

Universidade Federal da Paraíba – Campus I

Centro de Informática

Departamento de Informática

Big Data: conceitos e aplicações

Lab 01: Revisão: Modelo Relacional

Atividade 0) Acessar a conta no serviço <https://www.elephantsql.com/> (usuário: raonikulesza@gmail.com, senha: bigdata24), crie uma instancia do banco e execute o arquivo da base de dados “empresa” (ver no SIGAA no tópico "Laboratório Introdução ao Modelo Relacional “ o script "EMPRESA.SQL") no SQL Browser da aplicação:

The screenshot shows the ElephantSQL SQL Browser interface. On the left is a sidebar with navigation options: DETAILS, ALARMS, BROWSER (selected), STATS, SLOW QUERIES, BACKUP, LOG, METRICS, ADMIN, INTEGRATIONS, FIREWALL, and MAINTENANCE. The main area is titled 'SQL Browser' and shows the table structure for 'EMPREGADO'. The structure is defined by a CREATE TABLE statement with columns: ssn (char(9) NOT NULL, PRIMARY KEY), pnome (varchar(30) NOT NULL), inicialm (char(1) default NULL), unome (varchar(30) NOT NULL), sexo (char(1) NOT NULL), endereco (varchar(50) default NULL), datanasc (date default NULL), superssn (char(9) default NULL), dno (char(1) NOT NULL), and salario (float default NULL). Below the structure, there is a section for 'Records of EMPREGADO' with an INSERT INTO statement showing sample data.

Atividade 1) Restrições de Integridade. Considere o estado do banco de dados "empresa" apresentado na figura a seguir:

Empregado

ssn	pnome	inicialm	unome	sexo	endereco	datanasc	superssn	dno	salario
123456789	Joao	B	Souza	M	R. 24 de maio, 1500 - Curitiba - PR	1965-01-09	333445555	5	30000
333445555	Fabio	T	Will	M	R. Alagoas, 325 - Curitiba - PR	1955-12-08	888665555	5	40000
999887777	Alice	J	Zebra	F	R. Ribeira, 98 - Pinhais - PR	1968-01-19	987654321	4	25000
987654321	Jennifer	S	Wallace	F	R. Berry, 291 - Colombo - PR	1941-06-20	888665555	4	43000
666884444	Ricardo	K	Nantes	M	Av. Figueira, 55 - Almirante Tamandare - PR	1962-09-15	333445555	5	38000
453453453	Jussara	A	Pereira	F	Trav. da Lapa, 23 - Curitiba - PR	1972-07-31	333445555	5	25000
987987987	Alberto	V	Medeiros	M	R. Ceará, 1245 - Curitiba - PR	1969-03-29	987654321	4	25000
888665555	Joaquim	E	Brito	M	R. Stone, 450 - Curitiba - PR	1937-11-10	(Null)	1	55000

Dependente

essn	nomedep	sexodep	datanascdep	parentesco
333445555	Alice	F	1986-04-05	FILHA
333445555	Teodoro	M	1983-10-25	FILHO
333445555	Joana	F	1958-05-03	CONJUGE
987654321	Abdala	M	1942-02-28	CONJUGE
123456789	Michel	M	1988-01-04	FILHO
123456789	Alice	F	1988-12-30	FILHA
123456789	Elizabete	F	1967-05-05	CONJUGE

Localização

dnum	dlocalizacao
1	Curitiba
4	Colombo
5	Araucaria
5	Curitiba
5	Pinhais

Departamento

dnumero	dnome	gerssn	gerdatainicio
5	Pesquisa	333445555	1988-05-22
4	Administracao	987654321	1995-01-01
1	Sede Administrativa	888665555	1981-06-19

Projeto

pnumero	pnome	plocal	dnum
1	ProdutoX	Pinhais	5
2	ProdutoY	Araucaria	5
3	ProdutoZ	Curitiba	5
10	Automatizacao	Pinhais	4
20	Reorganizacao	Curitiba	1
30	Novos Beneficios	Pinhais	4

Trabalha

essn	pno	horas
123456789	1	32,5
123456789	2	7,5
666884444	3	40
453453453	1	20
453453453	2	20
333445555	2	10
333445555	3	10
333445555	10	10
333445555	20	10
999887777	30	30
999887777	10	10
987987987	10	35
987987987	30	5
987654321	30	20
987654321	20	15
888665555	20	(Null)

Analisar todas as restrições de integridade violadas em cada operação, se houver alguma, e os diferentes mecanismos para impor essas restrições. Escreva o comando DML adequado (mesmo que 'erros' sejam gerados em função da falta de restrição).

