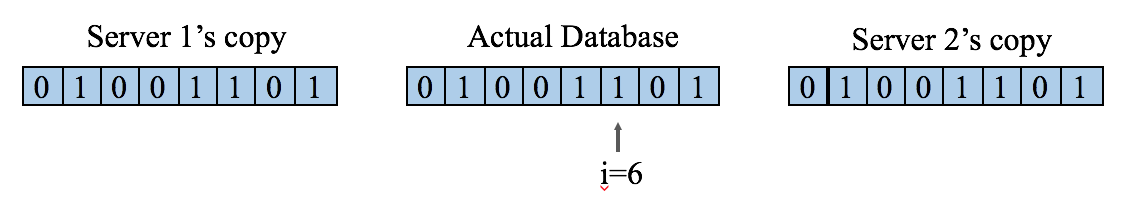
Top of Form

1. For the following table determine ---
   1. Quasi-identifiers and Sensitive identifiers. [2 points]
   2. Then construct a generalization lattice **indicating** nodes that achieve 3-anonymity and 3-diversity. [10 points]
   3. Finally, **write down** the actual 3-anonymous and 3-diverse table of your choice from the lattice. [3 points]

|  |  |  |  |
| --- | --- | --- | --- |
| ID | ZIP Code | Age | Disease |
| 1 | 35567 | 22 | colon cancer |
| 2 | 35502 | 31 | stomach cancer |
| 3 | 35560 | 24 | lung cancer |
| 4 | 35817 | 45 | stomach cancer |
| 5 | 35810 | 63 | diabetes |
| 6 | 35812 | 40 | aids |
| 7 | 35502 | 35 | aids |
| 8 | 35568 | 30 | flu |
| 9 | 35505 | 32 | lung cancer |

1. Write a filtering rule to block the following URL patterns.
   1. Block any non-image objects originating from *ads.com* [5 points]
   2. Block any script originating from *googleads.com,* while such requests are being made on any domain other than *google.com* [5 points]
   3. An **exception** rule that allows any script (previously blocked) to be loaded from *ad.com* [5 points]
2. D is the dataset containing annual salaries of all NC State employees. *bsdcount(D)* returns the number of entries in D that are greater than $1,000,000; *mode(D)* returns the **most frequent** salary in the dataset. Let, *San* be the standard Laplacian mechanism for ε-differential privacy. Given any function *f*, *San* generates random ξ from the Laplacian distribution with variance that depends on the sensitivity of function *f* and the privacy parameter ε, and returns *f*(D) + ξ.
3. What is the sensitivity of *bsdcount* and *max*? State all assumptions you needed to calculate the answers. [5 points]
4. For the “same level of privacy” (i.e., ε) which function requires “more noise” to be added and why? [5 points]
5. For the following 8-bit database show all the steps involved in retrieving the i-th (6-th) position bit using a **3-server** PIR (Private Information Retrieval) protocol. [10 points]



where random query strings for the first two servers are Q1 = < 0 1 0 1 0 0 1 1>, Q2 = < 1 0 0 1 1 0 0 1>

1. Explain the Tor circuit creation process. How does Tor ensure forward secrecy? [6+4 points]
2. Assume you are conducting a random response survey where you are trying to approximate - “What fraction of the population use drug?” To preserve user’s anonymity each participant is asked to answer using the following protocol:

* Flip a coin.
* If **heads**, then respond truthfully
* If **tails**, then respond untruthfully

1. Draw the response graph. [5 points]
2. If 60% of the observed participants answer ‘Yes’ to the question ‘Do you use drugs?’ then what is the actual fraction of the population using drugs? [5 points]
3. How are cookies spawned through HTML5 localstroage? How can you use ETags to track users? What is the basic difference between them? [4+4+2 points]