JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

Make-up Examination-Nov-2025

COURSE CODE (CREDITS): 20B1WCI531 (2) MAX. MARKS: 25

COURSE NAME: FOUNDATION FOR DATA SCIENCE AND VISUALIZATION

COURSE INSTRUCTORS: RBT MAX. TIME: 1 Hour 30 Minutes

Note: Note: (a) All questions are compulsory.

(b) The candidate is allowed to make Suitable numeric assumptions wherever required

for solving problems (Calculators allowed)

Q.No	Question	CO	Marks
Q1	The following are the ages and weights of a random sample of 10 high school male students <table><tr><td>Age</td><td>Weight</td></tr><tr><td>14</td><td>129</td></tr><tr><td>16</td><td>173</td></tr><tr><td>18</td><td>188</td></tr><tr><td>15</td><td>121</td></tr><tr><td>17</td><td>190</td></tr><tr><td>16</td><td>166</td></tr><tr><td>14</td><td>133</td></tr><tr><td>16</td><td>155</td></tr><tr><td>15</td><td>152</td></tr><tr><td>14</td><td>115</td></tr></table> Assuming a simple linear regression model, give an interval that, with 95 percent confidence, will contain the average weight of all 17 year old male high school students. b) How do you handle missing values in a dataset?	6	3 + 2 = 5
Q2	a) Write Python program to solve the below Fizz Buzz programming challenge. Print the numbers 1 to 100, except that if the number is divisible by 3, print "fizz"; if the number is divisible by 5, print "buzz"; and if the number is divisible by 15, print "fizzbuzz". b) Suppose in the simple linear regression model $Y = \alpha + \beta x + e$ that $0 < \beta < 1$. (a) Show that if $x < \alpha / (1 - \beta)$, then $x < E[Y] < \alpha / (1 - \beta)$ (b) Show that if $x > \alpha / (1 - \beta)$, then $x > E[Y] > \alpha / (1 - \beta)$ and conclude that $E[Y]$ is always between x and $\alpha/(1 - \beta)$.	2	2.5 + 2.5 = 5
Q3	The following data indicate the gain in reading speed versus the	6	5

number of weeks in the program of 10 students in a speed-reading program.

Number of Weeks	Speed Gain (wds/min)
2	21
3	42
8	102
11	130
4	52
5	57
9	105
7	85
5	62
7	90

(a) Plot a scatter diagram to see if a linear relationship is indicated.
(b) Find the least squares estimates of the regression coefficients.
(c) Estimate the expected gain of a student who plans to take the program for 7 weeks.

Q4	a) What relationship exists between row space, column space, and null space of the matrix? b) Find the eigenvector of the matrix. Row1=[3 0], row2 = [8 -1] c) What is the Data Science Life Cycle?	6	1 + 2 + 2 = 5
Q5	a) Compare each pair of distributions to decide which one has the greater mean and the greater standard deviation. You do not need to calculate the actual values of μ and σ, just how they compare with each other. 3, 5, 5, 5, 8, 11, 11, 11, 13 and 3, 5, 5, 5, 8, 11, 11, 11, 20. b) What is the test statistic used to test the significance of difference between small sample mean and population mean. c) Define Type I and Type II errors	5	2 + 2 + 1