



Disciplina de Gestão de Dados e de Bases de Dados

Ano Letivo 2020/2021

Normalização Relacional

Tópicos

- Etapas da modelização da informação
- Normalizar porquê?
- Objetivos da normalização
- Etapas da normalização
- Como decompor uma relação
- Exemplo de normalização

Normalizar Porquê?

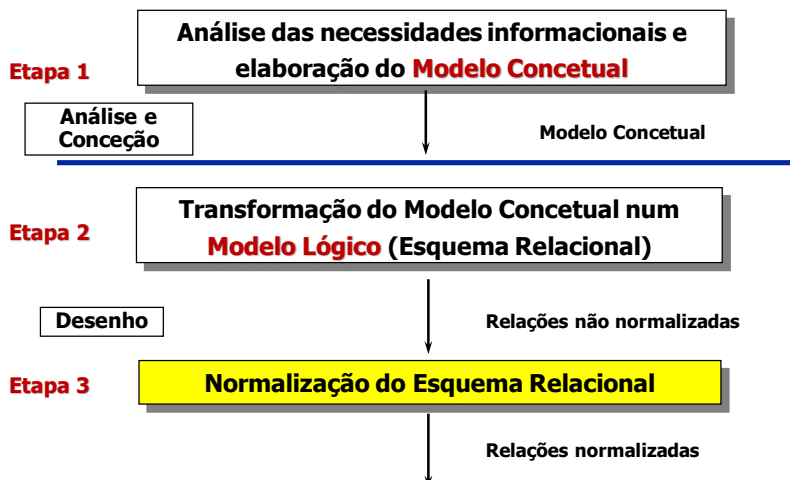
Após a construção do modelo concetual dos dados (Diagrama de Classes) é feita a transformação para o modelo lógico (Esquema Relacional).

O Esquema Relacional obtido representa a estrutura da informação de um modo natural e completo

Mas terá o mínimo de redundância e será o mais estável possível ?

A Normalização tem como objetivo **avaliar a qualidade do Esquema Relacional** e decompô-lo, em caso de necessidade, noutro Esquema Relacional equivalente, **menos redundante e mais estável**

Etapas da Modelização da Informação



Objetivos da Normalização

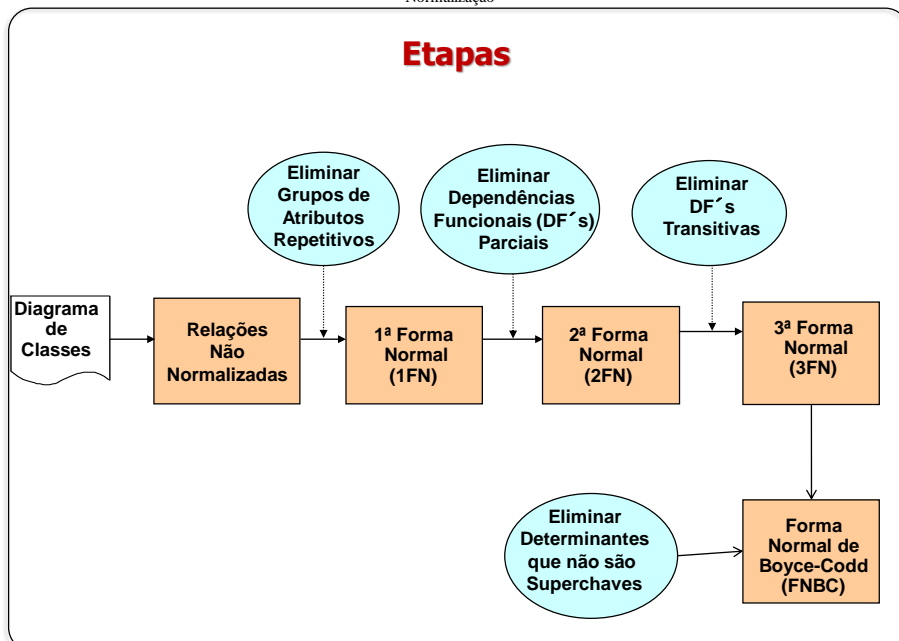
- **Minimizar a redundância de dados**
- **Conseguir que cada relação possa ser consistente e corretamente atualizada**

N_Func	Nome	Categoria	Vencimento	Disciplina
10234	Tiago Silva	Prof. Associado	4.000 €	GF1
10345	Tomás Ribeiro	Prof. Auxiliar	3.000 €	TI
11908	Maria Marques	Assistente	2.000 €	SIG
16000	Carlos Antunes	Prof. Auxiliar	3.000 €	TI
17241	Raquel Silva	Assistente	2.000 €	SIG