- a) Insira >'943775543', 'Roberto', 'F', 'Silva', 'M', 'Rua X, 22 – Araucária – PR', '1952-06-21', '888665555', '1', 58000< em EMPREGADO.
- b) Insira >'4', ProdutoA', 'Araucaria', '2'< em PROJETO.
- c) Insira >'4', 'Produção', '943775543', '1998-10-01'< em DEPARTAMENTO.
- d) Insira >'677678989', null, 40.0< em TRABALHA.
- e) Insira >'453453453', 'Joao', 'M', '1970-12-12', 'CONJUGE'< em DEPENDENTE.
- f) Remova as tuplas de TRABALHA com ESSN = '333445555'.
- g) Remova a tupla de EMPREGADO com SSN = '987654321'.
- h) Remova a tupla de PROJETO com PJNOME = 'ProdutoX'
- i) Modifique o GERSSN e GERDATAINICIO da tupla de DEPARTAMENTO com DNUMERO = 5 para '123456789' e '1999-01-10', respectivamente.
- j) Modifique o atributo SUPERSSN da tupla EMPREGADO com SSN = '999887777' para '943775543'.
- l) Modifique o atributo HORAS da tupla de TRABALHA com ESSN = '999887777' e PNO = '10' para 5,0.

Atividade 2) Comandos DML: Considerando o banco de dados "empresa" apresentado na **Atividade 1**, escreva o comando DML adequado que:

- a) Mostre o número de segurança social do empregado, o nome do dependente e o parentesco, ordenado por ssn ascendente e parentesco descendente.
- b) Mostre o nome e endereço de todos os empregados que trabalham para o departamento Pesquisa.
- c) Para todo projeto localizado em Araucaria, liste o nome do projeto, o nome do departamento de controle e o último nome, endereço e data de nascimento do gerente do departamento.
- d) Recupere os nomes de todos os empregados que trabalhem mais de 10 horas por semana no projeto Automatizacao.
- e) Mostre o nome dos empregados que têm os três maiores salários.
- f) Mostre o nome e o salário dos supervisores com aumento de 20%.
- g) Mostre os departamentos e suas localidades.
- h) Mostre os departamentos e seus projetos.
- i) Mostre os empregados do sexo feminino e a data de nascimento cujo salário é maior que 30.000.
- j) Mostre os projetos em que o empregado 'Fábio' trabalha.

Atividade 3) Joins: Dadas as tabelas A e B e seus registros (abaixo):

TABELA_A	
CODIGO	NOME
1	UM
2	DOIS
3	TRES
4	QUATRO
5	CINCO

TABELA_B		
ID	CODIGO	VALOR
1	1	1.000
2	1	2.000
3	1	5.000
4	2	4.000
5	2	9.000
6	3	7.000
7	5	4.000
8	8	7.000

Mostre como ficariam as pesquisas (SELECT's) com os JOIN's (INNER, [LEFT| RIGHT | FULL] OUTER). Para tais pesquisas vamos usar a seguinte linguagem:

```
SELECT [CAMPOS]
FROM "TABELA_A"
[INNER] JOIN | {LEFT | RIGHT | FULL } [OUTER]} JOIN "TABELA_B"
```

Observação: Para a relação entre as tabelas, tem-se:

- 3 registros para a empresa 1 (que existe na tabela A);
- 2 registros para a empresa 2 (que existe na tabela A);
- 1 registros para a empresa 3 (que existe na tabela A);
- 0 registros para a empresa 4 (que existe na tabela A);
- 1 registros para a empresa 5 (que existe na tabela A);
- 1 registros para a empresa 8 (que NÃO existe na tabela A)

Atividade 4) Operações CRUD (Complementar)

Suponha que cada uma das seguintes operações de atualização seja aplicada diretamente ao estado do banco de dados "empresa", apresentado na **Atividade 1**. Escreva o comando DML adequado que:

- a) Encontre todos os empregados cujo supervisor esteja alocado em um departamento diferente. Mostre o número de segurança social e o nome completo do empregado.
- b) Mostre o número de segurança social do empregado, o nome do dependente o e parentesco, ordenado por ssn ascendente e parentesco descendente.
- c) Encontre os nomes de todos os empregados que são diretamente supervisionados por "Joaquim E Brito".

d) Faça uma lista dos números e nomes de projetos que envolvam um empregado cujo último nome seja 'Will', mesmo que esse trabalhador seja o gerente do departamento que controla o projeto.

e) Encontre o nome dos empregados que trabalham em algum dos projetos controlados pelo departamento número 5.

f) Encontre o nome e o endereço de todos os empregados que trabalhem em pelo menos um projeto localizado em 'Curitiba', mas cujo departamento não se localiza em 'Curitiba'.)

