Trabalho Prático - ELC119 Fundamentos de Banco de Dados

Isabella Sakis e Rhauani Fazul —

Fonte de dados abertos



http://dados.gov.br/dataset/mec-prouni

Bolsas concedidas e perfil dos beneficiários do ProUni - 2016

Detalhamento das bolsas concedidas pelo Prouni por ano, contendo a região; UF; município; instituição de educação superior; curso; modalidade de ensino (presencial ou EAD); turno e tipo de bolsa. Detalhamento do perfil dos beneficiários do Prouni por sexo; raça/cor; data de nascimento e pessoas com deficiência.

Arquivo CSV

ProUni							
Nome	Campo	Tipo	Descrição				
Ano da concessão da bolsa	ANO_CONCESSAO_BOLSA	Numérico	Ano da concessão da bolsa ProUni (início da vigência).				
Código do e-MEC da IES que concedeu a bolsa	CODIGO_EMEC_IES_BOLSA	Numérico	Código do e-MEC referente a IES que concedeu a bolsa ProUni.				
Nome da IES	NOME_IES_BOLSA	Alfanumérico	Nome/Razão Social da Instituição de Ensino Superior que concedeu a bolsa ProUni.				
Tipo da Bolsa	TIPO_BOLSA	Alfanumérico	Descrição do tipo da bolsa concedida ao beneficiário do ProUni (integral – 100% ou parcial – 50%).				
Modalidade de ensino	MODALIDADE_ENSINO_BOLSA	Alfanumérico	Descrição da modalidade de ensino da bolsa concedida ao beneficiário do ProUni (presencial ou ensino à distância – EAD).				
Nome do Curso	NOME_CURSO_BOLSA	Alfanumérico	Nome do curso do beneficiário da bolsa ProUni.				
Turno do Curso	NOME_TURNO_CURSO_BOLSA	Alfanumérico	Descrição do turno do curso do beneficiário da bolsa ProUni.				
CPF do beneficiário	CPF_BENEFICIARIO_BOLSA	Alfanumérico	CPF do beneficiário da bolsa ProUni (verificar regras de concessão e sobreposição de bolsas abaixo).				
Sexo do beneficiário	SEXO_BENEFICIARIO_BOLSA	Alfanumérico	Sexo informado pelo beneficiário da bolsa ProUni.				
Raça/Cor	RACA_BENEFICIARIO_BOLSA	Alfanumérico	Raça/Cor informado pelo beneficiário da bolsa ProUni.				
Data de nascimento do beneficiário	DT_NASCIMENTO_BENEFICIARIO	Data	Data de nascimento do beneficiário da bolsa ProUni.				
Indicação se o beneficiário é portador de deficiência	BENEFICIARIO_DEFICIENTE_FISICO	Alfanumérico	Indicação se o beneficiário da bolsa ProUni é portador de algum tipo de deficiência (sim ou não).				
Região	REGIAO_BENEFICIARIO_BOLSA	Alfanumérico	Nome da região de residência do beneficiário da bolsa ProUni.				
UF	SIGLA_UF_BENEFICIARIO_BOLSA	Alfanumérico	Sigla da UF de residência do beneficiário da bolsa ProUni.				
Município	MUNICIPIO_BENEFICIARIO_BOLSA	Alfanumérico	Nome do Município de residência do beneficiário da bolsa ProUni.				

