JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2025

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

COURSE CODE (CREDITS): 22M1WCI136 (3) MAX. MARKS: 35

COURSE NAME: Data Visualization

COURSE INSTRUCTORS: Dr. Ramesh Narwal MAX. TIME: 2 Hours

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	Discuss the purpose and application of different visualization charts like heatmaps, scatter plots, box plots, and time-series charts. Using the dataset: Sales: [100, 120, 130, 150, 180, 200] over months Jan–Jun Draw a time-series line chart and annotate the highest and lowest values.	2	7
Q2	Explain techniques used to process unstructured text data. Take the following text: "Data Science is exciting. Data scientists love data!" Perform: a) Tokenization b) Stop-word removal c) Stemming d) Lemmatization Compare the results.	3	7
Q3	Discuss how data cleansing and wrangling influence machine learning model accuracy. Given the following dataset: Age = [18, 19, None, 21, 200, 23, -5, 22] Apply cleansing by: a) Removing outliers b) Filling missing values with median c) Normalizing valid values (min-max normalization)	6	7
Q4	Evaluate the risks and responsibilities associated with deploying data science solutions, especially regarding bias, transparency, data privacy, and algorithmic accountability. Provide examples from domains such as hiring, policing, and healthcare.	1	7
Q5	Explain the importance of probability distributions in predictive analytics. Perform the following: a) Calculate the binomial probability of getting exactly 3 heads when a coin is tossed 8 times. b) Using Poisson distribution, compute the probability of 2 website failures in a day if the average failure rate is 0.6/day.b) Describe actions required under the IT Act 2000/IT Amendment Act 2008. c) Recommend at least three preventive security controls.	4,5	7