends on its _____.

Group - B

2. (a) Find the weighted average mean of the following test grades. [[CO3](ANALYSE/ANALYSE/HOCQ)]
- Given two school classes — one with 20 students, one with 30 students — and test grades in each class as follows:
 - Morning class = {62, 67, 71, 74, 76, 77, 78, 79, 79, 80, 80, 81, 81, 82, 83, 84, 86, 89, 93, 98}
 - Afternoon class = {81, 82, 83, 84, 85, 86, 87, 87, 88, 88, 89, 89, 89, 90, 90, 90, 90, 91, 91, 91, 92, 92, 93, 93, 94, 95, 96, 97, 98, 99}
- (b) Find the mean, median, mode, max and min of the following data. [[CO3](APPLY/IOCQ)]

Name	Weight
Tom	62
Dick	57
Harry	52
John	65
Jane	62
Jenny	59
Ram	82
Shyam	67
Jadu	56
Madhu	52

- (c) Which is better measure of central tendency mean or median and why?
 [[CO3](REMEMBER/LOCQ)]
6 + 4 + 2 = 12
3. (a) Define analytics. [[CO1](REMEMBER/LOCQ)]
 (b) Give 4 benefits of analytics. [[CO3](APPLY/IOCQ)]
 (c) Explain any 3 domains of analytics with their functionalities. [[CO2](APPLY/IOCQ)]
2 + 4 + 6 = 12

Group - C

4. We have an online training program designed to upgrade the supervisory skills of the managers. Since the program is self-administered, different supervisors require different completion times. A study of the past participants indicates that the mean length of the time spent on the program is 500 hours and that this normally distributed random variable has a standard deviation of 100 hours.
- (i) What is the probability that a particular participants selected in random will require more than 500 hours to complete the program? [[CO4](REMEMBER/LOCQ)]
 (ii) What is the probability a candidate selected at random will take between 500 and 650 hours to complete the training program? [[CO4](REMEMBER/LOCQ)]
 (iii) What is the probability a candidate selected at random will take less than 580 hours to complete the training program? [[CO4](APPLY/IOCQ)]
 (iv) What is the probability a candidate selected at random will take between 420 and 570 hours to complete the training program? [[CO4](APPLY/IOCQ)]
- From Z table:**
 $P(Z < 1.5) = 0.9332$
 $P(Z < 0) = 0.5$
 $P(Z < 0.8) = 0.7881$
 $P(Z < 0.7) = 0.7580$
 $P(Z < -0.8) = 0.2119$
- (2 + 2 + 4 + 4) = 12**

5. Critically evaluate the limitations of applying normal distribution to real-world business problems.
- (i) List three limitations. [[CO4](ANALYSE/HOCQ)]
 (ii) Suggest alternatives or adjustments for these cases. [[CO4](ANALYSE/HOCQ)]
(6 + 6) = 12

Group - D

6. Explain the concept of exponential smoothing in time series forecasting.
- (i) Define simple exponential smoothing with its formula. [[CO5](REMEMBER/IOCQ)]
 (ii) When is it appropriate to use this method? [[CO5](REMEMBER/IOCQ)]
 (iii) Compare it with moving average in terms of responsiveness to change. [[CO5](REMEMBER/IOCQ)]
(4 + 4 + 4) = 12

7. Find the sales for PC based on trend

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2023	100	121	142	162	179	200	225	258	272	297	319	330
2024	139	154	189	228	251	290	318	347	383	408	422	438

Find the mean yearly forecast and monthly sales forecast for year 2025.

[[CO5](ANALYSE/HOCQ)]

12

Group - E

8. Calculate the two regression equations of X on Y and Y on X from the data given below, taking deviations from actual means of X and Y.

Price (Rs.)	10	12	13	12	16	15
Amount demanded	40	38	43	45	37	43

Estimate the likely demand when the price is Rs. 20.

[[CO5](ANALYSE/HOCQ)]

(8 + 4) = 12

9. (a) Define linear regression.
 (b) Write the equation linear regression.
 (c) The formula used for linear regressions is, $y = a + bx$.
 Give the values of coefficient a and b
 (d) Write 4 properties of linear regression.

[[CO5](REMEMBER/LOCQ)]

[[CO5](REMEMBER/LOCQ)]

[[CO5](APPLY/IOCQ)]

[[CO5](REMEMBER/LOCQ)]

2 + 2 + 4 + 4 = 12

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	16.67	39.58	43.75