Assumindo que a cada categoria corresponde um determinado salário

- **Problemas de Atualização** – Para alterar o salário de uma categoria têm de se alterar várias linhas (todos os docentes com essa categoria)
- **Problemas de Remoção** - Se se apagar o registo do Tiago Silva deixa de se ter informação relativa ao salário de um Prof. Associado

Etapas



Exemplo

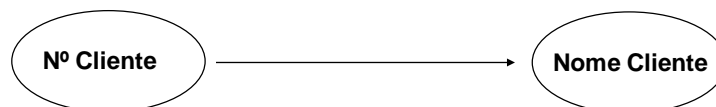
Numa Escola, pretende-se manter informação sobre:

- os Estudantes da escola (nº interno, nome e curso a que pertencem)
- as Disciplinas que são ministradas na escola (nº disciplina e nome)
- os Professores contratados pela escola (código, nome e categoria)
- Inscrições dos Alunos em Disciplinas
- Notas obtidas pelos alunos nas disciplinas em que estão inscritos
- Considera-se que a uma Disciplina está afeto um único Professor (responsável), mas que um Professor pode ser responsável por várias Disciplinas

Dependências Funcionais

Por vezes, dois atributos estão intrinsecamente ligados entre si.

Ex:



Num determinado instante, em qualquer ponto da base de dados onde figurem estes dois atributos, a um mesmo nº de cliente corresponderá necessariamente o mesmo nome (*).

Isto significa que é válida a Dependência Funcional (DF)

Nº Cliente → Nome Cliente

E diz-se que Nº Cliente é Determinante para Nome Cliente

(*) Note que o inverso poderá não ser verdade!

Axiomas de Armstrong

As dependências funcionais satisfazem um conjunto de regras de inferência

1. Reflexividade

Se $A \subseteq B$ então $B \rightarrow A$

$nFatura \subseteq nFatura, data$ então $nFatura, data \rightarrow nFatura$

2. Aumento - Dependências Funcionais Parciais

Se $nFatura \rightarrow nFornecedor$ existe

também $nFatura, data \rightarrow nFornecedor$ existe

A DF $nFatura, data \rightarrow nFornecedor$ diz-se **parcial** porque $nFornecedor \rightarrow nFatura$

Axiomas de Armstrong

3. Transitividade: Dependências Funcionais Transitivas

Se $nFatura \rightarrow nFornecedor$ e $nFornecedor \rightarrow nomeFornecedor$

também $nFatura \rightarrow nomeFornecedor$ que se designa por **DF**

Transitiva

- **Chaves Candidatas**

São os atributos de uma relação que identificam, de forma inequívoca, uma ocorrência específica dessa relação, distinguindo-a das restantes. **A chave tem de conter todos os atributos necessários à identificação de todos os restantes atributos da relação**, ou seja Chave → Atributo, para todos os atributos da relação

- **Atributos Primos**

São os **atributos** que pertencem a uma **chave candidata**

- **Chave Primária**

É a chave candidata escolhida

- **Superchave**

Qualquer conjunto de atributos que contém uma chave candidata

Ex: Linha_fatura (num_fatura, num_linha, cod_produto, quantidade, preço unitário)

Chave primária= num_fatura, num_linha

Uma superchave= num_fatura, num_linha, cod_produto

- **Descritores**

São os atributos que apenas descrevem ou caracterizam as ocorrências de uma relação

Definição de 1ª FORMA NORMAL

Uma relação está na 1ª Forma Normal (1FN) se e só se cada linha contém exatamente um valor para cada atributo

Dado que as Tabelas são estruturas bidimensionais, então no cruzamento de uma linha com uma coluna (atributo) só é possível armazenar valores atômicos

Relação Não Normalizada

Relação NOTAS

Nº Estudante	Nome Estudante	Curso	Nº Disciplina	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor	Nota
21934	Antunes	Informática	04	Álgebra	21	Gil Alves	PA	15
			14	Análise Sist.	87	Ana Lopes	PC	-
			23	Progr.Linear	43	Paulo Pinto	AS	16
42346	Bernardo	Matemática	08	Topologia	32	Nuno Neves	AE	10
			04	Álgebra	21	Gil Alves	PA	-
			12	Geometria	21	Gil Alves	PA	18
			16	Lógica	32	Nuno Neves	AE	13
54323	Correia	Estatística	04	Álgebra	21	Gil Alves	PA	11
			08	Topologia	32	Nuno Neves	AE	10
...

