



## School of Computer Science and Engineering

Fall Semester 2023-2024

Continuous Assessment Test – 1

SLOT – F1+TF1

Programme Name & Branch : B.Tech. CSE Spec. in Information Security

Course Name & code: Data Privacy – BCSE318L

Class Number (s): VL2024250102051 & VL2024250102052

Faculty Name (s): Dr. Geetha Mary A & Dr. Senthil Kumar K

Exam Duration: 90 Min.

Maximum Marks: 50

| Q.No. | Question                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Max Marks | CO  | BL  |
|-------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|-----|-----|
| 1.    | <p>Discuss the various stakeholders in a bank based on the usage of different formats of data. Present the responsibilities and their contribution of data and its transformation.</p> <pre>graph TD     C[Customer/Record owner] --&gt; Co[Company—Banks, insurance, healthcare, or retail]     G[Government—Compliance and regulations] --&gt; Co     A[Adversary/Data snooper] --&gt; Co     Co --&gt; PD[(Production databases)]     PD --&gt; DA[Data anonymizer]     DA --&gt; AD[(Anonymized databases)]     AD --&gt; DA_A[Data analysts]     AD --&gt; T[Testers]     BOE[Business operations employee] --&gt; AD     BOE --&gt; PD</pre> | 10        | CO1 | BL1 |
| 2.    | <p>Present with relevant example about the various ways of measuring privacy and utility of anonymized data.</p> <ul style="list-style-type: none"><li>▪ Measure privacy : Statistical and Probabilistic methods</li><li>▪ Statistical method : Variance → larger the variance, the better the privacy of the perturbed data</li><li>▪ Probabilistic method : distribution of the data in the original data set and background information about some tuples in the data set.</li></ul>                                                                                                                                                            | 10        | CO1 | BL2 |

|       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |        |        |        |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
|-------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|--------|--------|-------|------|-------|-------|--------|-------|-------|--------|-------|-------|------|-------|-------|--------|-------|-------|------|-------|----|-----|-----|
|       | Measure Utility: Does the correlation between QI and SD fields in D' still exist? → Each anonymization function will provide different levels of distortion                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |        |        |        |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
| 3.    | <p>a. Consider the microdata provided below and calculate the global file level disclosure. [5]</p> <table><tr><td>Age</td><td>Gender</td><td>Salary</td></tr><tr><td>21-30</td><td>Male</td><td>50000</td></tr><tr><td>21-30</td><td>Female</td><td>60000</td></tr><tr><td>21-30</td><td>Female</td><td>55000</td></tr><tr><td>31-40</td><td>Male</td><td>60000</td></tr><tr><td>31-40</td><td>Female</td><td>70000</td></tr><tr><td>31-40</td><td>Male</td><td>65000</td></tr></table> <p>Solution: Global File level disclosure=<math>\frac{1}{1} + \frac{1}{2} + \frac{1}{1} + \frac{1}{2} = 3</math></p> <p>b. Present the various types of disclosures. [5]<br/>Answer:<br/>Identity disclosure takes place when a correct record re-identification operation is achieved by an intruder by comparing a target individual in a sample with an available list of units that contains individual identifiers such as name and address.</p> <p>Attribute disclosure takes place when sensitive information contained in the microdata about a specific individual unit is revealed through the release of the file.</p> <p>Inference disclosure takes place when one can determine the value of some characteristics of an individual from the released data.</p> | Age    | Gender | Salary | 21-30 | Male | 50000 | 21-30 | Female | 60000 | 21-30 | Female | 55000 | 31-40 | Male | 60000 | 31-40 | Female | 70000 | 31-40 | Male | 65000 | 10 | CO2 | BL3 |
| Age   | Gender                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Salary |        |        |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
| 21-30 | Male                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 50000  |        |        |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
| 21-30 | Female                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 60000  |        |        |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
| 21-30 | Female                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 55000  |        |        |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
| 31-40 | Male                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 60000  |        |        |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
| 31-40 | Female                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 70000  |        |        |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
| 31-40 | Male                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 65000  |        |        |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
| 4.    | <p>Identify the below mentioned non-perturbation technique and describe about it.</p> <p>a. Attribute value is updated from Male to Male or Female. [3]</p> <p>b. Attribute values of age in the range [0-100] is categorized as young, middle aged and Old. [4]</p> <p>c. Attribute values are replaced by ‘?’ [3]</p> <p>Solution:</p> <p>a. Global Recoding</p> <p>b. Top-Bottom Coding</p> <p>c. Suppression</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 10     | CO2    | BL2    |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |
| 5.    | Consider the dataset mentioned below.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 10     | CO2    | BL1    |       |      |       |       |        |       |       |        |       |       |      |       |       |        |       |       |      |       |    |     |     |

| PatientID | Gender | Age | Zip code | Test     |
|-----------|--------|-----|----------|----------|
| 55998     | M      | 19  | 15723    | Negative |
| 88557     | F      | 35  | 15674    | Positive |
| 55868     | F      | 35  | 15674    | Positive |
| 44551     | M      | 45  | 15623    | Negative |
| 58524     | M      | 45  | 15623    | Negative |
| 25584     | F      | 61  | 15633    | Negative |
| 58744     | F      | 61  | 15643    | Positive |

Discuss the different methods to protect the explicit identifier in the dataset.

Solution:

One way tokenization and two way tokenization