1. Criação de um banco de dados desnormalizado

```
DROP DATABASE IF EXISTS prouni2016;
      CREATE DATABASE IF NOT EXISTS prouni2016 DEFAULT CHARACTER SET = utf8;
      USE prouni2016:
    CREATE TABLE IF NOT EXISTS bolsa desnormalizado
 6
          codigo registro bolsa
                                            INT
                                                          NOT NULL AUTO INCREMENT.
          ano concessao bolsa
                                                         NOT NULL.
          codigo emec ies bolsa
                                                         NOT NULL.
          nome ies bolsa
                                            VARCHAR (100)
10
          tipo bolsa
                                            VARCHAR (30)
                                                         NOT NULL.
          modalidade ensino bolsa
                                            VARCHAR (30)
          nome curso bolsa
                                            VARCHAR (100)
                                                         NOT NULL.
13
          nome turno curso bolsa
                                            VARCHAR (20)
                                                         NOT NULL.
14
          cpf beneficiario bolsa
                                            VARCHAR (20)
                                                         NOT NULL.
15
          sexo beneficiario bolsa
                                            VARCHAR (15)
16
          raca beneficiario bolsa
                                            VARCHAR (15)
                                                         NOT NULL.
                                                          NOT NULL,
          dt nascimento beneficiario
                                            DATE
18
          beneficiario deficiente fisico
                                            VARCHAR(3)
                                                             NULL.
19
          regiao beneficiario bolsa
                                            VARCHAR (15)
                                                             NULL,
20
          sigla uf beneficiario bolsa
                                            VARCHAR(2)
                                                         NOT NULL,
21
          municipio beneficiario bolsa
                                            VARCHAR (60)
                                                         NOT NULL.
          PRIMARY KEY (codigo registro bolsa)
23
        ENGINE = InnoDB DEFAULT CHARACTER SET = utf8;
```



2. Normalização do esquema desnormalizado

PRIMEIRA FORMA NORMAL (1FN)

" diz-se que uma tabela está na primeira forma normal, quando ela não contém tabelas aninhadas"

bolsa_desnormalizado (codigo_registro_bolsa, ano_concessao_bolsa, codigo_emec_ies_bolsa, nome_ies_bolsa, tipo_bolsa, modalidade_ensino_bolsa, nome_curso_bolsa, nome_turno_curso_bolsa, cpf_beneficiario_bolsa, sexo_beneficiario_bolsa, raca_beneficiario_bolsa, dt_nascimento_beneficiario, beneficiario_deficiente_fisico, regiao_beneficiario_bolsa, sigla_uf_beneficiario_bolsa, municipio_beneficiario_bolsa)

Já está na 1FN!

SEGUNDA FORMA NORMAL (2FN)

" uma tabela encontra-se na segunda forma normal, quando, além de estar na 1FN, não contém dependências parciais*"

* Uma dependência (funcional) parcial ocorre quando uma coluna depende apenas de **parte** de uma **chave primária composta**.

Passagem à 2FN: Tabela com uma coluna na chave primária

1FN

bolsa_desnormalizado (**codigo_registro_bolsa**, ano_concessao_bolsa, codigo_emec_ies_bolsa, nome_ies_bolsa, tipo_bolsa, modalidade_ensino_bolsa, nome_curso_bolsa, nome_turno_curso_bolsa, cpf_beneficiario_bolsa, sexo_beneficiario_bolsa, raca_beneficiario_bolsa, dt_nascimento_beneficiario, beneficiario_deficiente_fisico, regiao_beneficiario_bolsa, sigla_uf_beneficiario_bolsa, municipio_beneficiario_bolsa)

2FN

bolsa_desnormalizado (**codigo_registro_bolsa**, ano_concessao_bolsa, codigo_emec_ies_bolsa, nome_ies_bolsa, tipo_bolsa, modalidade_ensino_bolsa, nome_curso_bolsa, nome_turno_curso_bolsa, cpf_beneficiario_bolsa, sexo_beneficiario_bolsa, raca_beneficiario_bolsa, dt_nascimento_beneficiario, beneficiario_deficiente_fisico, regiao_beneficiario_bolsa, sigla_uf_beneficiario_bolsa, municipio_beneficiario_bolsa)

Já está na 2FN!

TERCEIRA FORMA NORMAL (3FN)

" uma tabela encontra-se na terceira forma normal, quando, além de estar na 2FN, não contém dependências transitivas*"

^{*} Uma dependência (funcional) transitiva ocorre quando uma coluna, além de depender da chave primária da tabela, depende de outra coluna ou conjunto de colunas da tabela.

