

Instituto Federal de São Paulo

Banco de Dados 2 – Aula 5

NORMALIZAÇÃO

Prof. Dr. Cecilio Merlotti Rodas

Email: cecilio.rodas@ifsp.edu.br



Normalização

- Consiste da **decomposição de esquemas** para evitar anomalias de atualização, inserção e exclusão.
 - Algumas tabelas podem estar mal dimensionadas, com colunas na tabela errada. E a decomposição tem o objetivo de tornar o esquema mais otimizado.
- Bom design **evita redundância de dados e anomalias de atualização.**
- Mecanismo formal para analisar esquemas de relação (tabela) baseado nas suas chaves e nas **dependências funcionais** entre seus atributos.



Normalização

- Utilizada em:
 - Documentação de sistemas legados (antigos, não relacionais)
 - Migração para novas plataformas
 - Organização dos dados de um sistema
- Um projeto conceitual bem elaborado **termina naturalmente** em esquemas normalizados.



Normalização

- É um processo onde se aplica regras a todas as relações (tabelas) do banco de dados, a fim de evitar falhas no projeto como:
 - **redundância** de dados; **mistura de diferentes assuntos** numa mesma entidade; anomalias de atualização, inserção exclusão; entre outros problemas.



Normalização

- Basicamente, aplicando e respeitando as regras de cada uma dessas formas normais, poderemos garantir um **banco de dados mais íntegro**, com uma grande possibilidade de sucesso no seu projeto.
- As relações que atendem a um determinado conjunto de regras diz-se estarem em uma determinada **Forma Normal**.



Documento exemplo: Relatório de Alocação a Projeto

Código do projeto: LSC001			Tipo: Novo desenvolvimento		
Descrição:			Sistema de Estoque		
Código do empregado	Nome	Categ. Funcional	Salário	Dt início projeto	Tempo de alocação
2146	João	A1	4	1/11/01	24
3145	Sílvio	A2	4	2/10/01	24
6126	José	B1	9	3/10/02	18
1214	Carlos	A2	4	4/10/02	18
8191	Mário	A1	4	1/11/02	12
Código do projeto: PAG02			TIPO: Manutenção		
Descrição:			Sistema de RH		
Código do empregado	Nome	Categ. Funcional	Salário	Dt início projeto	Tempo de alocação
8191	Mário	A1	4	1/05/03	12
4112	João	A2	4	4/01/01	24
6126	José	B1	9	1/11/02	12



Tabela não normalizada

- Pode possuir tabelas aninhadas, ou seja, podemos identificar atributos que possam formar outras tabelas;
- Em uma tabela não normalizada é possível encontrar grupo **repetido**, **coluna multivalorada**, **coluna não atômica**.



Normalização das Relações

Problemas de relações mal normalizadas:

- Anomalia de Eliminação,
- Anomalia de Inserção e
- Anomalia de Modificação.

cliente = { nome, atividade, taxa }

Nome	Atividade	Taxa
José	Musculação	30,00
Pedro	Judô	35,00
Manoel	Judo	35,00



Problemas

- **Anomalia de eliminação** - Suponha a linha do aluno José, bem neste caso, se esta for excluída, **perdemos**, além do nome do aluno, as informações referentes a atividade Musculação, bem como seu valor.

Nome	Atividade	Taxa
José	Musculação	30,00
Pedro	Judô	35,00
Manoel	Judo	35,00



Problemas

- **Anomalia de inserção** - Outro problema ocorre quando a academia implanta um **novo curso** e não podemos inseri-lo até que um aluno tenha a disposição de fazê-lo.

Nome	Atividade	Taxa
José	Musculação	30,00
Pedro	Judô	35,00
Manoel	Judo	35,00



Problemas

- **Anomalia de modificação** - Agora, note que Judô, **está grafado de forma errada** na linha do aluno Manoel. Se uma busca for feita por Judô, só irá aparecer 1 aluno e não 2 alunos.

