COSC130 Fundamentals of Cybersecurity and Privacy
Tutorial Week 5

1. Suppose that a hospital has removed patients' names from the hospital records and intends to make these 'anonymised' patients' records in Table 1 available to a researcher. Suppose that the researcher has access to the external Table 2 and knows that every person with a record in Table 1 has a record in Table 2.
   a. Would this lead to record or attribute linkage of hospital patients?
   b. Which patients would have their privacy compromised?
   c. With what probability can an adversary infer that Betty has HIV?

Job	Sex	Age	Disease
Engineering	Female	31	Fracture
Scientist	Female	33	Flu
Scientist	Female	35	HIV
Lawyer	Female	32	Flu
Doctor	Female	31	Flu
Cricketer	Male	23	Fracture
Cricketer	Male	25	Fracture
Golfer	Male	20	HIV

Table 1

Name	Job	Sex	Age
Anne	Engineering	Female	31
Betty	Scientist	Female	33
Claire	Scientist	Female	35
Donna	Lawyer	Female	32
Anna	Doctor	Female	31
Bob	Cricketer	Male	23
Charlie	Cricketer	Male	25
Dennis	Golfer	Male	20
Peter	Doctor	Male	33
David	Lawyer	Male	32
Mark	Engineer	Male	24

Table 2

2. Consider Table 1 and Table 2.
   a. Identify the unique identifier ID and quasi identifier QID for both Table 1 and Table 2.
   b. Generate 3-anonymous tables from Table 1 and Table 2.
   c. What is the highest k you can achieve for each table?

3. Consider the anonymous data generated in Problem 2. Suppose the adversary knows that the target victim Betty is a scientist of age 33 and has a record in the Table 1. With what probability can an adversary infer that Betty has HIV? Compare this with the case when data was not k-anonymised.

4. Perform independent research to answer the following questions.
   a. What are data brokers?
   b. What is the difference between first-party data brokers and third-party data brokers?
   c. What are the main products and services offered by third-party data brokers in Australia?
   d. Identify at least 5 third-party data brokers operating in Australia.
   e. Identify at least 5 sources that third-party data brokers use to collect information.
   f. Identify at least 5 ways in which the collections and use of personal information by third-party data brokers can harm individuals.