

## SESION 1: ARCHIVOS

### Ejercicio (Importante: Tamaño de archivos, ordenación, organización de archivos, estimación de tuplas de una consulta, coste de buscar en una consulta sencilla)

Se dispone de un archivo  $r$  de datos que contiene información sobre estudiantes donde se almacena:

- Número de carnet: 8 bytes, todos los valores del campo son diferentes. El campo es un serial que comienza por valor 1.
- Nombre del alumno: 40 bytes
- Código de carrera: 2 bytes con valores incluidos (IN) (0,1,2,3,4,5,6,7,8,9)
- Edad: 8 bytes
- Índice académico: 2 bytes

El número de registros a almacenar es de 100.000 y cada bloque tiene un tamaño de 512 bytes. El grado de ocupación de cada bloque es del 65%. La longitud de puntero a bloque es de 6 bytes y la longitud de puntero a registro de 7 bytes. Cada bloque tiene 12 bytes de control. Se pide para cada una de las situaciones que se muestran a continuación lo siguiente:

1. Para un archivo que se organiza en forma de montículo (sin ordenación) calcular el tamaño de bloques. Para cada una de las siguientes consultas determinar el número de registros estimados y el coste de lectura.

a.  $\sigma_{\text{carnet}=2345}(r)$

b.  $\sigma_{\text{código\_carrera}=2}(r)$

2. Para un archivo que se organiza en forma secuencial (ordenado) por el campo carnet calcular el tamaño de bloques. Para cada una de las siguientes consultas determinar el número de registros estimados y el coste de lectura.

a.  $\sigma_{\text{carnet}=2345}(r)$

b.  $\sigma_{\text{código\_carrera}=2}(r)$

3. Para un archivo que se organiza en forma secuencial (ordenado) por el campo código de carrera, se ha añadido un puntero de 12 bytes a cada registro que sirve para formar la lista enlazada que marque la ordenación lógica del fichero según ese campo. Calcular el tamaño de bloques del archivo. Para cada una de las siguientes consultas determinar el número de registros estimados y el coste de lectura.

a.  $\sigma_{\text{carnet}=2345}(r)$

b.  $\sigma_{\text{código\_carrera}=2}(r)$

4. Para un archivo que se organiza en forma de árbol B+ por el campo carnet calcular el tamaño de bloques. Para cada una de las siguientes consultas determinar el número de registros estimados y el coste de lectura.

a.  $\sigma_{\text{carnet}=2345}(r)$

b.  $\sigma_{\text{código\_carrera}=2}(r)$

5. Para un archivo que se organiza por medio de asociación (hash) con función de asociación **carnet mod 100**, calcular el tamaño de bloques. Para cada una de las siguientes consultas determinar el número de registros estimados y el coste de lectura.

a.  $\sigma_{\text{carnet}=2345}(r)$

b.  $\sigma_{\text{código\_carrera}=2}(r)$

6. Cuestiones opcionales para resolver por los alumnos. Considerando los tipos de archivos del ejercicio anterior, determinar el número de registros a recuperar y el coste de buscar para cada una de las siguientes consultas.

a.  $\sigma_{\text{carnet} \geq 10.000}(r)$

b.  $\sigma_{\text{carnet} < 10.000}(r)$

c.  $\sigma_{\text{carnet} < > 10.000}(r)$

d.  $\sigma_{\text{carnet} > 10.000 \wedge \text{carnet} < 15.000}(r)$

e.  $\sigma_{\text{código\_carrera} > 6}(r)$

f.  $\sigma_{\text{código\_carrera} < 5}(r)$

g.  $\sigma_{\text{código\_carrera} < > 8}(r)$

h.  $\sigma_{\text{código\_carrera} > 7 \wedge \text{código\_carrera} \leq 9}(r)$

i.  $\sigma_{\text{carnet} = 53456 \wedge \text{código\_carrera} = 9}(r)$

j.  $\sigma_{\text{carnet} = 53456 \vee \text{código\_carrera} = 9}(r)$