A bloom filter consists of:

    1.

    2.

Bloom Filter: Insertion

Example: **S = { 16, 8, 4, 13, 29, 11, 22 }, |S| = n**
              h(k) = k % 7, |Array| = N

[0]	
[1]	
[2]	
[3]	
[4]	
[5]	
[6]	

Bloom Filter: Deletion

Bloom filters do not support the deletion operation because:

Bloom Filter: Search

With h(k) = k % 7, find(16), find(20), find(3)

[0]	**0**
[1]	**1**
[2]	**1**
[3]	**0**
[4]	**1**
[5]	**0**
[6]	**1**

Probabilistic Accuracy

What are the four possible outcomes for a probabilistic function?

    1.

    2.

    3.

    4.

Which outcome can a bloom filter NOT have?

Bloom Filter: Repeated Trial Insertion

Example: **S = { 16, 8, 4, 13, 29, 11, 22 }**
h1(x) = x % 10, h2(x) = 2x % 10, h3(x) = (5+3x) % 10

[0]	
[1]	
[2]	
[3]	
[4]	
[5]	
[6]	
[7]	
[8]	
[9]	

Given k hashes and a bit vector of length m, what is the run time for insert and find:

False Positive Rate

What is the expected false positive rate in terms of k (hashes), m (bit vector length), and n (number of inserted items)?

Optimal Bloom Filters

What equation describes the relationship between m bits, n items, and k hashes for an optimal bloom filter?

Combining bloom filters

Given the following bloom filters, write a bloom filter which 'contains' all the items present in both filters.

[0]	0	[0]	0	[0]	
[1]	1	[1]	1	[1]	
[2]	1	[2]	0	[2]	
[3]	0	[3]	0	[3]	
[4]	1	[4]	1	[4]	
[5]	0	[5]	1	[5]	
[6]	0	[6]	0	[6]	

Given the same bit vectors, write a filter which 'contains' only items that are present in both filters.