Dados em forma tabular, conforme são visualizados pelo utilizador

Esta relação foi desenhada para representar Inscrições e Aprovações de Alunos em Disciplinas. Cada estudante está inscrito e/ou já foi aprovado em várias disciplinas, com a informação correspondente. Assim, temos um grupo de atributos repetitivo :

{NºDisciplina, NomeDisciplina, Cod.Professor, Nome Professor, CategoriaProfessor, Nota}

Como Decompôr uma Relação

Teorema da Decomposição Binária

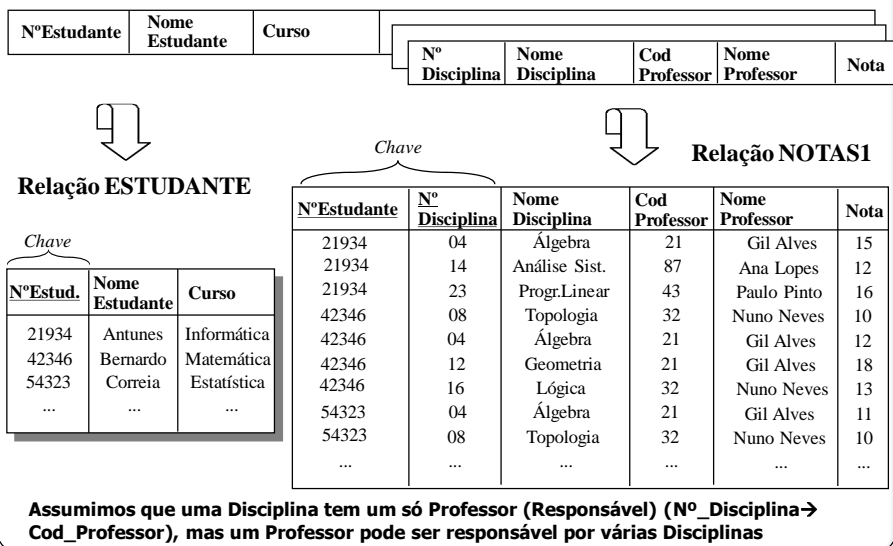
Dada uma relação $R(A, B, C, D, E, F)$ e dada a dependência funcional $A \rightarrow B$, a relação é sempre decomponível em $R_1(A, B)$ e $R_2(A, C, D, E, F)$ ou seja $R_2(A, U-A-B)$

Aplicando o Teorema da Decomposição Binária

Notas (N° Estudante, NomeEstudante, Curso, N° Disciplina, NomeDisciplina, Cod.Professor, NomeProfessor, CategoriaProfessor, Nota)

e N° Estudante \rightarrow NomeEstudante, Curso então Notas é decomponível em

Estudante (N° Estudante, NomeEstudante, Curso) e Notas1 (N° Estudante, N° Disciplina, NomeDisciplina, Cod.Professor, NomeProfessor, CategoriaProfessor, Nota)

1ª Forma Normal**1ª FORMA NORMAL****Problemas na relação NOTAS1**– **Remoção**

Se quisermos apagar a informação sobre todos os alunos aprovados ou inscritos numa determinada disciplina, então perdemos toda a informação dessa disciplina e do respetivo professor responsável

– **Atualização**

Se pretendermos modificar o nome de uma disciplina (por exemplo Álgebra passa a Álgebra e Aplicações) é necessário percorrer toda a relação e fazer essa modificação para todos os alunos inscritos ou aprovados nessa disciplina

No caso de falhar a aplicação de modificação em alguma ocorrência, então teríamos dados inconsistentes

Definição de 2ª FORMA NORMAL

Uma relação está na 2ª Forma Normal (2FN) se está na primeira e se todos os atributos não primos⁽¹⁾ dependerem da totalidade da chave e não apenas de parte dela (não podem existir Dependências Funcionais Parciais).

