1. What is/are the correct statement/s?

    a) Pseudonymised data are governed by GDPR.
    b) "Visibility and Transparency" in privacy by design means informing developers about data handling practices of their products.
    c) "Data Privacy" is only about protecting personal data
    d) K-anonymised data are governed by GDPR.

    i.  a
    ii. a, b, d
    iii. a, c
    iv. a, d
    v. c, d

2. What is incorrect regarding PETs?

   a) Some PETs are not suitable for practical implementations.
   b) Synthetic data can be used to train machine learning models.
   c) Scalability of PETs is considered when selecting them for privacy protection.
   d) Applying PETs will prevent Ransomware attacks.

3. What is true regarding Pseudonymisation?

   a) Pseudonymisation always generate a mapping table.
   b) Counter is the only way to generate pseudonyms.
   c) Pseudonymisation is a naive PET
   d) Pseudonymisation removes the link between individuals and personal data.

4. What are the correct statements regarding the privatised table?

a) It satisfies 2-Anonymity.
b) It satisfies 3-Anonymity.
c) Mondrian algorithm can be used to achieve the row partitions [0,3], [1,6], [2,4,5].
d) Preprocessing data might be applied before achieving the given outcome.

i. a
ii. b
iii. a, c
iv. a, c, d
v. c, d

| Row index | Age | Zip | Gender | Disease |
|---|---|---|---|---|
| 0 | 21 – 30 | 2141 | F | Cancer |
| 3 | 21 – 30 | 2141 | M | Infection |
| 1 | 31 – 35 | 213* | F | AIDS |
| 6 | 31 – 35 | 213* | F | AIDS |
| 2 | 41 - 50 | * | M | Cancer |
| 4 | 41 - 50 | * | M | Infection |
| 5 | 41 - 50 | * | M | Infection |

5. **Query**: "What is the total number of people in the dataset".

   This query was executed on dataset D. The real answer is 1000. But when differential private results were generated for epsilon values 0.1, 0.01, 0.001 one possible noisy result that was achieved is 1004.2. Suppose X and Y are greater than 0. Match the possible number of times 1004.2 was generated under each epsilon when the query was executed multiple times.

| Epsilon |
| --- |
| 0.1 |
| 0.01 |
| 0.001 |

| Number of times 1004.2 was generated |
| --- |
| X |
| X - Y |
| X + Y |

6. What is/are the correct statement/s regarding Differential Privacy?

   a) Privacy guarantee of a differentially private result is lost by post-processing it
   b) It can be used to mitigate non-repudiation threat
   c) Differentially private query results are randomised
   d) By using it we try to reduce the distinguishability of an individual included in a datasets


   i.  b. c
   ii. a, d
   iii. a, b, c
   iv. c, d
   v. b, c, d

7. Tracking John….

   Download the GPS data of one week in John's life (weekdays) and the notebook that contains the code to analyse those GPS data.

   Execute the given code : Tracking John.ipynb

   Form a group and answer the questions in the notebook. Follow the instructions in the **comment** sections when answering the questions.

   **Q1** : Go to https://www.gpsvisualizer.com/ to visualise the generated csv

   **Q2** : Refer the instructions **at the end of the notebook** to visualise the generated csv files

   Note : If the CSV files are downloaded as "xlsx", change the file extensions to csv after downloading them.

8. Research is crucial because it drives innovation and advances knowledge, leading to improvements in technology, health, and quality of life. University health center decides to develop a system where they release university counselling sessions data with outside researchers. Researchers aim to analyse the spread of different types of psychological issues of university students in the UK. Answer the following questions to decide the suitable PETs to improve the privacy of this new system. Health center records Uni Id, email address, residential address and all the issues discussed during the counselling sessions.

a) Draw a rough diagram to show the personal data flow in the new system

b) Using the answer to a) can you identify 3 threats to privacy in the new system? Categorise the threats according to the LINDDUN threat types.

c) Propose PETs to mitigate the threats identified in b)

QUESTIONS???