Passagem à 3FN

bolsa_desnormalizado (codigo_registro_bolsa, ano_concessao_bolsa, codigo_emec_ies_bolsa, nome_ies_bolsa, tipo_bolsa, modalidade_ensino_bolsa, nome_curso_bolsa, nome_turno_curso_bolsa, cpf_beneficiario_bolsa, sexo_beneficiario_bolsa, raca_beneficiario_bolsa, dt_nascimento_beneficiario, beneficiario_deficiente_fisico, regiao_beneficiario_bolsa, sigla_uf_beneficiario_bolsa, municipio_beneficiario_bolsa)

bolsa_prouni (<u>codigo_bolsa_prouni</u>, ano_concessao, tipo_bolsa, <u>codigo_curso</u>, <u>codigo_beneficiario</u>)

codigo_curso
codigo_beneficiario
REFERENCIA curso
REFERENCIA beneficiario

curso (codigo_curso, codigo_emec_ies, nome_ies_bolsa, modalidade_ensino, nome_curso, turno)

beneficiario (codigo_beneficiario, cpf, sexo, raca, dt_nascimento, deficiente_fisico, regiao, sigla_uf, municipio)

curso (codigo_curso, codigo_emec_ies, nome_ies_bolsa, modalidade_ensino, nome_curso, turno)



curso (codigo_curso, codigo_emec_ies, modalidade_ensino, nome_curso, turno) codigo_emec_ies REFERENCIA info_ies

Info_ies (codigo_emec_ies, nome_ies_bolsa)

beneficiario (codigo_beneficiario, cpf, sexo, raca, dt_nascimento, deficiente_fisico, regiao, sigla_uf, municipio)



beneficiario (codigo_beneficiario, cpf, sexo, raca, dt_nascimento, deficiente_fisico, codigo_municipio) codigo_municipio REFERENCIA municipio

municipio (codigo_municipio, nome_municipio, regiao, sigla_uf)

municipio (codigo_municipio, nome_municipio, regiao, sigla_uf)

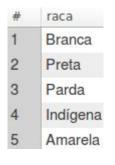


municipio (codigo_municipio, nome_municipio, cofigo_uf)

codigo_uf REFERENCIA info_uf

Info_uf (codigo_uf, regiao, sigla)

Aprimoramentos



anaficiario (godigo banaficiario enf. cava re

beneficiario (<u>codigo_beneficiario</u>, cpf, sexo, <u>raca</u>, dt_nascimento, deficiente_fisico, <u>codigo_municipio</u>) codigo_municipio REFERENCIA municipio



beneficiario (<u>codigo_beneficiario</u>, cpf, sexo, <u>codigo_raca</u>, dt_nascimento, deficiente_fisico, <u>codigo_municipio</u>)

codigo_racaREFERENCIA raçacodigo_municipioREFERENCIA municipio

raça (codigo raca, raca)

tipo_bolsa
BOLSA PARCIAL 50%
BOLSA INTEGRAL

Novos tipos? Modificações de nome?

bolsa_prouni (codigo_bolsa_prouni, ano_concessao, tipo_bolsa, codigo_curso, codigo_beneficiario)

codigo_curso REFERENCIA curso

codigo_beneficiario REFERENCIA beneficiario



bolsa_prouni (codigo_bolsa_prouni, ano_concessao, codigo_tipo, codigo_curso, codigo_beneficiario)

