

Normalização de Base Dados Relacional

TRABALHO PRÁTICO - APRESENTAÇÃO

CEBDSIG 2021/2022 – CRIAÇÃO E ESTRUTURAÇÃO DE BASES DE DADOS EM SIG

DOCENTES: PATRÍCIA ABRANTES E JORGE ROCHA

ALUNO: JORGE URBANO PEREIRA, N° 4687

Normalização de Base Dados Relacional

- ▶ Objetivos
- ▶ Preparação e tratamento da tabela
- ▶ Modelo Relacional
- ▶ Processo de normalização de uma BDR
- ▶ Tipo de Relações
- ▶ Esquema Relacional
- ▶ Notas conclusivas
- ▶ Referências bibliográficas consultadas

Objetivos

Este relatório tem enfoque na descrição e aplicação do processo de normalização de uma base dados relacional até atingir a terceira forma normal. Foca-se em três pontos:

- ▶ tratamento efetuado à base de dados;
- ▶ descrição dos processos de normalização da base de dados;
- ▶ modelo relacional, esquema relacional, tipo de relações e chaves utilizadas.

Tem por base o ficheiro *Base_de_dados_TP_2021_2022.xlsx*, que é uma tabela, e que será reconfigurada e subdividida em outras novas tabelas com a informação simplificada, não redundante. Foi organizada de modo a ser possível às novas tabelas relacionarem-se entre si através de atributos comuns, criando assim a estrutura do modelo relacional dando destaque às chaves primárias e estrangeiras utilizadas, ao tipo de relações e por fim ao esquema relacional.

Preparação e tratamento da tabela

Antes de iniciar o processo de normalização da base de dados relacional foi necessário em primeiro lugar reconfigurar, reorganizar e tratar os dados da tabela de base dada.

Actividade (Termo corrente, Termo específico, Designação original, Espaço, Cidade, Periodicidade, Duração, Distrito, Termo chave)

9 atributos (é o grau da relação)

2590 tuplos (é a cardinalidade da relação).

O tipo de todos os nove atributos é *string*, texto

Registos eliminados:

- ▶ 28 registos sem valor para o atributo [Designação original].
- ▶ 23 registos com a palavra **APAGAR?! e *FICHA ERRADA!!!*.
- ▶ 714 registos duplicados.
- ▶ 143 registos com 4 ou menos colunas preenchidas.

Preparação e tratamento da tabela

- ▶ No trabalho de revisão dos atributos foi eliminado o atributo [**Duração**]. Este apresenta a percentagem mais baixa (35,5%) de preenchimento dos registos e a informação apresentada é do tipo “5 Dias”. Assim, não sendo um atributo do tipo *number* (para valores numéricos), não é possível inquirir a Base de Dados Relacional (BDR) para, por exemplo, saber qual o evento com maior duração.
- ▶ Também foram revistos os atributos [**Distrito**] que, juntamente com [**cidade**], apresentam percentagens elevadas de preenchimento dos registos com valores muito próximos (95,3% e 80,1% respetivamente). Os registos destes dois atributos são repetidos, e assim sendo fez-se um *merge* a estes dois atributos e denominou-se [**cidade**]. Desta forma a taxa mais elevada de preenchimento dos registos manteve-se sem prejuízo da integridade dos dados.
- ▶ caracteres especiais do tipo () “ \$? ! / * ... [] ?, estes foram retirados. Por fim, substituíram-se os campos vazios por *null*.

Modelo Relacional

- ▶ Em toda a BDR, não pode haver dois campos com o mesmo nome. Cada campo é identificado de modo único.
- ▶ Não devem existir campos vazios. Caso o valor de um campo seja desconhecido ou não aplicável, então deve ser preenchido com um valor *null*.
- ▶ O domínio de todos os campos deve ser constituído por valores atômicos. Não é permitido incluir mais do que um valor em cada campo.
- ▶ Cada linha da tabela representa uma entidade ou ocorrência única. Por isso, não podem haver registos duplicados

Processo de normalização de uma BDR

► Primeira forma normal (1FN)

- todos os seus campos sejam atômicos
- não exista grupos de valores repetidos.

