

Nome do aluno: _____ Matrícula do aluno: _____

1) (1,5 pontos) Suponha que se deseja fazer uma **seleção por igualdade** sobre um atributo, e existe um **índice secundário** sobre esse atributo. Duas das possíveis estratégias de busca são a **busca linear** e o **uso de índice**. Suponha que a tabela ocupe **1.000 blocos** e os registros que atendem à seleção estão espalhados em **N blocos**. Indique a partir de **qual valor de N** passa a **valer a pena** usar a estratégia de **busca linear**. O cálculo do custo deve levar em consideração a soma dos tempos de **seek** e **transferência**, sendo que um **seek** é **50 vezes** mais lento que uma transferência. Considere que a árvore do índice tenha altura igual a **quatro**, e **nenhum dos nós** esteja na memória.

2) (1,5 pontos) Suponha que se deseja usar a **ordenação externa por intercalação** para **ordenar** um conjunto de registros divididos em **1024 partições** com **dois registros** cada. A intercalação utiliza um buffer de entrada com capacidade para **N partições** de entrada. Qual deve ser o tamanho **mínimo** de **N** para que a ordenação seja concluída sem que sejam necessárias mais do que **cinco passadas**.

3) (1,5 pontos) Considere que se deseja fazer uma **junção** entre as tabelas A e B. A tabela A ocupa **200 blocos** e a tabela B ocupa **500 blocos**. Duas estratégias de junção possíveis são **block nested loop join** e **indexed nested loop join**. Supondo que a tabela A seja usada no **nível externo**, indique para qual **número de registros de A** passa a valer a pena usar **block nested loop join**. Considere o custo como o tempo necessário para realizar **os seeks e transferências**, sendo que um **seek** é **50 vezes** mais lento que uma transferência e que **nenhum nó do índice** sobre B esteja na memória. Ainda, **nenhuma das tabelas** cabe na memória e a altura do índice é 4.

Considere o esquema relacional da Figura 1 e as estatísticas da Figura 2 para responder as questões de 4 até 7.

4) (1,0 pontos) Marque V para as consultas que precisam se preocupar com remoção de duplicatas e F caso contrário? Para acertar a questão, **todas** marcações precisam estar corretas. Responda na **folha de respostas**.

- a) Select distinct idFunc from func natural join depto
- b) Select distinct idDepto from func natural join depto
- c) Select idFunc from func natural join depto
- d) Select idDepto from func natural join depto
- e) Select distinct idFunc, idDepto from func natural join depto

5) (1,5 pontos) Em álgebra relacional, a expressão $\Pi_L (E1 \cap E2)$ e a expressão $(\Pi_L E1) \cap (\Pi_L E2)$ não são equivalentes. Demonstre um caso que prove essa afirmação usando consultas SQL sobre a tabela **func**. Para auxiliar na explicação, mostre os **registros da tabela**, bem como os **registros retornados pelas consultas**.

6) (1 ponto) O processamento de consultas compostas por subconsultas pode ser valer de uma técnica conhecida como **des correlação**, onde **tabelas temporárias** são criadas para que o resultado possa ser encontrado mais rapidamente. Ao aplicar essa técnica, qual **script SQL** seria criado para responder à consulta abaixo? Considere como parte do script tanto a **criação** quanto o **uso** da tabela temporária.

```
SELECT * FROM func
WHERE salario IN (
    SELECT custo
    FROM projeto )
```

7) (2 pontos) Considere a consulta abaixo:

```
SELECT *
FROM func NATURAL JOIN aloc NATURAL JOIN proj
WHERE salario = 2.000 AND custo = 4.000
```

Com base nas estatísticas disponíveis, estime quantos registros seriam retornados por essa consulta.

depto (<u>idDepto</u> , nomeDepto, setor)
func (<u>idFunc</u> , idDepto, nomeFunc, salario)
idDepto referencia depto
proj (<u>idProj</u> , nomeProj, custo, duracao)
aloc (<u>idFunc</u> , idProj, <u>função</u>)
idFunc referencia func
idProj referencia proj

Figura 1 – Esquema Relacional

Chave	Valor
N. registros func	1.000
N. registros aloc	10.000
N. registros proj	50
V(salario, func)	5
V(custo, proj)	5

Figura 2 – Estatísticas das tabelas