Assignment 1a

Problem Statement: Given a set of records $R_1, \ldots, R_n$ identified by primary keys $K_1, \ldots, K_n$ and a search key K, decide whether the record corresponding to key K exists or not.

Input: *n* random integers, K is another random integer where n = 10, 100, 1000, 10000, 100000.

Tasks:

1. Show the number of comparisons required to decide the result.
2. Find the time required for each of the instances.
3. Plot the graph for the inputs considering both the above cases where time is in Y axis and *n* is in X axis.
4. Write your observations, compare with the corresponding scenario of Assignment 1 and derive possible conclusions.

Algorithm: Do the following changes to your programs of Assignment 1

1. **Scenario 1:** The input records are not sorted based on the primary keys and each key has equal probability of getting searched.
   a. Now, append the search key K at the end of the array. $K_{n+1} \leftarrow K,\ i \leftarrow 1$
   b. If $K_i = K$, go to Step d.
   c. $i \leftarrow 1 + 1$, go to Step b
   d. If $i \leq n$, output "Yes". Otherwise, output "No" and Halt.

   Output: For the above program do the tasks for both when
   - The key is not present
   - Key is present at some random location

2. **Scenario 2:** Each key K_i, $1 \leq i \leq n - 1$, has a probability $p_i = \frac{1}{2^i}$ of getting searched. The input records are sorted in descending order based on the search probability. That is, the maximum searched element is put first, and so on.
   a. Store probability for each index positions of the primary key array
   b. Create large streams of search keys maintaining this probability

   Output:
   - Find the average number of comparisons for successful search when the stream of search keys generated is atleast five times of larger than the number of records in the input.
   - Repeat this experiment 10 times and note down the differences in observation, if any.