codigo_tipo
REFERENCIA tipo_bolsa

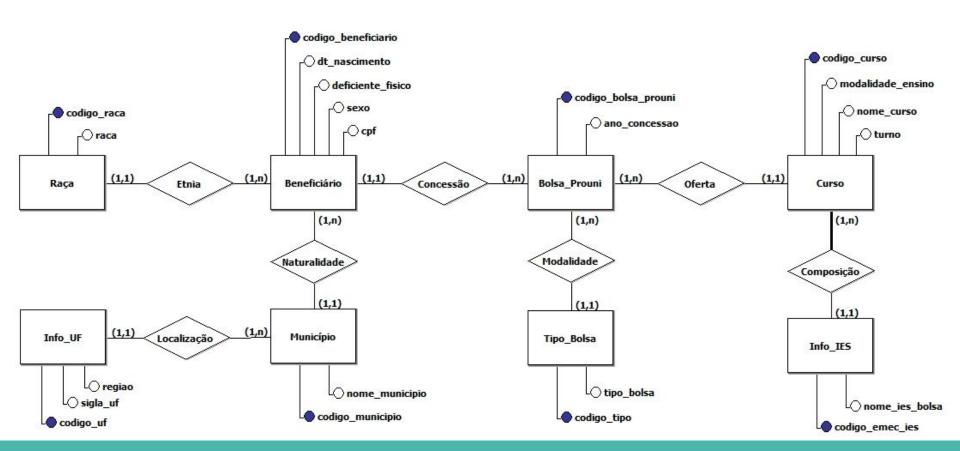
codigo_curso REFERENCIA curso

codigo_beneficiario REFERENCIA beneficiario

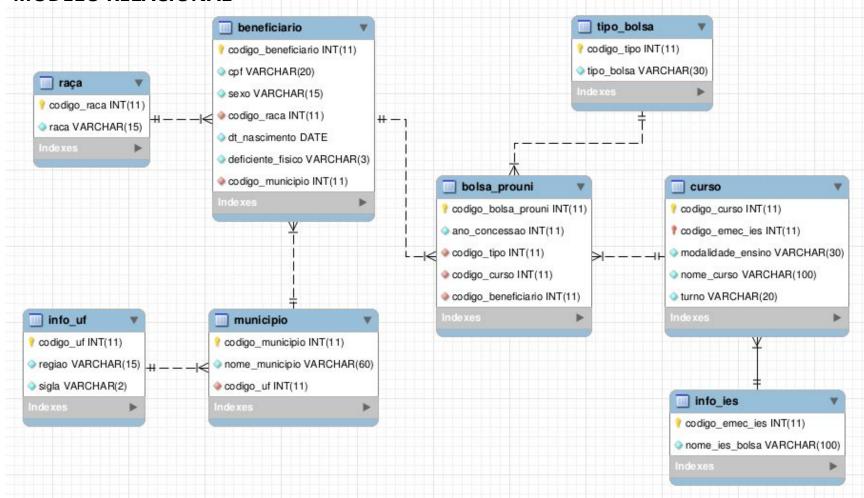
tipo_bolsa (codigo_tipo, tipo_bolsa)

3. Modelagem conceitual do esquema normalizado

3.1 MODELO ER



3.2 MODELO RELACIONAL



4. Carga dos dados para o esquema normalizado

a) Extrator para carregar os dados da fonte para a tabela desnormalizada

```
USE prouni2016;
       LOAD DATA LOCAL INFILE '~/Desktop/amostra pda prouni 2016.csv'
       INTO TABLE bolsa desnormalizado
       CHARACTER SET utf8
       FIELDS TERMINATED BY ';'
       OPTIONALLY ENCLOSED BY
       LINES TERMINATED BY '\n'
      IGNORE 1 LINES
     □ (
           ANO CONCESSAO BOLSA,
           CODIGO EMEC IES BOLSA,
13
           NOME IES BOLSA, TIPO BOLSA,
14
           MODALIDADE ENSINO BOLSA,
15
           NOME CURSO BOLSA.
16
           NOME TURNO CURSO BOLSA,
17
           CPF BENEFICIARIO BOLSA,
18
           SEXO BENEFICIARIO BOLSA,
19
           RACA BENEFICIARIO BOLSA,
20
           @DT NASCIMENTO BENEFICIARIO,
21
           BENEFICIARIO DEFICIENTE FISICO,
22
           REGIAO BENEFICIARIO BOLSA,
23
           SIGLA UF BENEFICIARIO BOLSA.
24
           MUNICIPIO BENEFICIARIO BOLSA
25
       SET dt nascimento beneficiario = STR TO DATE(@DT NASCIMENTO BENEFICIARIO, '%d-%m-%Y');
       -- dt nascimento beneficiario = coluna tipo date da tabela prouniTabelao
27
       -- SET dt nascimento beneficiario = STR TO DATE(@DT NASCIMENTO BENEFICIARIO, '%d-%m-%Y');
       -- beneficiario deficiente fisico = IF (@BENEFICIARIO DEFICIENTE FISICO = 'SIM', 'S', 'N');
```

