

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST-1 EXAMINATION- February, 2023

B.Tech. (CSE, IT) VI Semester

COURSE CODE: 18B1WCI635(2)

MAX. MARKS: 15

COURSE NAME: Data Mining and Data Warehousing

COURSE INSTRUCTORS: Dr. Ekta Gandotra

MAX. TIME: 1 Hour

Note: All questions are compulsory. Marks are indicated against each question in square brackets.

- Q1. a. Briefly outline how to compute the dissimilarity between objects described by binary attributes. [2] CO3
- b. For the following vectors, x and y , calculate the indicated similarity or distance measures. [3] CO3
- i. $x = (0, 1, 0, 1)$, $y = (1, 0, 1, 0)$ Cosine, Correlation, Jaccard
- ii. $x = (22, 1, 42, 10)$, $y = (20, 0, 36, 8)$ Euclidean, Manhattan, Supremum
- Q2. a. Draw a boxplot for the following data [2] CO1
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.
- b. Give any two methods to determine outliers in the data? [2] CO1
- Q3. In real-world data, tuples with missing values for some attributes are a common occurrence. Describe different methods for handling this problem with the help of suitable examples. [3] CO3
- Q4. Consider the following table presenting stock prices recorded at different times for Company-A and Company-B. Assume that the stocks are affected by the same industry trends. Using the covariance formula, determine whether their prices rise or fall together? [3] CO1

Time Point	Company-A	Company-B
TP-1	5	10
TP-2	6	20
TP-3	4	14
TP-4	3	5
TP-5	2	5