Quiz 2

Name:

Suppose we have t hash functions h1, h2, $\cdots$, ht mapping from objects to b buckets {0, $\cdots$, b-1}. For j in [1, t], let sj be a hash function from objects to {+1, -1} and let cj be an array of counters of size b. While processing the stream, each time we encounter an item oi, update the counter cj[hj[oi]] += sj[oi].

Then, we estimate the count of oi with median$_{j=1,\cdots,t}$ of ~~cj[hj[oi] * sj[oi]]~~.

➔ cj[hj[oi]] * sj[oi]

Let

t = 5, b = 5

hash function h:

	o1	o2	o3	o4	o5	o6	o7
h1	3	1	1	4	1	0	2
h2	3	3	0	2	0	4	0
h3	4	1	4	2	4	2	2
h4	3	1	1	0	4	1	4
h5	1	0	4	0	3	3	2

hash functions s:

	o1	o2	o3	o4	o5	o6	o7
s1	1	1	-1	-1	-1	-1	1
s2	-1	-1	1	-1	-1	1	1
s3	-1	-1	1	1	-1	-1	-1
s4	1	-1	1	-1	-1	1	-1
s5	-1	1	1	-1	-1	1	-1

Data stream (the numbers below are the indexes i of object ID oi):

1	1	2	6	5	2	2	5	2	5	7	1	4	4

➔ Estimate the count of objects {o1, o2, o3}.

Input: 1

	0	1	2	3	4
c1	0	0	0	1	0
c2	0	0	0	-1	0
c3	0	0	0	0	-1
c4	0	0	0	1	0
c5	0	-1	0	0	0

Input: 1

	0	1	2	3	4
c1	0	0	0	2	0
c2	0	0	0	-2	0
c3	0	0	0	0	-2
c4	0	0	0	2	0
c5	0	-2	0	0	0

Input: 2

	0	1	2	3	4
c1	0	1	0	2	0
c2	0	0	0	-3	0
c3	0	-1	0	0	-2
c4	0	-1	0	2	0
c5	1	-2	0	0	0

Input: 6

	0	1	2	3	4
c1	-1	1	0	2	0
c2	0	0	0	-3	1
c3	0	-1	-1	0	-2
c4	0	0	0	2	0
c5	1	-2	0	1	0

Input: 5

	0	1	2	3	4
c1	-1	0	0	2	0
c2	-1	0	0	-3	1
c3	0	-1	-1	0	-3
c4	0	0	0	2	-1
c5	1	-2	0	0	0

Input: 2

	0	1	2	3	4
c1	-1	1	0	2	0
c2	-1	0	0	-4	1
c3	0	-2	-1	0	-3
c4	0	-1	0	2	-1
c5	2	-2	0	0	0

Input:2

	0	1	2	3	4
c1	-1	2	0	2	0
c2	-1	0	0	-5	1
c3	0	-3	-1	0	-3
c4	0	-2	0	2	-1
c5	3	-2	0	0	0

Input: 5

	0	1	2	3	4
c1	-1	1	0	2	0
c2	-2	0	0	-5	1
c3	0	-3	-1	0	-4
c4	0	-2	0	2	-2
c5	3	-2	0	-1	0

Input: 2

	0	1	2	3	4
c1	-1	2	0	2	0
c2	-2	0	0	-6	1
c3	0	-4	-1	0	-4
c4	0	-3	0	2	-2
c5	4	-2	0	-1	0

Input: 5

	0	1	2	3	4
c1	-1	1	0	2	0
c2	-3	0	0	-6	1
c3	0	-4	-1	0	-5
c4	0	-3	0	2	-3
c5	4	-2	0	-2	0

Input: 7

	0	1	2	3	4
c1	-1	1	1	2	0
c2	-2	0	0	-6	1
c3	0	-4	-2	0	-5
c4	0	-3	0	2	-4
c5	4	-2	-1	-2	0

Input: 1

	0	1	2	3	4
c1	-1	1	1	3	0
c2	-2	0	0	-7	1
c3	0	-4	-2	0	-6
c4	0	-3	0	3	-4
c5	4	-3	-1	-2	0

Input: 4

	0	1	2	3	4
c1	-1	1	1	3	-1
c2	-2	0	-1	-7	1
c3	0	-4	-1	0	-6
c4	-1	-3	0	3	-4
c5	3	-3	-1	-2	0

Input: 4

	0	1	2	3	4
c1	-1	1	1	3	-2
c2	-2	0	-2	-7	1
c3	0	-4	0	0	-6
c4	-2	-3	0	3	-4
c5	2	-3	-1	-2	0

o1 -> median of < 3, 7, -1, 0, 3 > = 3 (count = 3)

o2 -> median of < 1, 7, 4, 3, 2 > = 3 (count = 4)

o3 -> median of < -1, -2, -6, -3, 0 > = -3 (count = 0)