Atividade 5) Funções agregadas e nativas - Utilize funções agregadas e/ou nativas SQL para responder as seguintes questões:

a) Recupere a média salarial de todos os empregados do sexo feminino.

b) Mostre o número de empregados por supervisor.

c) Mostre o maior número de horas envolvido em projetos.

d) Para cada projeto, liste o nome do projeto e o total de horas por semana (de todos os empregados) gastas no projeto.

e) Para cada departamento, recupere o nome do departamento e a média salarial de todos os empregados que trabalham nesse departamento.

f) Liste os nomes de todos os empregados com dois ou mais dependentes.

g) Mostre o nome do departamento que possui o menor número de projetos associados.

h) Escreva uma consulta que retorne do 10o ao 22o caractere do endereço do empregado.

i) Escreva uma consulta que retorne apenas o mês de nascimento de cada funcionário.

j) Escreva uma consulta que retorne a idade (em anos) que o empregado tinha quando o dependente de parentesco filhou ou filha nasceu.

k) Escreva uma consulta que conte o número de dependentes por ano de nascimento.

l) Escreva uma consulta que mostre o nome de empregados supervisor que tenham 2 ou mais supervisionados.

m) Escreva uma consulta que mostre o valor mensal a ser pago por projeto (considere que a coluna 'salário' de empregado é mensal).

Atividade 6) – Subconsultas - Escreva subconsultas SQL para as questões:

a) Recupere nome (pnome e unome) de cada um dos empregados que tenham um dependente cujo primeiro nome e sexo sejam o mesmo do empregado em questão.

b) Recupere os nomes dos empregados (pnome e unome) cujos salários são maiores que a média dos salários dos empregados do departamento 5.

c) Retorne o número do seguro social (SSN) de todos os empregados que trabalham com a mesma combinação (projeto, horas) em algum dos projetos em que o empregado 'Fabio Will' (SSN= 333445555) trabalhe.

- d) Recupere os nomes de todos os empregados que não trabalham em nenhum projeto.
- e) Recupere o nome de empregados que não tenham dependentes.
- f) Liste o último nome de todos os gerentes de departamento que não tenham dependentes.
- g) Liste os nomes dos gerentes que tenham, pelo menos, um dependente.

Atividade 7) Índices - Escreva comandos SQL para responder as seguintes questões:

- a) Crie a tabela tempEmpregado contendo dos dados dos atributos ssn, pnome, sexo, endereco e datanasc da tabela empregado. Realizar a copia dos dados durante o processo de criação da nova tabela.
- b) Utilize o comando EXPLAIN ANALYSE e anote o tempo de execução da consulta `SELECT * FROM TEMPEMPREGADO`.
- c) Crie um índice com o campo ssn que não permite dados duplicados para a tabela tempEmpregado, utilizando o algoritmo de ordenação BTREE.
- d) Crie um índice com o campo pnome para a tabela tempEmpregado, em ordem decrescente e que utilize o algoritmo de ordenação HASH.
- e) Utilize o comando EXPLAIN ANALYSE, anote e compare com o tempo de execução da consulta `SELECT * FROM TEMPEMPREGADO`. O que achou do resultado?
- f) Remova os índices criados nos itens c e d.
- g) Remova a tabela temporária criada no item a.

Atividade 8) Visões - Escreva comandos SQL para responder as seguintes questões

- a) Crie a visão chamada TRABALHA_EM que deverá conter os campos pnome e unome da tabela empregado, o campo pjnome da tabela projeto e o campo horas da tabela trabalha
- b) Crie uma consulta SQL na visão implementada no item a que retorne o último e o primeiro nome de todos os empregados que trabalham no 'ProdutoX'.
- c) Exclua a visão criada no item a.
- d) Crie uma visão chamada DEPTO_INFO que deverá conter os campos dnome da tabela departamento, e o total de empregados e somatório dos salários dos empregados da tabela empregado por departamento
- e) Crie uma consulta SQL na visão implementada no item c que retorne a lista de informações por departamentos ordenados pelo somatório dos salários.
- f) Exclua as visões criadas nos itens a.