

EDA – Preguntas lab 6

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

Req1 – El mejor índice para este requerimiento es por el año de nacimiento de los artistas debido a que se puede ir uno por uno en el rango y se encuentran directamente.

Req2 – En este caso el agrupar por fechas es poco eficiente debido a que es poco común que varias obras compartan una misma fecha específica, se podría utilizar solo el año para no tener una cantidad excesiva de mapas individuales

Req3 - La mejor opción de inicio es el medio de la obra ya que se pueden agrupar varias obras en un solo mapa y encontrar una cantidad específica de obras del artista

Req4 - en este caso es muy útil el clasificar los artistas por ID debido a que se tienen que recorrer las obras y esta es la forma más fácil de acceder a sus artistas

Req5 - Aquí se pueden organizar las obras con índice de su departamento para encontrar las obras a mover más fácil

Req6 - se necesita cambiar un poco la carga de datos, pero si se añade un valor de cuantas obras tiene se puede organizar por este criterio el mapa

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

Linear – R2, R4, R6 – como la llave varía demasiado entre elementos es más fácil adoptar esta medida

Separate – R1, R3, R5 - En este caso nos importa recorrer los elementos que cumplan cierta condición rápidamente, por eso el separate chaining es más efectivo.

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

En los lineal buscamos un factor de 0.7 lo que equivale a $15224 / 21751$ (primo de 21748)

Y en los sepárate buscamos un factor de 3 (o hasta más dependiendo del caso) lo que equivale a $15224 / 5077$ (primo de 5074)