CARGA 1

b) Usar um script SQL para que os dados da tabela desnormalizada sejam copiados para as respectivas tabelas normalizadas

```
-- INFO IES
                      info ies (codigo emec ies, nome ies bolsa)
      INSERT INTO
      SELECT DISTINCT codigo emec ies bolsa, nome ies bolsa
                       bolsa desnormalizado:
 6
      FROM
      -- CURSO
                      curso (modalidade ensino, codigo emec ies, nome curso, turno)
9 .
      INSERT INTO
      SELECT DISTINCT modalidade ensino bolsa, codigo emec ies bolsa, nome curso bolsa, nome turno curso bolsa
10
                       bolsa desnormalizado;
11
      FROM
12
13
      -- INFO UF
14 .
                       info uf (regiao, sigla)
      INSERT INTO
      SELECT DISTINCT regiao beneficiario bolsa, sigla uf beneficiario bolsa
15
                       bolsa desnormalizado;
16
      FROM
17
18
      -- MUNICIPIO
                      municipio (nome municipio, codigo uf)
19 •
      INSERT INTO
      SELECT DISTINCT t1.municipio_beneficiario_bolsa, t2.codigo_uf
20
                       bolsa desnormalizado AS t1
21
      FROM
22
      INNER JOIN
                      info uf
                                            AS t2 ON t2.sigla = t1.sigla uf beneficiario bolsa;
23
24
      -- RACA
25 .
      INSERT INTO
                       'raca' (raca)
      SELECT DISTINCT raca beneficiario bolsa
26
27
                       bolsa desnormalizado;
      FROM
20
```

```
29
     -- TIPO BOLSA
30 •
     INSERT INTO
                    tipo bolsa (tipo bolsa)
31
     SELECT DISTINCT tipo bolsa
32
     FROM
                    bolsa desnormalizado;
33
34
     -- BENEFICARIO **DISTINCT importante (mesmo beneficiario presente em + de um registro)**
35 •
     INSERT INTO
                    beneficiario (cpf, sexo, codigo raca, dt nascimento, deficiente fisico, codigo municipio)
36
     SELECT DISTINCT t1.cpf beneficiario bolsa, t1.sexo beneficiario bolsa, t2.codigo raca,
                    tl.dt nascimento beneficiario, tl.beneficiario deficiente fisico, t3.codigo municipio
37
38
     FROM
                    bolsa desnormalizado AS t1
                    `raça`
39
                                        AS t2 ON t2.raca = t1.raca beneficiario bolsa
     INNER JOIN
40
                                        AS t3 ON t3.nome municipio = t1.municipio beneficiario bolsa;
     INNER JOIN
                    municipio
41
42
     -- BOLSA PROUNT
43 •
     INSERT INTO
                    bolsa prouni(ano concessao, codigo tipo, codigo curso, codigo beneficiario)
                    tl.ano concessao bolsa, t2.codigo tipo, t3.codigo curso, t4.codigo beneficiario
44
     SELECT
45
     FROM
                    bolsa desnormalizado AS tl
46
     INNER JOIN
                    tipo bolsa AS t2 ON t2.tipo bolsa = t1.tipo bolsa
                                    AS t3 ON t3.modalidade ensino = t1.modalidade ensino bolsa
                    curso
47
     INNER JOIN
48
                    t3.nome curso
                                      = t1.nome curso bolsa
     AND
                    t3.turno
                                      = tl.nome turno curso bolsa
49
     AND
50
                    t3.codigo emec ies = t1.codigo emec ies bolsa
     AND
                    beneficiario AS t4 ON t4.cpf
51
     INNER JOIN
                                                                     = tl.cpf beneficiario bolsa
52
                    t4.dt nascimento = t1.dt nascimento beneficiario;
     AND
53
```

