

# Svolgimento Esercizi Costi Database

## Esercizio 1: Tabella VISITA

### Query SQL

```
SELECT COUNT(*)
FROM VISITA
WHERE DATA = '2024-05-10' AND CF = '...'
```

### Statistiche e Parametri

- $N_T = 3000$
- $N_B = 300$

### Indici Disponibili

#### 1. Indice su CF (Unclustered)

- $N_K = 1000$
- $N_F = 150$
- Fattore di selettività:  $F_{CF} = \frac{1}{N_K} = \frac{1}{1000}$

#### 2. Indice su DATA (Unclustered)

- $N_K = 2000$
- $N_F = 100$
- Fattore di selettività:  $F_{DATA} = \frac{1}{N_K} = \frac{1}{2000}$

### Calcolo dei Costi

#### 1. Scansione Sequenziale ( $C_{seq}$ )

$$C_{seq} = N_B = 300$$

#### 2. Accesso tramite Indice su CF ( $C_{CF}$ )

La formula per l'indice unclustered è:  
 $[F \cdot N_F] + [F \cdot N_T]$ .

$$C_{CF} = \left\lceil \frac{1}{1000} \cdot 156 \right\rceil + \left\lceil \frac{1}{1000} \cdot 3000 \right\rceil$$
$$C_{CF} = \lceil 0.156 \rceil + \lceil 3 \rceil = 1 + 3 = 4$$

### 3. Accesso tramite Indice su DATA ( $C_{DATA}$ )

$$C_{DATA} = \left\lceil \frac{1}{2000} \cdot 100 \right\rceil + \left\lceil \frac{1}{2000} \cdot 3000 \right\rceil$$

$$C_{DATA} = \lceil 0.05 \rceil + \lceil 1.5 \rceil = 1 + 2 = 3$$

**Conclusione:** L'accesso più efficiente è tramite l'indice su **DATA** (Costo = 3).

## Esercizio 2: Tabella PROFESSIONISTA

### Query SQL

```
SELECT ID_CENTRO, STIPENDIO  
FROM PROFESSIONISTA  
WHERE ID_CENTRO = 1  
ORDER BY ID_CENTRO, STIPENDIO NETTO
```

### Statistiche e Parametri

- $N_T = 1000$
- $N_B = 100$
- $C_{seq} = N_B = 100$  (Costo scansione sequenziale)

### Indici Disponibili

- Indice su ID\_CENTRO (Unclustered)
  - $N_K = 50$
  - $N_F = 10$
  - Fattore di selettività:  $F_{ID\_CENTRO} = \frac{1}{50}$

### Calcolo dei Costi

Accesso tramite Indice su ID\_CENTRO ( $C_{ID\_CENTRO}$ ) La formula utilizzata è:  
 $\lceil F \cdot N_F \rceil + \lceil F \cdot N_T \rceil$ .

$$C_{ID\_CENTRO} = \left\lceil \frac{1}{50} \cdot 10 \right\rceil + \left\lceil \frac{1}{50} \cdot 1000 \right\rceil$$
$$C_{ID\_CENTRO} = \lceil 0.2 \rceil + \lceil 20 \rceil = 1 + 20 = 21$$

### Conclusione

Confrontando i costi:

$$21 < 100 \quad (C_{Indice} < C_{seq})$$

Pertanto, l'utilizzo dell'indice **CONVIENE**.