Nome	Atividade	Taxa
José	Musculação	30,00
Pedro	Judô	35,00
Manoel	Judo	35,00



Normalização – 1FN

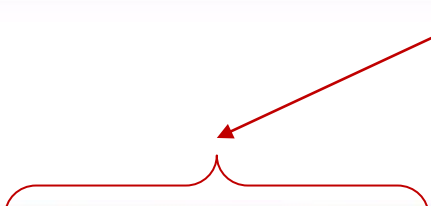
- Para o modelo relacional, a Forma Normal (FN) mais importante é a chamada **1 forma normal (1FN)**.
- *Uma relação está na 1FN quando todos os seus atributos são **Atômicos e Monovalorados**.*
- Um atributo atômico é aquele que não é tratado em partes separadas. (**≠ composto**)
- Um atributo monovalorado é aquele que possui somente um valor. (**≠ multivalorado**)

1ª Forma Normal

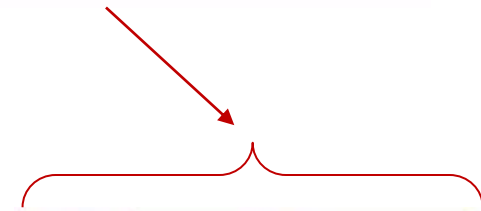
Composto

Multivalorado

<u>CPF</u>	Nome	Sexo	Localização	Telefone
333	Lia	F	Rio de Janeiro, RJ, Brasil	999-555, 777-333
222	Caio	M	Brasília, DF, Brasil	555-888, 333-222
111	Ana	F	São Paulo, SP, Brasil	444-999
444	Beto	M	Porto Alegre, RS, Brasil	888-222



<u>CPF</u>	Nome	Sexo	Cidade	UF	País
333	Lia	F	Rio de Janeiro	RJ	Brasil
222	Caio	M	Brasília	DF	Brasil
111	Ana	F	São Paulo	SP	Brasil
444	Beto	M	Porto Alegre	RS	Brasil



<u>CPF</u>	<u>Telefone</u>
333	999-555
333	777-333
222	555-888
222	333-222
111	444-999
444	888-222



Considere a seguinte relação:

cliente={CPF, nome, endereço, (telefone)}

- A relação Cliente não está na 1FN, porque Nome e Endereço são atributos compostos e telefone é um atributo multivalorado.
- A normalização desta relação resulta nas seguintes relações:
 - *cliente={CPE, nome, sobrenome, rua, numero_casa, bairro}*
 - *cliente_telefone = {CPE, DDD, numero tel}*
- Note que Telefone foi para uma nova relação composta pela chave primária de cliente mais o telefone (decomposto).



Outros exemplos: 1 FN

Curso = {Cod_Curso, descricao_curso, ra_aluno, nome_aluno}

Normalizada:

Curso = {Cod_Curso, descricao_curso}

Aluno = {RA_Aluno, Nome_Aluno, Cod_Curso}

Obras = {cod_Obra, descricao, cidades}

Normalizada:

Obras = {cod_Obra, descricao}

Obras_Cidades = {Cód_Obra, cidades}

Cidades = {codcidade, nome}



Dependência Funcional

- Em uma tabela relacional, diz-se que uma coluna C2 depende funcionalmente de uma coluna C1 (ou que C1 determina C2) quando, em todas as linhas da tabela, para cada valor de C1 aparece o mesmo valor de C2.

C1 \longrightarrow C2

Dependência funcional

Pessoa

CPF	Nome
333	Lia
222	Caio
111	Ana
444	Beto

O atributo nome
depende funcionalmente
de CPF

~~Pessoa~~

CPF	Nome
333	Lia
222	Caio
111	Ana
444	Beto
444	Igor

Pessoa

CPF	Nome
333	Lia
222	Caio
111	Ana
444	Beto
555	Ana

CPF → NOME

Ou seja, CPF determina
Nome



Exemplo de Dependência Funcional

...	Código	...	Salário	...
	E1		10	
	E3		10	
	E1		10	
	E2		5	
	E3		10	
	E2		5	
	E1		10	