Eventos

ID	Desig_orig	Espaco	Cidade	Periodicidade
1	feira de s. bartolomeu	transferida do cais das ameias para o Rossio	Coimbra	null
2	espetáculos	no teatro d'esta cidade	Santarém	única
3	sarau	Antigo convento de Santo António, extra-mu	Évora	null
4	tourada	arena na praça de s. joão	Setúbal	null
5	soirée	assembleia viseense	Viseu	null
6	reunião	assembleia viseense	Viseu	null
7	procissão com a imagem de	até a sé; no dia seguinte para a igreja de stª r	Castelo Branco	null
8	Baile	Ateneu Comercial	Portalegre	única
9	feira de S. Bartolomeu	Av. Navarro	Coimbra	anual
10	arraial	Bairro da Capelinha	Évora	null

Registo: 1 de 1682 Sem Filtro Procurar

Termos

ID	Termo_especif	Termo_corr	Termo_chav
1	trabalho	feiras e exposições	feira
2	cultura e lazer	artes e espectáculos	espectáculo
3	cultura e lazer	artes e espectáculos	sarau
4	cultura e lazer	artes e espectáculos	tourada
5	cultura e lazer	festa	festas
6	cultura e lazer	festa	festas
7	religião	festa	procissão
8	cultura e lazer	festa	baile
9	trabalho	feiras e exposições	feira
10	cultura e lazer	festa	arraial

Registo: 1 de 1682 Sem Filtro Procurar

Processo de normalização de uma BDR

► Segunda forma normal (2FN)

- todos os seus campos estejam na 1FN
- exista a total dependência da chave primária composta

ID	Cod_TE	Cod_TC	Cod_TCh
+	1 TE05	TC07	TCh20
+	2 TE01	TC01	TCh17
+	3 TE01	TC01	TCh34
+	4 TE01	TC01	TCh37
+	5 TE01	TC08	TCh21
+	6 TE01	TC08	TCh21
+	7 TE04	TC08	TCh31
+	8 TE01	TC08	TCh04
+	9 TE05	TC07	TCh20
+	10 TE01	TC08	TCh03

Registo: 1 de 1682 Sem Filtro

Tab lig (ID, Cod_TE, Cod_TC, Cod_TCh)

Termo_especif
(Cod_TE, Termo_especif)

Cod_TE	Termo_especif
+	TE01 cultura e lazer
+	TE02 mutualidade e beneficência
+	TE03 política e cívica
+	TE04 religião
+	TE05 trabalho

Registo: 1 de 5 Sem Filtro

Termo_chav
(Cod_TCh, Termo_chav)

Cod_TCh	Termo_chav
+	TCh01 almanaque
+	TCh02 animatógrafo
+	TCh03 arraial
+	TCh04 baile
+	TCh05 bazar
+	TCh06 bodo
+	TCh07 carnaval entrudo
+	TCh08 ciclismo
+	TCh09 cinema
+	TCh10 circo
+	TCh11 comemoração
+	TCh12 concerto
+	TCh13 conferência
+	TCh14 coreto
+	TCh15 correio telégrafo
+	TCh16 desporto outros
+	TCh17 espectáculo
+	TCh18 excursão
+	TCh19 exposição
+	TCh20 feira
+	TCh21 festas
+	TCh22 futebol
+	TCh23 ginástica
+	TCh24 hipismo
+	TCh25 hotel
+	TCh26 jornal
+	TCh27 livraria
+	TCh28 mercado
+	TCh29 música
+	TCh30 passeio
+	TCh31 procissão
+	TCh32 revista
+	TCh33 romagem
+	TCh34 sarau
+	TCh35 subscrição
+	TCh36 teatro
+	TCh37 tourada
+	TCh38 variedades

Registo: 1 de 38 Sem Filtro

Termo_corr
(Cod_TC, Termo_corr)

Cod_TC	Termo_corr
+	TC01 artes e espectáculos
+	TC02 comércio
+	TC03 conferência congresso
+	TC04 culto
+	TC05 desporto
+	TC06 edições
+	TC07 feiras e exposições
+	TC08 festa
+	TC09 hospedagem
+	TC10 passeio excursão
+	TC11 subscrição pública
+	TC12 tipografia imprensa
+	TC13 transportes e comunicações

Registo: 1 de 13 Sem Filtro

Processo de normalização de uma BDR

► Terceira forma normal (3FN)

- todos os seus campos estejam na 2FN
- não existam dependências funcionais entre os atributos que não são chave primária

ID	Desig_orig	Periodicidade
1	feira de s. bartolomeu	null
2	espetáculos	única
3	sarau	null
4	tourada	null
5	soirée	null
6	reunião	null
7	procissão com a imagem	null
8	Baile	única
9	feira de S. Bartolomeu	anual
10	arraial	null

Registo: 1 de 1682 Sem Filtro

Eventos (ID, Desig_orig, Periodicidade)

ID	Espaco	Cidade
1	transferida do cais das ameias para o R	Coimbra
2	no teatro d'esta cidade	Santarém
3	Antigo convento de Santo António, extra	Évora
4	arena na praça de s. joão	Setúbal
5	assembleia viseense	Viseu
6	assembleia viseense	Viseu
7	até a sé no dia seguinte para a igreja de	Castelo Branco
8	Ateneu Comercial	Portalegre
9	Av. Navarro	Coimbra
10	Bairro da Capelinha	Évora

