BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(QUIZ-2 CONTINUOUS INTERNAL ASSESSMENT)

CLASS: BCA
BRANCH: COMPUTER SCIENCE

SEMESTER: IV
SESSION: SP/2023

SUBJECT: CA273 DATA ANALYTICS

TIME: 40 MINUTES

FULL MARKS: 10

INSTRUCTIONS:
1. The total marks of the questions are 10.
2. Candidates attempt for all 10 marks.
3. Before attempting the question paper, be sure that you have got the correct question paper.
4. The missing data, if any, may be assumed suitably.
5. Tables/Data handbook/Graph paper etc. to be supplied to the candidates in the examination hall.

			CO	BL
Q1	Write a short note on random sampling.	[2]	3	1
Q2	Suppose you are asked to conduct a survey on the smoking habits of ABC University teachers. How will you proceed?	[2]	3	4
Q3	6% of the people from the city Noida are grey eyed and 3% of the people from the city New Delhi are grey eyed. A random sample of 1000 people from the city Noida and 1200 people from the city New Delhi are selected. What is the probability that the difference in the sample proportion is less than 4%?	[2]	3	5
Q4	Differentiate between null hypothesis and alternative hypothesis.	[2]	3	4
Q5	A plastic part is designed to tolerate an average pressure of 100 units. A random sample of size 150 is tested and it is found that the average pressure which these parts can withstand is 98 units with a standard deviation of 10 units. Test that the lot meets the specification with 1% significance level.	[2]	3	5

:::::: 05/04/2023 :::::: E

BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(MID SEMESTER EXAMINATION SP2023)

CLASS: BCA
BRANCH: BCA

SEMESTER : IV
SESSION : SP2023

TIME: 02 Hours

SUBJECT: CA273 DATA ANALYTICS

FULL MARKS: 25

INSTRUCTIONS:
1. The question paper contains 5 questions each of 5 marks and total 25 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
4. Tables/Data handbook/Graph paper etc., if applicable, will be supplied to the candidates

---

			CO	BL
Q.1(a)	Differentiate interval data and ratio data.	[2]	1	1
Q.1(b)	Find the mean of the marks obtained by 40 students in an exam given below.	[3]	1	6

Marks	25	30	15	20	24
Number of Students	8	12	10	6	4

Q.2(a)	Differentiate skewness and kurtosis.	[2]	1	4
Q.2(b)	Construct a box plot for the following data set. 3,5,8,8,9,11,12,12,13,13,16	[3]	1	6
Q.3(a)	In real-world data, tuples with missing values for some attributes are a common occurrence. Describe various methods for handling this problem.	[2]	2	2
Q.3(b)	Use smoothing by bin means to smooth the data, using a bin depth of 3. Illustrate your steps for the following data set: 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70	[3]	2	

Q.4(a) Find Karl Pearson's coefficient of correlation for the following data: [2] 2 6

Price (Rs.)	10	12	14	16	18
Quantity (Units)	20	29	21	22	28

Q.4(b) Discuss any two types of data transformation methods with suitable examples. [3] 2 2

Q.5(a)	Differentiate between sampling and non-sampling errors.	[2]	3	3
Q.5(b)	Discuss any two methods each from random sampling and non-random sampling types.	[3]	3	3

::::::23/02/2023::::::M

BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)

CLASS: BCA
BRANCH: BCA

SEMESTER: IV
SESSION: SP/2023

SUBJECT: CA 273 Data Analytics

TIME: 3 Hours

FULL MARKS: 50

INSTRUCTIONS:
1. The question paper contains 5 questions each of 10 marks and total 50 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
4. Before attempting the question paper, be sure that you have got the correct question paper.
5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.

			CO	BL
Q.1(a)	Explain the term Data Analytics. Under the descriptive data analytics explain the data distribution's standard deviation and skewness with suitable example.	[5]	1	1
Q.1(b)	Draw the following chart for the data of yearly expenditure on various items as: (in Rupees) given in the table. a) % pie chart b) Vertical bar chart.	[5]	1	2

Items	Food	Clothes	Entertainment	Transport
Expenditure	40000	5000	2000	3000

Q.2(a)	Explain the term data correlation between two quantitative data. Find the Pearson's correlation coeff. for the following data given in the table below:	[5]	2	2

X	3	5	8	10
Y	5	6	10	12

Q.2(b)	What is the significance of Z transformation in data analytics? Explain the method to find the Z transformation with suitable example.	[5]	2	2
Q.3(a)	Explain the usefulness of sampling in data analytics. Also explain the various non-random sampling methods with suitable examples.	[5]	3	3
Q.3(b)	In a sample of 100 plastic chairs produced by a company is found to be 1550 days with standard deviation of 80 days. Test the hypothesis that, the average life of the chairs produced by the company is **less than or equal to** 1570 days with 5% significance Level. The critical value z (alpha = 0.05) is 1.645. Show the necessary diagram also.	[5]	3	4
Q.4(a)	Explain the social network and its applications with suitable examples. Also Explain the various terms used to measure the size of social network.	[5]	4	2
Q.4(b)	Explain the various steps applied in text data analysis for documents similarities task.	[5]	4	3
Q.5(a)	What are the outliers in a given dataset? Explain the various type of outliers with suitable example.	[5]	5	3
Q.5(b)	Explain the method based on interquartile range to find out the outliers present in the univariate data set. Show the suitable sketch of box plot to mention the position of outliers.	[5]	5	4

:::::27/04/2023 E:::::