5. Consultas sobre o esquema normalizado

5.1 Listar quantos cursos ofertaram bolsas e quantas bolsas foram deferidas por instituição de ensino.

```
4 • SELECT nome_ies_bolsa,
COUNT(DISTINCT codigo_curso) AS cursos_ofertados, COUNT(codigo_bolsa_prouni) AS bolsas_deferidas
FROM curso
NATURAL JOIN info_ies
NATURAL JOIN bolsa_prouni
GROUP BY nome_ies_bolsa;
```

#	nome_ies_bolsa	cursos_ofertados	bolsas_deferidas
1	CENTRO UNIVERSITÁRIO DE B	15	82
2	CENTRO UNIVERSITÁRIO PAR	2	4
3	CENTRO UNIVERSITÁRIO SALE	2	14
4	PONTIFÍCIA UNIVERSIDADE CA	1	6
5	UNIVERSIDADE DA REGIÃO DE	5	32
6	UNIVERSIDADE DE PASSO FU	5	25

Função de agregação

5.2 Listar os códigos dos tipos de bolsas que foram ofertada para o curso de Sistemas de Informação OU os os códigos dos tipos de bolsas que foram ofertada para o curso de Administração.

```
codigo tipo
15 • ☐ ( SELECT
16
        FROM
                     bolsa prouni
17
        NATURAL JOIN curso
18
        WHERE
                     nome curso LIKE 'Sistemas de Informação' )
19
      UNION
20
    ☐ ( SELECT
                     codigo tipo
21
                     bolsa prouni
        FROM
22
        NATURAL JOIN curso
23
        WHERE
                     nome curso LIKE 'Administração');
24
```

5.3 a) Listar as informações das bolsas concedidas e as informações do curso quando o curso pertence à uma instituição de ensino específica.

```
SELECT
32
      FROM
              bolsa prouni
33
      NATURAL LEFT OUTER JOIN
34
35
          SELECT *
36
                 curso
37
          WHERE
                 codigo emec ies = 349
38
     AS t intra;
```

b) Listar as informações sobre as bolsas ofertadas para o curso de Sistemas de Informação e as informações dos tipos de bolsas.

```
SELECT
44
      FROM
              tipo bolsa
      NATURAL LEFT OUTER JOIN
46
    □(
47
          SELECT
48
                        bolsa prouni
49
                        curso
50
                        nome curso LIKE 'Sistemas de Informação'
          WHERE
51
        AS t intra;
```

Junção externa

5.4 Listar dados completos sobre determinada bolsa a partir de CPF do beneficiario

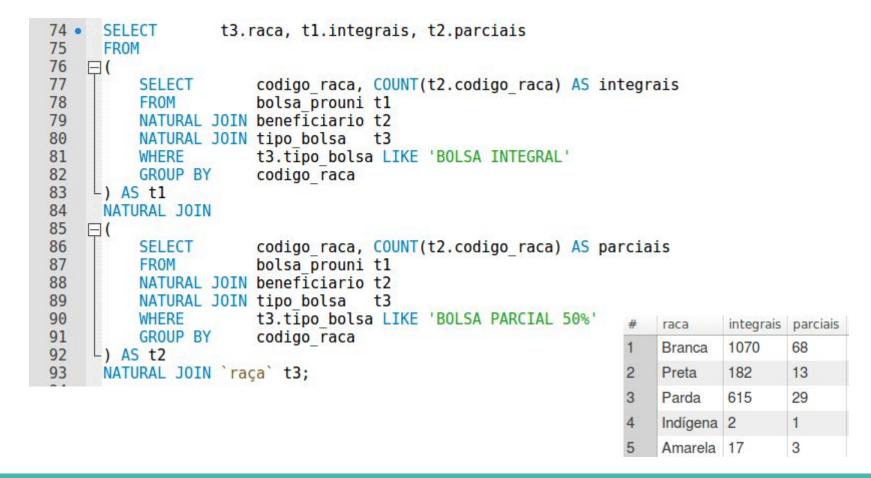
```
59 •
      SELECT
                   tl.ano concessao, t2.codigo emec ies, t3.nome ies bolsa, t4.tipo bolsa, t2.modalidade ensino,
                   t2.nome curso, t2.turno, t5.cpf, t5.sexo, t6.raca, t5.dt nascimento, t5.deficiente fisico,
60
                   t8.regiao, t8.sigla, t7.nome municipio
61
                   bolsa prouni tl
62
      FROM
63
      NATURAL JOIN curso
      NATURAL JOIN info ies
                                 t3
64
65
      NATURAL JOIN tipo bolsa
66
              JOIN beneficiario t5
      NATURAL JOIN 'raca'
67
                                 t6
                                 t7
68
      NATURAL JOIN municipio
69
      NATURAL JOIN info uf
                   t5.cpf LIKE '%882932%';
      WHERE
70
```