Registo: 1 de 1682 Sem Filtro Procurar

Localiz (ID, Espaco, Cidade)

Processo de normalização de uma BDR

► Terceira forma normal (3FN)

Cod_Esp	Espaco
+	AH01 Teatro Angrense
+	AV01 liceu de aveiro
+	AV02 museu de aveiro
+	AV03 praça do município
+	AV04 sede do grémio aveirense
+	AV05 teatro aveirense
+	AV06 teatro da rua do rato
+	AV07 teatro da rua dos condes
+	BC01 casa de d. candida furtado
+	BC02 coreto
+	BC03 igreja da sé
+	BC04 sede da Assembleia Brigantina
+	BC05 teatro brigantino
+	BC06 teatro camões
+	BG01 edificio municipal
+	BG02 Monte Sameiro
+	BG03 num monte próximo
+	BG04 passeio público

DTCC	Cidade
+	0105 Aveiro
+	0205 Beja
+	0303 Braga
+	0402 Bragança
+	0502 Castelo Branco
+	0603 Coimbra
+	0701 Évora
+	0805 Faro
+	0907 Guarda
+	1009 Leiria
+	1106 Lisboa
+	1214 Portalegre
+	1312 Porto
+	1416 Santarém
+	1512 Setúbal
+	1609 Viana do Castelo
+	1714 Vila Real
+	1823 Viseu

Cod_Period	Periodicidade
+	PER01 anual
+	PER02 bianual
+	PER03 mensal
+	PER04 trimestral
+	PER05 semanal
+	PER06 quinzenal
+	PER07 bissemanal
+	PER08 trissemanal
+	PER09 diária
+	PER10 única
+	PER11 irregular

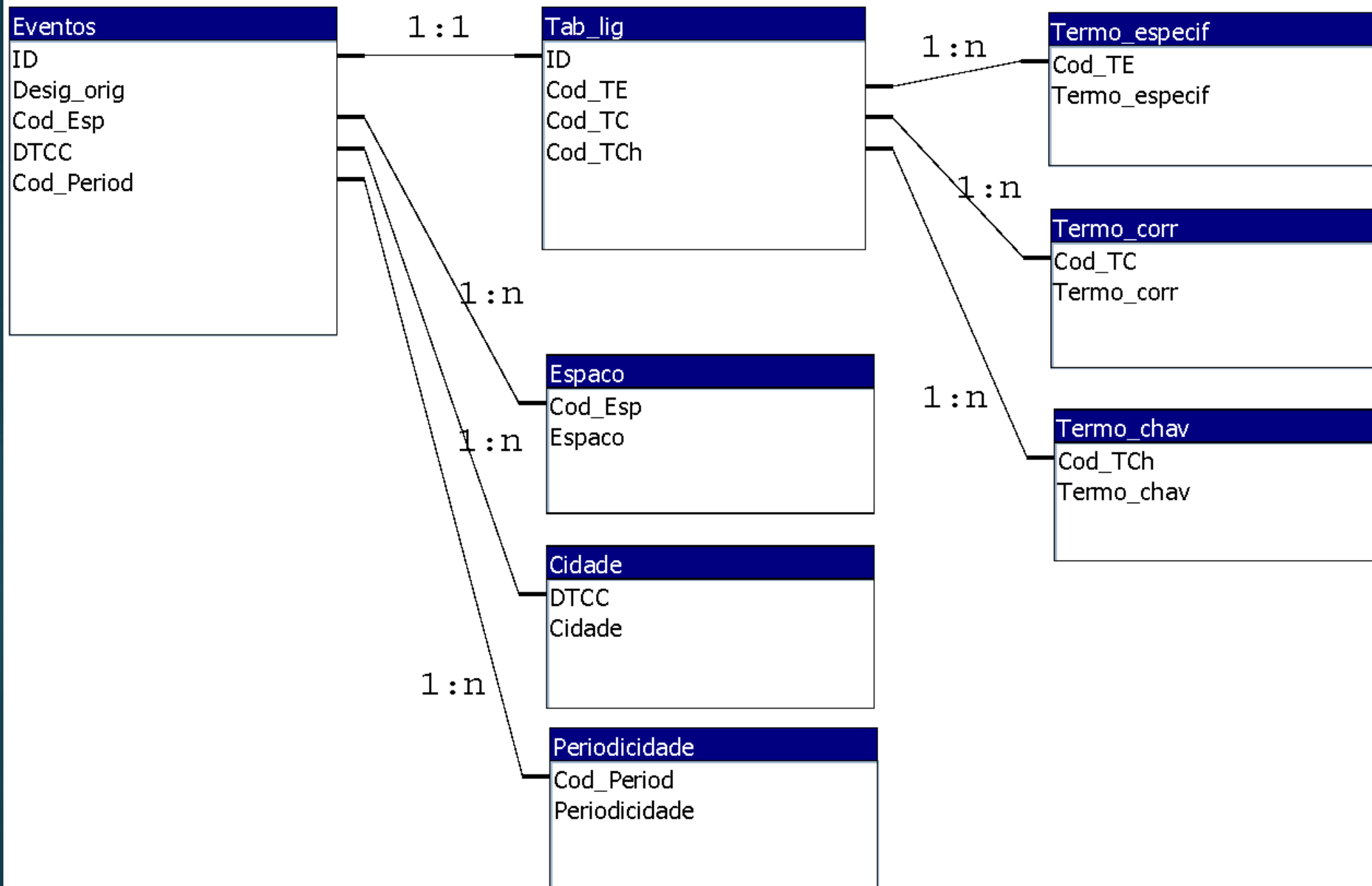
Eventos (ID, Desig_orig, Cod_Period, Cod_Esp, DTCC)