Segunda Forma Normal

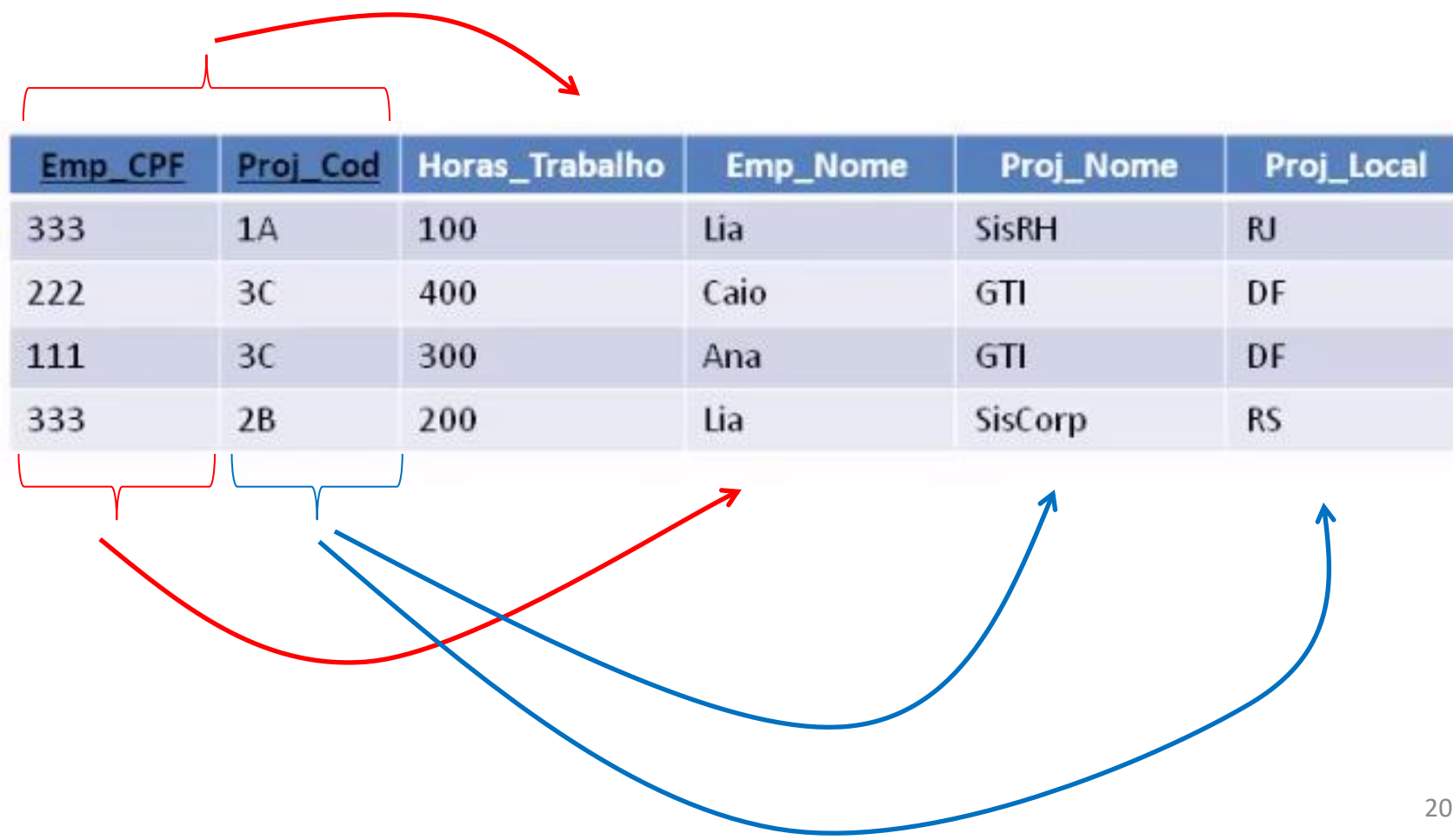
- 2FN: Uma tabela está na 2FN, quando, além de estar na 1FN, não contém dependências funcionais parciais

Turma b

- **Dependências funcionais parciais:** ocorre quando uma coluna **depende apenas de parte** de uma chave primária composta
- Uma relação, para estar na 2FN, não deve possuir **atributo não-chave funcionalmente determinado por “parte” da chave-primária**.

Segunda Forma Normal (2NF)

Tabela Empregado_Projeto



The diagram illustrates functional dependencies for the 'Empregado_Projeto' table. A red bracket groups 'Emp_CPF' and 'Proj_Cod' as a composite key. A red arrow points from this bracket to 'Horas_Trabalho'. A red arrow points from 'Emp_CPF' to 'Emp_Nome'. A blue arrow points from 'Proj_Cod' to 'Proj_Nome'. A blue arrow points from 'Proj_Cod' to 'Proj_Local'. A blue arrow points from 'Emp_CPF' to 'Proj_Local'.

<u>Emp_CPF</u>	<u>Proj_Cod</u>	Horas_Trabalho	Emp_Nome	Proj_Nome	Proj_Local
333	1A	100	Lia	SisRH	RJ
222	3C	400	Caio	GTI	DF
111	3C	300	Ana	GTI	DF
333	2B	200	Lia	SisCorp	RS

Segunda Forma Normal (2NF)



Tabela Empregado_Projeto

<u>Emp_CPF</u>	<u>Proj_Cod</u>	Horas_Trabalho	Emp_Nome	Proj_Nome	Proj_Local
333	1A	100	Lia	SisRH	RJ
222	3C	400	Caio	GTI	DF
111	3C	300	Ana	GTI	DF
333	2B	200	Lia	SisCorp	RS

Tabela Empregado Tabela Empregado_Projeto

<u>CPF</u>	Nome
333	Lia
222	Caio
111	Ana

<u>Emp_CPF</u>	<u>Proj_Cod</u>	Horas_Trabalho
333	1A	100
222	3C	400
111	3C	300
333	2B	200

Tabela Projeto

<u>Cod</u>	Nome	Local
1A	SisRH	RJ
3C	GTI	DF
2B	SisCorp	RS

Segunda Forma Normal (2NF)

- Podemos observar que a integridade referencial foi mantida.