- Equivalente à consulta:

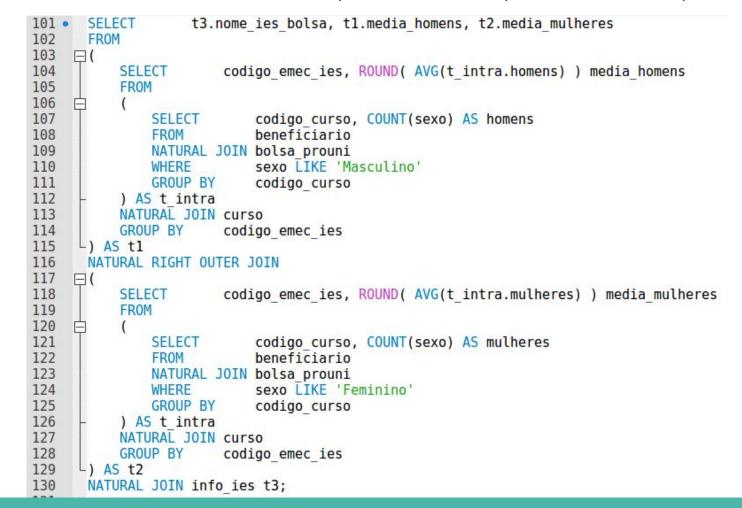
```
SELECT *
FROM bolsa_desnormalizado
WHERE cpf_beneficiario_bolsa LIKE '%882932%';
```

Junções entre três ou mais tabelas

5.5 a) Listar quantos beneficiários de cada raça foram contemplados com cada tipo de bolsa.



b) Listar a média de alunos de cada sexo que foram contemplados com bolsas por universidade.



Subconsulta

#	nome_ies_bolsa	media_homens	media_mulheres
1	PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ	HULL	6
2	UNIVERSIDADE DE PASSO FUNDO	3	4
3	CENTRO UNIVERSITÁRIO PARA O DESENVOLVIMENTO DO ALTO VALE DO ITAJAÍ	2	1
1	UNIVERSIDADE DA REGIÃO DE JOINVILLE	4	3
5	UNIVERSIDADE DO VALE DO ITAJAÍ	4	6
3	UNIVERSIDADE ESTÁCIO DE SÁ	2	2
7	UNIVERSIDADE METODISTA DE PIRACICABA	4	5
3	UNIVERSIDADE PITÁGORAS UNOPAR	2	1
)	UNIVERSIDADE PRESIDENTE ANTÔNIO CARLOS	2	3
0	UNIVERSIDADE PAULISTA	5	7
1	CENTRO UNIVERSITÁRIO DE BELO HORIZONTE	3	3
2	UNIVERSIDADE TIRADENTES	2	3
3	UNIVERSIDADE REGIONAL INTEGRADA DO ALTO URUGUAI E DAS MISSÕES	3	4
4	UNIVERSIDADE LUTERANA DO BRASIL	4	8
5	UNIVERSIDADE DO SUL DE SANTA CATARINA	3	4
6	CENTRO UNIVERSITÁRIO SALESIANO DE SÃO PAULO	6	3

Obrigado!