ID	Desig_orig	Cod_Period	Cod_Esp	DTCC
1	feira de s. bartolomeu	null	CO01	0603
2	espetáculos	PER10	ST01	1416
3	sarau	null	EV001	0701
4	tourada	null	SB01	1512
5	soirée	null	VI01	1823
6	reunião	null	VI01	1823
7	procissão com a imagem	null	CB01	0502
8	Baile	PER10	PL01	1214
9	feira de S. Bartolomeu	PER01	CO02	0603
10	arraial	null	EV002	0701

Espaco (Cod_Esp, Espaco) Cidade (DTCC, Cidade) Periodicidade (Cod_Period, Periodicidade)

Tipo de Relações

- ▶ **Relações um para um (1:1):** quando um elemento de uma entidade só tem ligação com um elemento de outra entidade (por exemplo relação entre o ID de Eventos e o ID de Tab_lig).
- ▶ **Relações um para vários (1:n):** quando um elemento de uma entidade tem ligação com vários elementos da outra entidade (por exemplo relação entre o Termo_especif e Cod_TE).
- ▶ **Relações de vários para vários (n:n):** quando um elemento da entidade A tem vários elementos na entidade B, e um elemento da entidade B tem vários elementos na entidade A. Este tipo de relação terá que ser desmontado em duas relações do tipo um para vários, recorrendo a uma chave composta. Este tipo de relação não se verifica neste trabalho de normalização.



Esquema Relacional

Notas conclusivas

- ▶ Procedeu-se neste trabalho a um processo de normalização de uma BDR. Com o objectivo da normalização alcançado, podemos avaliar a qualidade do desenho de tabelas e transformá-lo num desenho de conjunto de tabelas equivalente, menos redundante, mais estável e que permite consultas mais satisfatórias à BDR e também a possibilidade de alterar, apagar e introduzir novos dados de modo ágil, sem duplicação e incoerências.
- ▶ Esta BDR normalizada possibilita a distribuição espacial dos eventos através de atributos com características espaciais, como o código INE para distrito e concelho (DTCC) associado à cidade onde decorre o evento. Podemos assim representar graficamente a sua distribuição num mapa.
- ▶ Outro modo de apresentar a informação num mapa seria a de geocodificar, através dum processo de *Geocoding – Address matching* (disponível por exemplo no *Batch geocoder* em <https://geocode.localfocus.nl/> ou em alternativa o *Batchgeo* em <https://pt.batchgeo.com/>), onde o atributo *Espaco* poderia fornecer o local mais preciso do evento. Este método apresenta algumas limitações, mas é possível para grande parte dos registos do atributo *Espaco*.
- ▶ Ficheiros em <https://github.com/Norge0/CEBDSIG/>

Referências bibliográficas consultadas

Gilmore, W. Jason; Treat, Robert H. (2006), *Beginning PHP and PostgreSQL 8: From Novice to Professional*. Apress, Berkeley.

Groff, James R.; Weinberg, Paul N. (1999), *SQL: The Complete Reference*. Osborne/McGraw-Hill. Berkeley.

Yeung, Albert K.W.; Hall, G. Brent, (2007), *Spatial Database Systems. Design, Implementation and Project Management. The GeoJournal Library (Vol. 87)*. Springer, Dordrecht.

► Sites consultados

Oracle - What is a Relational Database (RDBMS),
<<https://www.oracle.com/pt/database/what-is-a-relational-database/>>

W3schools - Learn to Code, em <<https://www.w3schools.com/sql/>>