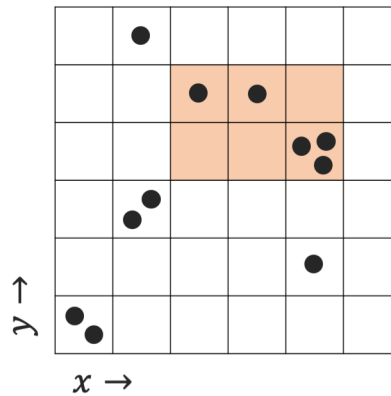


Problems with marked with an asterisk (\*) are more challenging or open-ended.

1. (Exercise 2.1 in the notes) Let  $F : [D]^n \rightarrow \mathbb{R}^{(D)}$  that takes a dataset  $\mathbf{x}$  and outputs a vector of answers containing  $f_{s,t}(\mathbf{x})$  for every  $1 \leq s \leq t \leq D$ . Prove that the global sensitivity of  $F$  is  $\Theta(D^2)$ .
2. (Exercise 2.4 in the notes) Let  $\mathcal{T}$  be the set of intervals in the binary tree mechanism with domain size  $D$  that is a power of 2. Prove that for every  $t$ , we can express the interval  $\{1, \dots, t\}$  as the union of at set of at most  $\log_2 D$  intervals in  $\mathcal{T}$ . For example,  $\{1, \dots, 11\} = \{1, \dots, 8\} \cup \{9, 10\} \cup \{11, 11\}$ .
3. In this question we'll generalize the ideas of the binary tree mechanism to answer *rectangle queries*. Here the data universe is the two-dimensional grid with side length  $D$ , and each datapoint is a pair  $(x_i, y_i) \in [D]^2$ . A rectangle query  $f_{s,t}^{u,v}$  is defined by ranges  $1 \leq s \leq t \leq D$  and  $1 \leq u \leq v \leq D$  and

$$f_{s,t}^{u,v}(\mathbf{x}) = \# \{i : s \leq x_i \leq t \text{ and } u \leq y_i \leq v\} \quad (1)$$

Here's an example depicting data in the domain  $[6]^2$ . Black dots are the data points (the position within the grid cells is irrelevant) and the orange shaded area represents the rectangle query  $f_{3,5}^{4,5}$ , whose answer on this dataset is 5.



- (a) How many rectangle queries are there? What is the global sensitivity of the set of all rectangle queries on the domain  $[D]^2$ ? If we use the Laplace mechanism to answer all such queries, how much noise do we add to each query?
- (b) Suppose we only want to answer the subset of queries called *corner-aligned rectangle queries*. This is the subset of rectangle queries that include the lower-left corner  $(1, 1)$ , and have the form  $f_{1,t}^{1,v}$ .

- i. Using the Laplace mechanism, how much noise would we add to answer all corner-aligned rectangle queries?
  - ii. How can you express any rectangle query  $f_{s,t}^{u,v}$  as a combination of a small number of corner-aligned rectangle queries?
  - iii. How much noise would we incur if we use the Laplace mechanism to answer all corner-aligned rectangle queries and then recover the answer to the other rectangle queries?
- (c) (\*) Can you generalize the binary-tree mechanism to answer all rectangle queries with error  $O(\frac{1}{\epsilon} \log^a D)$  for some constant exponent  $a$ ?
4. Consider the setting prefix-sum queries from Lectures 2 and 3 on reconstruction. The data set consists of an ordered list  $x_1, x_2, \dots, x_n$  in  $[0, 1]$ . Give a differentially private algorithm which answers all subset sum queries with expected additive error  $O(\frac{\log^3 n}{\epsilon})$  or better.