Tabela Empregado

<u>CPF</u>	Nome
333	Lia
222	Caio
111	Ana

Tabela Empregado_Projeto

<u>Emp_CPF</u>	<u>Proj_Cod</u>	Horas_Trabalho
333	1A	100
222	3C	400
111	3C	300
333	2B	200

Tabela Projeto

<u>Cod</u>	Nome	Local
1A	SisRH	RJ
3C	GTI	DF
2B	SisCorp	RS





Qual a importância da normalização?

- Evitar anomalias de:
 - Inserção
 - Atualização
 - Exclusão



Exemplo

- Projeto_Empregado(CodProj,CodEmp, Nome, Categoria, Salario, DataInicio, Descrição, Tipo, TempoAlocação)



Como fazer?

- Criar, caso ainda não exista, uma tabela na 2FN que tenha como PK a parte da chave que é determinante mais as colunas dependentes
- Resultado:
 - Projeto(CodProj, Tipo, Descrição)
 - ProjEmp(CodProj, CodEmp, DataInicio, TempAlocação)
 - Empregado(CodEmp, Nome, Categoria, Salario)



Terceira Forma Normal – (3FN)

- Uma relação está na 3FN se e somente se estiver na 2FN e nenhum atributo não-chave (que não faça parte da chave-primária) for **transitivamente** dependente da chave-primária.
- Está na 3FN a tabela que está na 2FN e não contém **dependências transitivas** (ou indiretas)



Terceira Forma Normal – (3FN)

- Para eliminar redundâncias como: o salário é determinado pela categoria funcional
- **Dependência Transitiva:** coluna não chave depende funcionalmente de outra coluna ou combinação de **colunas não chave**.



Terceira Forma Normal (3FN)

Tabela Empregado_departamento

<u>Emp_CPF</u>	Emp_nome	Emp_sexo	Dep_id	Dep_nome	Dep_Gerente
333	Caio	M	1A	RH	Jorge
222	João	M	2B	Adm	Fábio
111	Pedro	M	3C	TI	Melissa
444	Ana	F	1A	RH	Jorge

Terceira Forma Normal (3FN)

Tabela Empregado_departamento

<u>Emp_CPF</u>	Emp_nome	Emp_sexo	Dep_id	Dep_nome	Dep_Gerente
333	Caio	M	1A	RH	Jorge
222	João	M	2B	Adm	Fábio
111	Pedro	M	3C	TI	Melissa
444	Ana	F	1A	RH	Jorge



Anomalia de inserção

Tabela Empregado_departamento

<u>Emp_CPF</u>	Emp_nome	Emp_sexo	Dep_id	Dep_nome	Dep_Gerente
333	Caio	M	1A	RH	Jorge
222	João	M	2B	Adm	Fábio
111	Pedro	M	3C	TI	Melissa
444	Ana	F	1A	RH	Jorge
555	Maria	F	2B	RH	Fábio

Atenção: O RH acabou ficando com dois nomes de Gerentes

Anomalia de exclusão

Tabela Empregado_departamento

<u>Emp_CPF</u>	Emp_nome	Emp_sexo	Dep_id	Dep_nome	Dep_Gerente
333	Caio	M	1A	RH	Jorge
222	João	M	2B	Adm	Fábio
111	Pedro	M	3C	TI	Melissa
444	Ana	F	1A	RH	Jorge

Atenção: Observe que se o Pedro for excluído nessa tabela, o departamento de TI também será excluído, e isso não deveria acontecer

Anomalia de atualização

Tabela Empregado_departamento

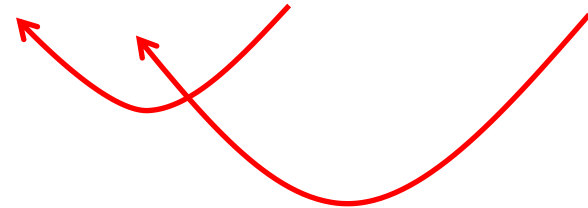
<u>Emp_CPF</u>	Emp_nome	Emp_sexo	Dep_id	Dep_nome	Dep_Gerente
333	Caio	M	1A	RH	Jorge
222	João	M	2B	Adm	Fábio
111	Pedro	M	3C	TI	Melissa
444	Ana	F	1A	RH	Jorge

Atenção: Observe que se tivermos que atualizar o nome do Gerente, deveremos percorrer a tabela buscando o item que necessitaria ser atualizado. Isso poderia gerar um retrabalho e erros.

