

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT
TEST -2 EXAMINATION- 2025

B.Tech-I Semester (CSE/IT/ECE/CE/BT/BI)

COURSE CODE (CREDITS): 22M1WCI136 (3)

MAX. MARKS: 25

COURSE NAME: Data Visualization

COURSE INSTRUCTORS: Dr. Ramesh Narwal

MAX. TIME: 1 Hour 30 Min

Note: (a) All questions are compulsory.

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

Q.No	Question	CO	Marks
Q1	Given a dataset: $X = [1, 2, 3, 4, 5]$ $Y = [2, 3, 4, 6, 8]$ a) Perform simple linear regression to find the regression equation $Y=a+bX$ b) Calculate a (intercept) and b (slope) manually.	4	5
Q2	Explain the process of data wrangling and data aggregation. Give a numerical example using sales data (e.g., summing daily sales by month using pandas groupby()).	5	5
Q3	You are given two categorical variables (Gender: Male/Female, Preference: Yes/No). Explain how to use a Chi-square test to determine whether there is an association between gender and preference. State the steps and null hypothesis clearly.	2	5
Q4	Given the dataset: Marks = [10, 20, 20, 30, 40, 40, 50] a) Compute the mean, median, and mode of the dataset. b) Find the variance and standard deviation .	5	5
Q5	You are given a CSV file with 1,000 student records containing marks and attendance. a) Describe how you would clean this data in Python using pandas. b) Write steps to handle missing values, outliers, and inconsistent entries.	1	5