Hashing

1. Hashing is a technique that is used to uniquely identify a specific object from a group of similar objects. It is also used for storing, retrieving & removing information as quick as possible.

2. It is also a technique to convert a range of key values into a range of indexes of an array.

3. It is a method to allocate a unique code for any object.

4. It is a searching technique used to store & retrieve the info in $O(1)$ time complexity i.e less than other searching techniques such linear search which as $O(n)$ & binary search which has $(\log n)$ time complexity.

5. It is also a process of converting a arbitary size key into a fixed size value. The conversion is done via special function called as hash function.

Hashtable:

→ Hashtable is a data structure which stores the data in an associative manner. In a hash table, data is stored in an array format, where each data value has its own unique index value. Access of data becomes very fast if we know the index of desired data; Thus it becomes the data structure in which insertion & search operations are very fast irrespective of size of data.

→ It uses array as a storage medium & uses hash technique to generate an index where an element is to be inserted (or) is to be located from.

→ A hash table is made up of two parts:

i) An array (the actual table where the data to be searched is stored)

2) Mapping function (Hash function)

→ It uses a hash function to compute an index into an array in which an element will be inserted or searched.

→ It is also a data structure which organises data using hash functions. In order to support quick insertion & search.

There are different kinds of hash tables.

1. The hashmap is one of the implementations of map structure to store (key, value) pairs.

2. The hashset is one of the implementations of set structure to store no repeated values.

Hash function

→ The hash function is a mapping from the input space to the integer space that defines the indices of the array, it provides a way for assigning number to the input data such that the data can be stored at the array index corresponding to the assigned number.

Input + hashfunction = hashvalue (or) hashcode.

→ It is a function that can be used to map a data set of an arbitrary size to a dataset of fixed size which falls into the hashtable.

→ Values returned by hash function are called hash values / codes / sums.

To achieve a good hashing mechanism, it is important to have good hash function.

1. Easy to compute
2. Uniform distribution

3. Less collisions
4. Fast
5. Hash function should point to same index in hash table for the same value always.
6. It should generate values only from $[0, n-1)$ n is size of hash table.
7. Hash function should be unbiased & should uniformly distribute keys to indexes in hashtable.