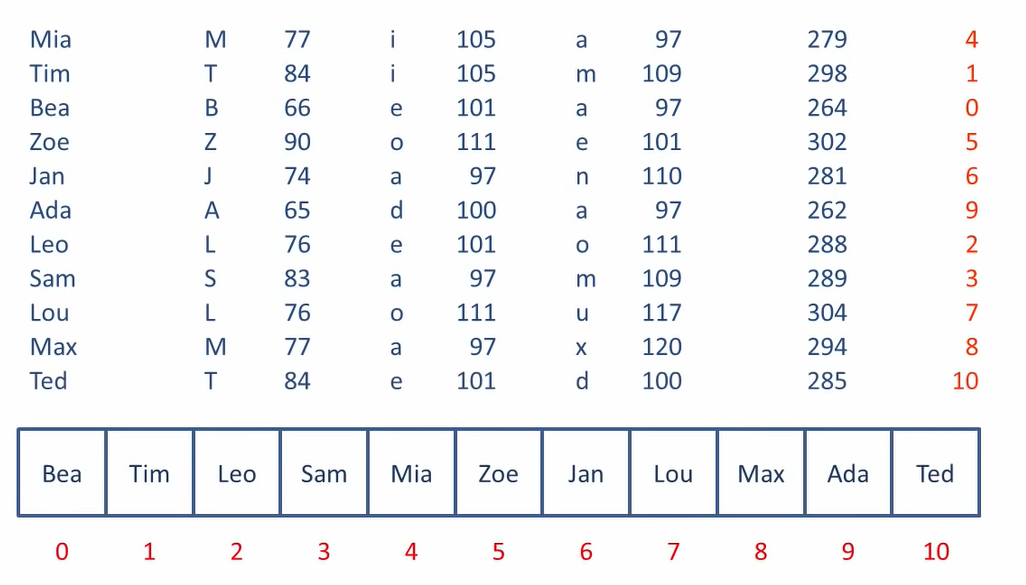
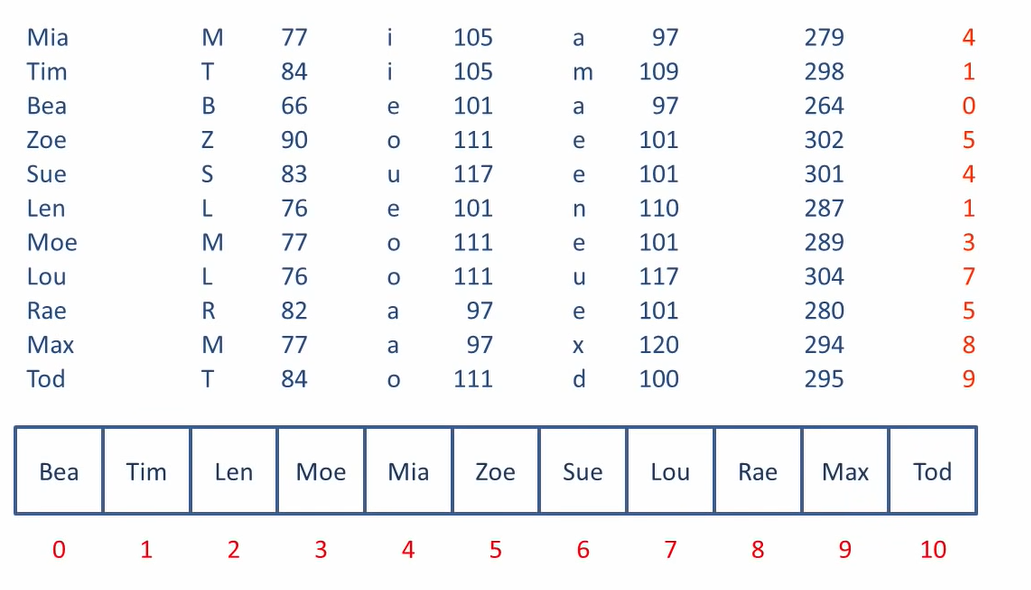
Bảng băm (Hash table):

* Dùng để đánh dấu chỉ số (index) cho một lượng lớn data
* Key (hay index) được tính bằng chính nó
* Ví dụ: Các tên được tính mã ascii của nó và chia mod cho tổng số phần tử cần index (11 phần tử - 259 mod n ) => KEY mod N

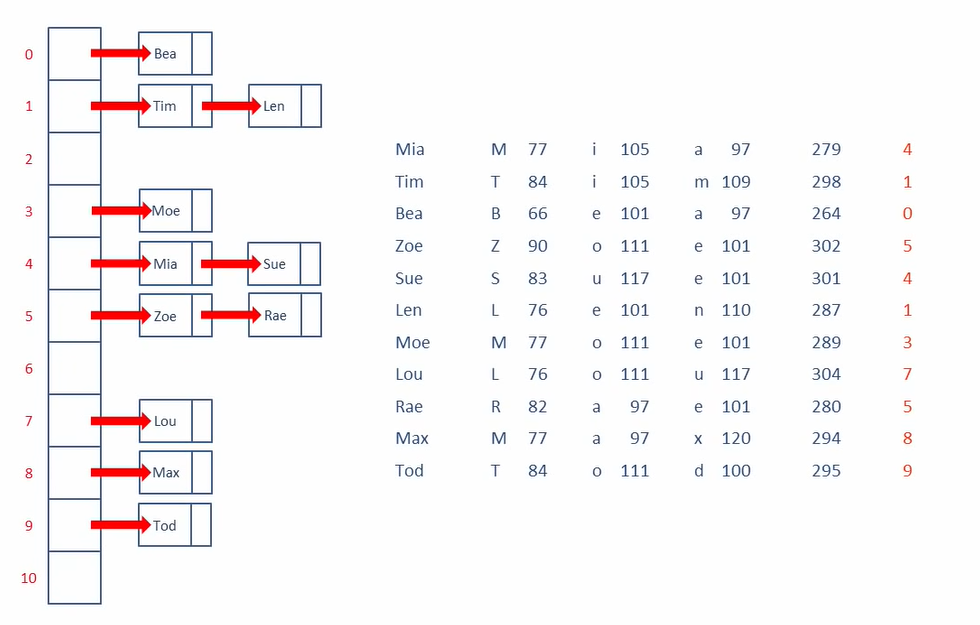


* Collisions thường được giải quyết bằng open addressing (linear search) hoặc closed addressing (linked list, etc)
* Collision là khi một value có index trùng với index khác.



+ Phía trên dùng cách linear probing, về cơ bản là traverse sang phải (hoặc toàn bộ mảng nếu worst case) để tìm vị trí chưa occupied => Open addressing, thường dùng cho các lượng data nhỏ

+ Hoặc có thể dùng chaining (Linked List, etc) để quản lý các index trùng nhau. Cách này được dùng cho lượng data lớn, nếu lượng data là nhỏ thì Open Addressing nhanh hơn và tối ưu hơn



* Open Addressing: Có thể dùng các cách sau cho linear probing

+ Linear probing

+ Plus 3 rehash (+3)

+ Quadratic probing (failed attempts)2

+ Duble hashing (Dùng thêm 1 bảng băm nửa để xem những vị trí collision còn trống tiếp theo)

* Closed Addressing: Chaining (Linked List hoặc Data structures khác)
* Hashing thường được dùng rộng rãi trong việc đánh dấu index trong database, compilers, caching, password authentification, etc
* Về mặt lý thuyết, có thể tạo ra một hash table hoàn hảo cho việc index các value