(1) Que não pertencem a nenhuma chave candidata

Dependências Funcionais Parciais

Chave

Relação NOTAS1

<u>Nº Estudante</u>	<u>Nº Disciplina</u>	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor	Nota
21934	04	Álgebra	21	Gil Alves	PA	15
21934	14	Análise Sist.	87	Ana Lopes	PC	12
21934	23	Progr.Linear	43	Paulo Pinto	AS	16
42346	08	Topologia	32	Nuno Neves	AE	10
42346	04	Álgebra	21	Gil Alves	PA	12
42346	12	Geometria	21	Gil Alves	PA	18
42346	16	Lógica	32	Nuno Neves	AE	13
54323	04	Álgebra	21	Gil Alves	PA	11
54323	08	Topologia	32	Nuno Neves	AE	10
...

Os atributos **Nome Disciplina**, **Cod-Professor**, **Nome-Professor** e **Categoria-Professor** dependem apenas do atributo **Nº Disciplina** (que está estritamente contido na chave da relação)

Aplicando o Teorema da Decomposição Binária

Notas1 (NºEstudante, NºDisciplina, NomeDisciplina, Cod.Professor, NomeProfessor, CategoriaProfessor, Nota)

Nº Disciplina → Nome Disciplina, Cod-Professor, Nome-Professor e Categoria-Professor então Notas1 é decomponível em

Disciplina (Nº Disciplina, Nome Disciplina, Cod-Professor, Nome-Professor e Categoria-Professor) e

Notas2 (NºDisciplina, NºEstudante, Nota)

2ª Forma Normal

Nº Estudante	Nº Disciplina	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor	Nota
--------------	---------------	-----------------	---------------	----------------	---------------------	------

Chave e atributos que dependem da totalidade desta

Relação NOTAS2

Nº Estudante	Nº Disciplina	Nota
21934	04	15
21934	14	12
21934	23	16
42346	08	10
42346	04	12
42346	12	18
42346	16	13
54323	04	11
54323	08	10
...

Atributos que dependem de parte da chave mais a referida parte da chave

Relação DISCIPLINA

Nº Disciplina	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor
04	Álgebra	21	Gil Alves	PA
14	Análise Sist.	87	Ana Lopes	PC
23	Progr.Linear	43	Paulo Pinto	AS
08	Topologia	32	Nuno Neves	AE
12	Geometria	21	Gil Alves	PA
16	Lógica	32	Nuno Neves	AE
...

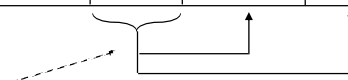
Definição de 3ª FORMA NORMAL

Uma relação está na 3ª Forma Normal (3FN) se está na 2ª Forma Normal e se não existirem atributos não primos⁽¹⁾ a dependerem funcionalmente de outros atributos (que não pertençam à chave)- não podem existir Dependências Funcionais Transitivas da chave. Assim sendo, cada atributo deve depender apenas das Chaves Candidatas da relação

(1) Que não pertencem a nenhuma chave candidata

Dependências Funcionais Transitivas

Chave		Relação DISCIPLINA		
Nº Disciplina	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor
04	Álgebra	21	Gil Alves	PA
14	Análise Sist.	87	Ana Lopes	PC
23	Progr.Linear	43	Paulo Pinto	AS
08	Topologia	32	Nuno Neves	AE
12	Geometria	21	Gil Alves	PA
16	Lógica	32	Nuno Neves	AE
...



Os atributos **Nome Professor** e **Categoria Professor** dependem do atributo **Cod.Professor** (que não é chave da relação) e portanto as DFs $N^{\circ}Disciplina \rightarrow Nome_Professor$ e $N^{\circ}Disciplina \rightarrow Categoria_Professor$ são Transitivas, porque, p.ex., $N^{\circ}Disciplina \rightarrow CodProfessor$ e $CodProfessor \rightarrow Nome_Professor$).

Aplicando o Teorema da Decomposição Binária

Disciplina (Nº Disciplina, Nome Disciplina, Cod-Professor, Nome-Professor e Categoria-Professor)

Cod-Professor → Nome-Professor, Categoria-Professor então
Disciplina é decomponível em

Professor (Cod-Professor, Nome-Professor, Categoria-Professor)

Disciplina1 (Cod_Professor, Nº Disciplina, Nome Disciplina)

3ª Forma Normal

<u>Nº Disciplina</u>	Nome Disciplina	Cod Professor	Nome Professor	Categoria Professor
----------------------	-----------------	---------------	----------------	---------------------

Atributos que dependem do atributo não chave + esse atributo

Atributos que apenas dependem da chave

Relação PROFESSOR

<u>Cod Professor</u>	Nome Professor	Categoria Professor
21	Gil Alves	PA
87	Ana Lopes	PC
43	Paulo Pinto	AS
32	Nuno Neves	AE
...