Terceira Forma Normal – 3FN

Tabela Empregado_departamento

<u>Emp_CPF</u>	Emp_nome	Emp_sexo	Dep_id	Dep_nome	Dep_Gerente
333	Caio	M	1A	RH	Jorge
222	João	M	2B	Adm	Fábio
111	Pedro	M	3C	TI	Melissa
444	Ana	F	1A	RH	Jorge



Atenção: Verificamos que há uma dependência transitiva →
Dep_nome e Dep_gerente dependem **transitivamente** de
Dep_cod.

Terceira Forma Normal – 3FN

Tabela Empregado_departamento

<u>Emp_CPF</u>	Emp_nome	Emp_sexo	Dep_id	Dep_nome	Dep_Gerente
333	Caio	M	1A	RH	Jorge
222	João	M	2B	Adm	Fábio
111	Pedro	M	3C	TI	Melissa
444	Ana	F	1A	RH	Jorge

Tabela Empregado

<u>Emp_CPF</u>	Emp_nome	Emp_sexo	Dep_cod
333	Caio	M	1A
222	João	M	2B
111	Pedro	M	3C
444	Ana	F	1A

Tabela Departamento

Dep_cod	Dep_nome	Dep_Gerente
1A	RH	Jorge
2B	Adm	Fábio
3C	TI	Melissa
1A	RH	Jorge



Como?

- Criar, caso ainda não exista, uma tabela no esquema na 3FN que tenha como PK a coluna da qual há a dependência indireta.
- Copiar a coluna dependente para a tabela criada
- A coluna determinante deve permanecer também na tabela original



Exemplo

Projeto(CodProj, Tipo, Descrição)

Projeto_Empregado(CodProj, CodEmp, DataInicio,
TempAlocacao)

Empregado(CodEmp, Nome, codCat)

Categoria(codCat, Sal)

Exercício



- No contexto de um sistema de controle acadêmico, considere a tabela abaixo:
- Matricula(CodAluno, CodTurma, CodDisciplina, NomeDisciplina, NomeAluno, CodLocalNascAluno, NomeLocalNascAluno)
- * O CodDisciplina identifica a disciplina da turma

2



<u>Codaluno</u>	<u>codTurma</u>	CodDisc	NomeDisc	NomeAlu	CodCidade	NomeCidade

<u>Codaluno</u>	NomeAlu	codCidade

<u>codCidade</u>	NomeCidade

<u>codTurma</u>	<u>CodAluno</u>	codDisc

<u>codDisc</u>	NomeDisc

<u>codTurma</u>	DescriTurma

Exercício

- Considere o esquema

Numero Nota Fiscal:	Data da Nota:
Codigo do Cliente:	
Nome do Cliente:	Telefone Cliente:
Logradouro Cliente	
(rua, bairro, etc):	

Cod.Prod	Desc.Prod	Qtidade	Val.Prod	Val.Total
----------	-----------	---------	----------	-----------



Exercício

- Considere que um analista mapeou a nota para o seguinte esquema relacional:
- *nota_fiscal* = {*num_nota*, *cod_cliente*, *nome_cliente*, *logradouro_cliente*, *telefone_cliente*, *data_nota_fiscal*, (*cod_produto*, *desc_produto*, *qtdade*, *valor_produto*, *valor_total*)}
- Normalize o exercício, passando pelas 3 formas normais.



Exercício

- Considere a tabela abaixo e obtenha a 2FN e 3FN
- ItemVenda(NumeroNF, CodigoTipoProd, NumeroProd, DescricaoProj, DataVenda, CodReg, CodEmp, QtdeItem, PrecoItem, NomeEmp, DescricaoTipoProd)



Bibliografia

- DATE, C.J. Introdução a Sistemas de Bancos de Dados. Ed. Campus, 8ª ed.
- RAMEZ ELMASRI, SHAMKANT B. NAVATHE. Sistemas de Banco de Dados. Ed. Pearson Education, 4ª ed.
- www.devmedia.com.br/cursos/engenharia (acessado em 08/06/2010).