

Exercício 2.3

- Dados desde 2001
- Milhões de pacientes e atestados
- Milhares de médicos e doenças
- Todas as tabelas tem índice primário pelo respectivo código
- Todas as tabelas tem índice auxiliar por cada outro atributo mencionado na consulta.

SELECT M.CRM, M.nome, COUNT(A.*) AS nAtests,
COUNT(DISTINCT P.codP)

FROM Atestado A, Medico M, Paciente P,
Doenca D

WHERE A.CRM = M.CRM AND A.codP = P.codP
AND A.codD = D.codD AND A.data >= '01/01/12'
AND D.espec <> M.espec AND
M.estado = 'SC' AND P.estado = 'SC'

GROUP BY M.CRM, M.nome

HAVING nAtests > (SELECT COUNT(*) /
COUNT(DISTINCT CRM) FROM Atestado)

ORDER BY nAtests DESCENDING

* As projeções da árvore foram omitidas por espaço, mas assume-se que foram feitas nos nó-folha.

ORDER BY
nAtests DESCENDING (8)

| K_F

π M.CRM, M.nome, nAtests, nPac (7)

K_{ADP}

\bowtie A.codP = P.codP (6)

K_{AD}

σ M.espec (>) D.espec
| K_{AD}

\bowtie A.codD = D.codD (5)

K_{MA}

\bowtie M.CRM = A.CRM (4)

COMUTAT.
DE σ e \bowtie

K_M

K_A

σ M.estado
'SC' (1)

MEDICO

σ A.data
'01/01/12' (2)

ATESTADO

DOENCA

σ P.estado
'SC' (3)

PACIENTE

COUNT
(Distinct P.codP)
AS nPac (11)

CASCATA
DE SELEÇÃO

Atestado

COUNT (A.*)
AS nAtests (9)

COUNT
(DISTINCT
CRM)
AS distCRM (10)

nAtests/
distCRM
AS mediz

● Busca por índice $\rightarrow O(\log |N|)$

(1) $O(\log |M|)$

(2) $O(\log |A|)$

(3) $O(\log |P|)$

● Join por igualdade

(Hash join) $\rightarrow O(|R| + |S|)$

(4) $O(|K_M| + |K_A|)$

(5) $O(|K_{MA}| + |D|)$

(6) $O(|K_P| + |K_{AP}|)$

● Algoritmo semelhante ao hash join $\rightarrow O(|N|)$

(7) $O(|K_{ADP}|)$

● Merge Sort $\rightarrow O(|R| \cdot \log |R|)$

(8) $O(|K_F| \cdot \log |K_F|)$

● Busca exaustiva $\rightarrow O(|N|)$

(9) $\rightarrow O(|A|)$

(10) \rightarrow

(11) $O(|P|)$

Complexidade total

$$- (K_A \approx K_P) > K_M \gg K_{MA} \gg K_{AD} \gg K_{ADP} \gg K_F$$

$$- (|A| \approx |D|) > (|M| \approx |D|)$$

$$O(|A| + |D| + \log |M| + \log |D| + K_M + K_P + (K_F \cdot \log |K_F|))$$