Relação DISCIPLINA1

<u>Nº Disciplina</u>	Nome Disciplina	Cod Professor
04	Álgebra	21
14	Análise Sist.	87
23	Progr.Linear	43
08	Topologia	32
12	Geometria	21
16	Lógica	32
...

Outra Definição de 3ª FORMA NORMAL

Uma relação está na 3ª Forma Normal se cada atributo depende apenas da chave primária (1FN), de toda a chave (2FN) e de nada mais do que da chave (3FN)

Forma Normal de Boyce-Codd (FNBC)

Uma relação está na Forma Normal de Boyce-Codd se está na 3FN e todos os Determinantes⁽¹⁾ são Superchaves

A FNBC só é diferente da 3FN quando:

- **A relação tem mais do que uma chave candidata**
- **As chaves candidatas têm atributos em comum**

(1) Se num_cliente → nome_cliente, então num_cliente é determinante

Forma Normal de Boyce-Codd (FNBC)

Exemplo Turma (UC, Estudante, Professor)

Com **Professor** → UC (um Professor é docente de uma única UC, mas o inverso não é verdadeiro. Professor é determinante e UC é primo – pertence a uma chave candidata)

Chaves Candidatas:

UC, Estudante → Professor

Professor, Estudante → UC (Propriedade do Aumento)

Está na 3FN mas não na FNBC porque Professor (Determinante) não é superchave da relação

Como **Professor** → UC

Vamos partir a relação em:

Professor (Professor, UC) e Estudante (Professor, Estudante)

Anomalias na relação Turma ?

Problemas na Relação Turma

Turma (UC, Estudante, Professor)

Professor → UC

Outra Chave Candidata (Professor, Estudante)

– Inserção

Não é possível inserir um Professor e atribuí-lo a uma UC enquanto não houver pelo menos um aluno inscrito nessa UC

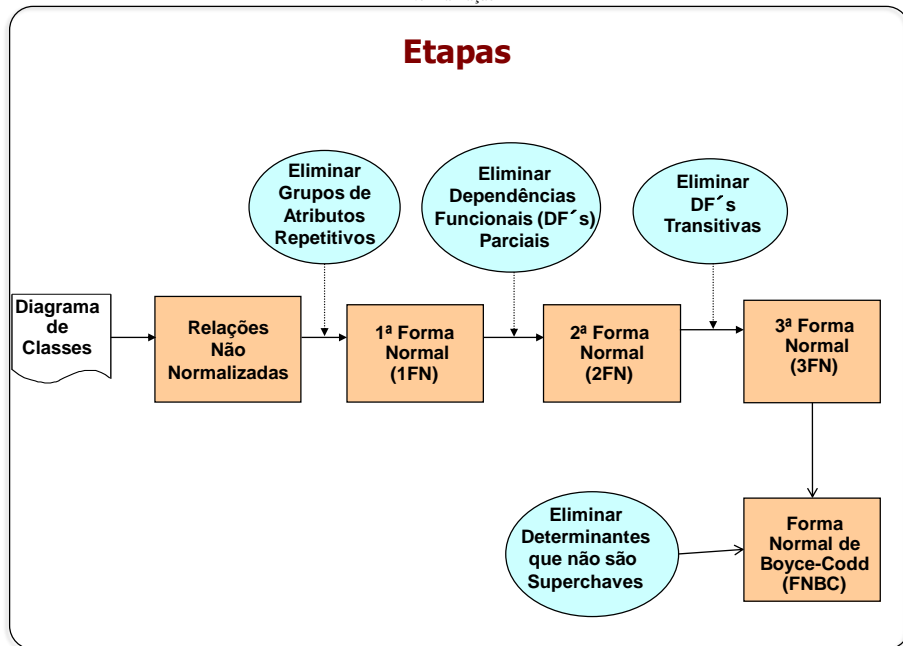
– Remoção

Se quisermos apagar a informação sobre todos os alunos inscritos numa UC perdemos a informação sobre os Professores afetos a essa UC

– Atualização

Se pretendermos modificar o nome de uma UC é necessário percorrer toda a relação e fazer essa modificação para todos os estudantes afetos a essa UC

No caso de falhar a aplicação de modificação em alguma ocorrência, então teríamos dados inconsistentes



Pretendemos com esta aula sobre Normalização que os alunos compreendessem:

- Etapas da modelização da informação
- Normalizar porquê?
- Objetivos da normalização
- Etapas da normalização