

## Observaciones laboratorio #6

No tengo muy claro si con índices se refieren a cada llave, o a los datos de cada llave, lo haré con la segunda opción.

Primera parte:

a) Teniendo en cuenta cada uno de los requerimientos ¿Cuántos índices implementaría en el Reto? y ¿Por qué?

Respuesta teniendo en cuenta el archivo small:

En caso de los Medios utilizaría cerca de 500 datos para poder guardar todos sin necesidad de hacer rehash.

En caso de los Nacionalidades utilizaría cerca de 200 datos que son casi la cantidad de países que existen.

Y en caso de llaves que son únicas como el nombre de los artistas o los id's utilizaría la cantidad de datos del archivo.

b) Según los índices propuestos ¿en qué caso usaría Linear Probing o Separate Chaining en estos índices? y ¿Por qué?

Para las llaves que se repiten mucho como la nacionalidad y/o los medios usaría Separate Chaining para no tener problemas con las llaves repetidas a la hora de guardar los datos.

Para las llaves que son únicas utilizaría Linear Probing para facilitar la búsqueda de los valores de cada dato.

c) Dado el número de elementos de los archivos MoMA, ¿Cuál sería el factor de carga para estos índices según su mecanismo de colisión?

En caso de utilizar los Separate Chaining utilizaría un factor de carga por encima de 1 para que se puedan guardar la mayor cantidad de datos.

En caso de utilizar Linear Probing utilizaría un factor de carga de 0.7 para guardar todos los datos posibles antes de hacer un rehash

Segunda parte:

Tiempos de ejecución con archivo small:

Se probó con la función hecha sobre las nacionalidades pasando por parámetro “German” que tiene 80 valores.

Separate Chaining:

2.0 31.25

4.0 46.875

8.0 31.25

Linear Probing:

0.2 15.625

0.5 15.625

0.8 15.625

Conclusiones:

Para esta cantidad de datos se demora un poco menos el Linear Probing y también noto que Linear tiene menos variabilidad de tiempos así se cambie el factor de carga.