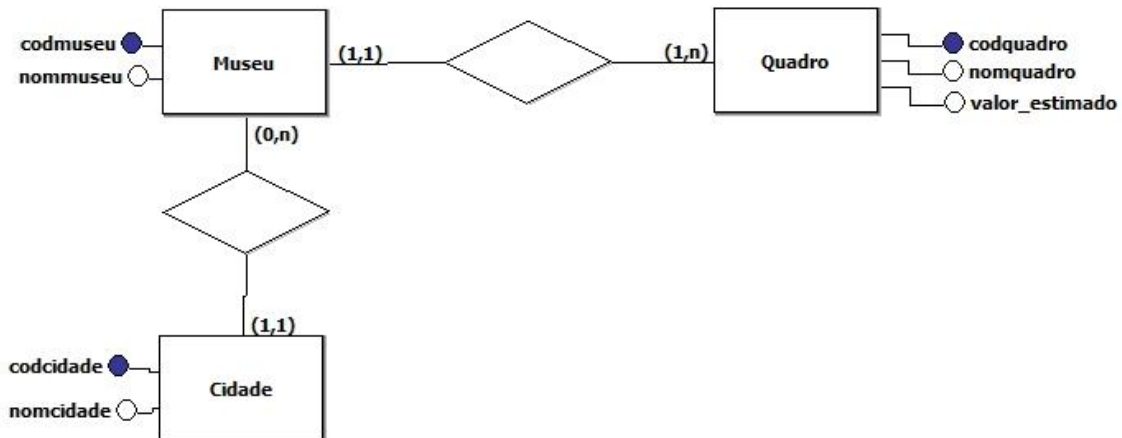


Gabarito tarefa 3 – último exercício

Considere o modelo ER a seguir, as tabelas geradas a partir do modelo



CIDADE (CODCIDADE, NOMCIDADE)

MUSEU (CODMUSEU, NOMMUSEU, CODCIDADE)
CODCIDADE REFERENCIA CIDADE

QUADRO (CODQUADRO, NOMQUADRO, CODMUSEU, VALOR_ESTIMADO)
CODMUSEU REFERENCIA MUSEU

Considere a consulta apresentada na sequência (em SQL e Álgebra).

Consulta

<pre>SELECT NOMMUSEU FROM QUADRO, MUSEU WHERE QUADRO.CODMUSEU = MUSEU.CODMUSEU AND VALOR_ESTIMADO=100.000,00;</pre>	<pre>π NOMMUSEU (σ VALOR_ESTIMADO=10.00,00 ^ QUADRO.CODMUSEU = MUSEU.CODMUSEU (QUADRO X MUSEU))</pre>
---	--

Levando em conta as estatísticas sobre a base de dados mostradas a seguir (ver estatísticas na folha 5 da prova) e a consulta
 σ VALOR_ESTIMADO=10.000,00 (QUADRO).

Calcule e descreva o cálculo do custo (isto é, o número de acessos a disco) para executar esta consulta, para cada uma das possibilidades apresentadas a seguir.

4.1) Custo para executar a consulta varrendo toda tabela QUADRO

Custo: 200 acessos a disco

4.2) Custo para executar a consulta usando o índice para a coluna VALOR_ESTIMADO da tabela QUADRO

Existem 500 valores distintos para VALOR_ESTIMADO da tabela QUADRO.

Como não existe a informação sobre distribuição de frequência (frequência absoluta - quantas vezes cada valor aparece), é assumido que a distribuição dos dados é uniforme. Assim, considerando as 4.000 linhas da tabela quadro, pode-se estimar que existem 8 linhas onde o VALOR_ESTIMADO da tabela QUADRO é igual a 10.000,00 (4.000 linhas da tabela quadro/500 valores distintos = 8 linhas).

No pior caso, cada uma destas 8 linhas está em um bloco de disco (8 acessos a blocos de disco).

Considerando que existe um índice para VALOR_ESTIMADO e que o nível é 1 (raiz + filhos) deverá ser feito acesso ao bloco com o nó raiz e mais um acesso ao bloco do nós da árvore onde estão os ponteiros para os blocos de disco com valor_estimado = 10.000,00. Assim para o índice, são dois acessos a disco.

O número de acessos é 8 (blocos com os dados) + 2 (blocos de árvore) o custo para executar a consulta σ VALOR_ESTIMADO=10.000,00 (QUADRO) é de 10 acessos a disco. Pode-se considerar no curso ainda a tabela temporária (1 acesso a disco) o que totalizaria um curso de 11 acessos.

Neste caso, vale a pena usar o índice e não varrer sequencialmente toda tabela (isto é, todos 200 blocos).

Estatísticas sobre a base de dados

(Obs. Leve em conta que o otimizador assume uma distribuição uniforme dos dados)

Informações sobre as colunas

NOME_TABELA	NOME_COLUNA	VALORES DISTINTOS	MENOR VALOR	MAIOR VALOR
QUADRO	VALOR_ESTIMADO	500	10.000,00	700.000,00
QUADRO	CODQUADRO	4.000	1	4.000
QUADRO	NOMQUADRO	3.800		
QUADRO	CODMUSEU	100	1	100
MUSEU	CODMUSEU	100	1	100
MUSEU	NOMMUSEU	100		
MUSEU	CODCIDADE	500	1	500
CIDADE	CODCIDADE	500	1	500
CIDADE	NOMCIDADE	500	1	500

Informações sobre as tabelas

NOME_TABELA	NÚMERO_LINHAS	BLOCOS DO DISCO OCUPADOS
QUADRO	4.000	200
MUSEU	100	10
CIDADE	500	15

Informações sobre os índices

NOME_DO_ÍNDICE	ÚNICO	NÍVEL B	CHAVES_DISTINTAS
QUADRO_CODQUADRO	SIM	1	4.000
QUADRO_VALOR_ESTIMADO	NÃO	1	500

NÍVEL B: é o número de níveis da árvore excluindo a raiz