

FIAP GRADUAÇÃO

DATA SCIENCE: BIG DATA, BI & DATA ENGINEERING

BUILDING RELATIONAL DATABASE

PROF. TADEU KANASHIRO proftadeu.kanashiro@fiap.com.br
PROFa. RITA DE CÁSSIA rita@fiap.com.br

PROFESSOR: PERÍODO NOTURNO



- **EXPERIÊNCIA PROFISSIONAL:** Mais de 15 anos atuando na área de tecnologia, focado em análise e estruturação de dados;
- **ÁREAS DE ATUAÇÃO:** Big Data, Analytics Engineering, Business Intelligence e Database Marketing;
- **INDÚSTRIAS:** Saúde, setor imobiliário farmacêutica, fintech, financeiro, Internet, telecomunicações, educação e filantropia.
- **MBA:** Big Data (Data Science);
- **GRADUAÇÃO:** Sistemas de Informações e Gestão Financeira.

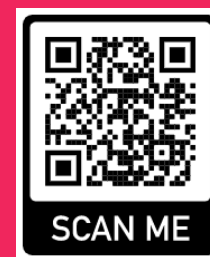
Professor: Tadeu Kanashiro



proftadeu.kanashiro@fiap.com.br



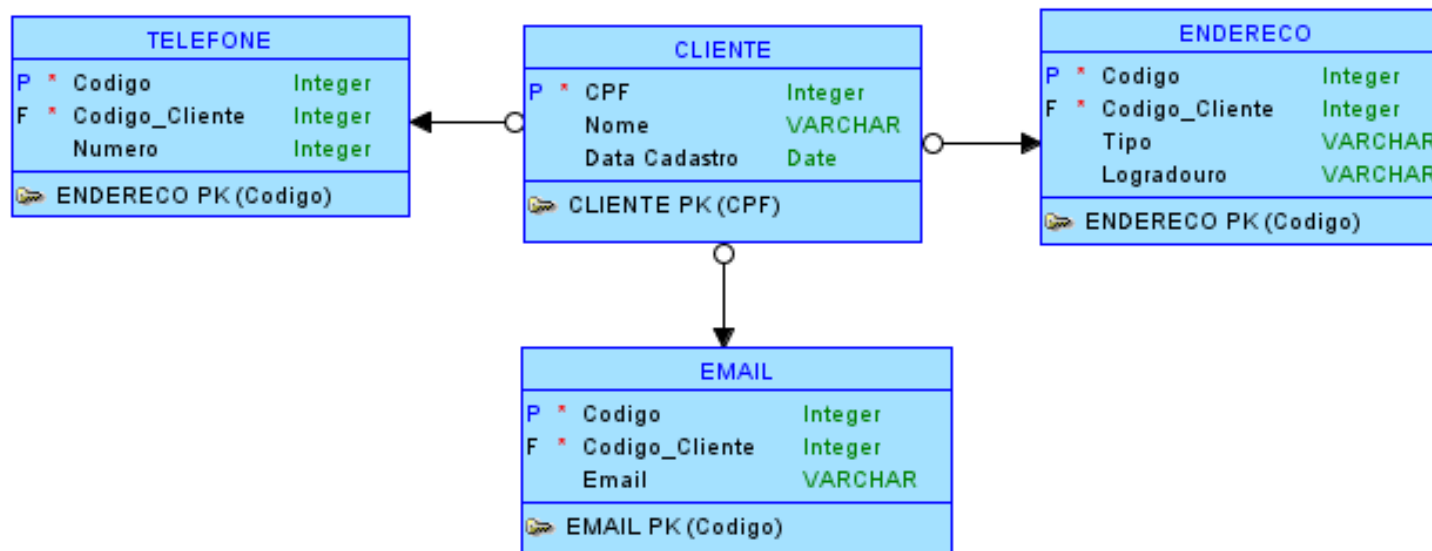
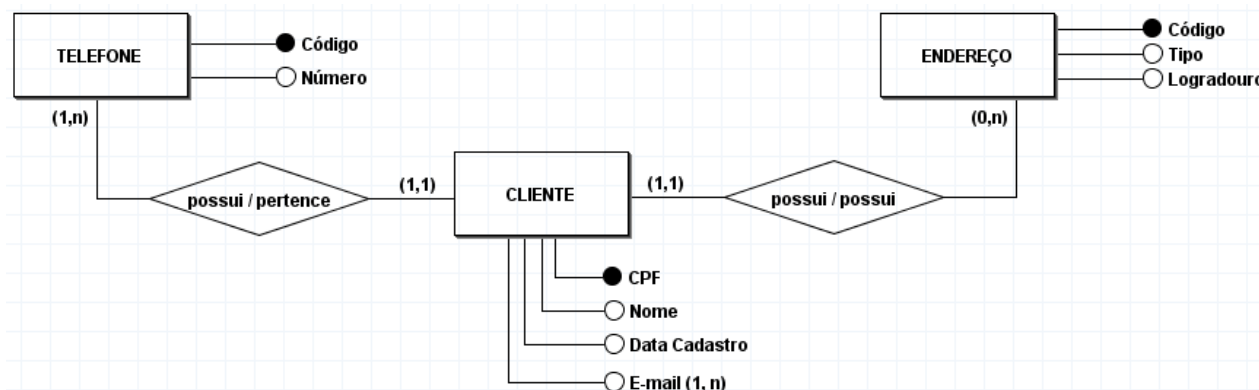
<https://www.linkedin.com/in/tadeukanashiro/>

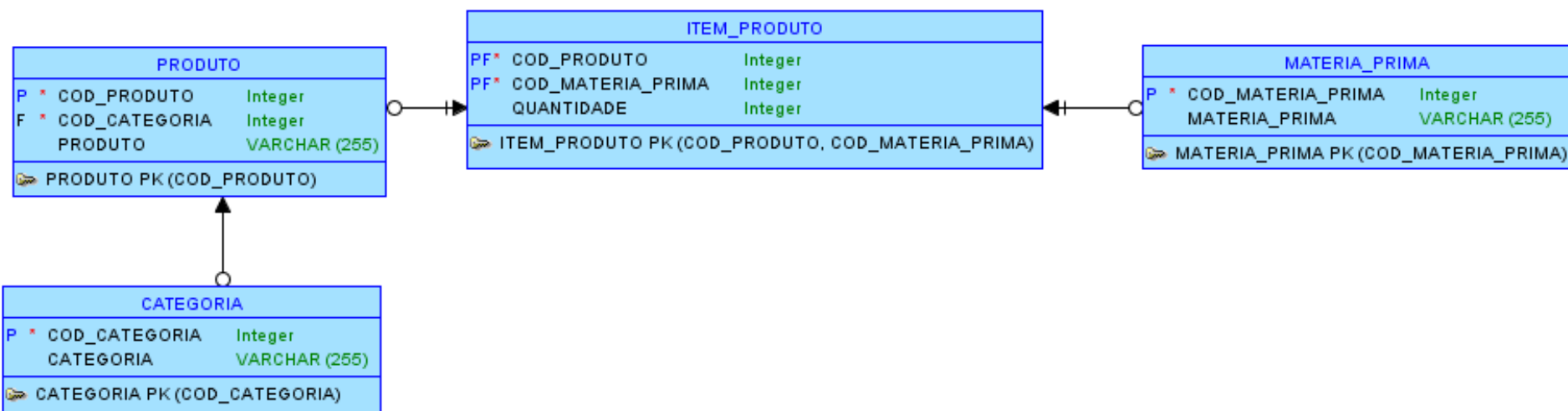
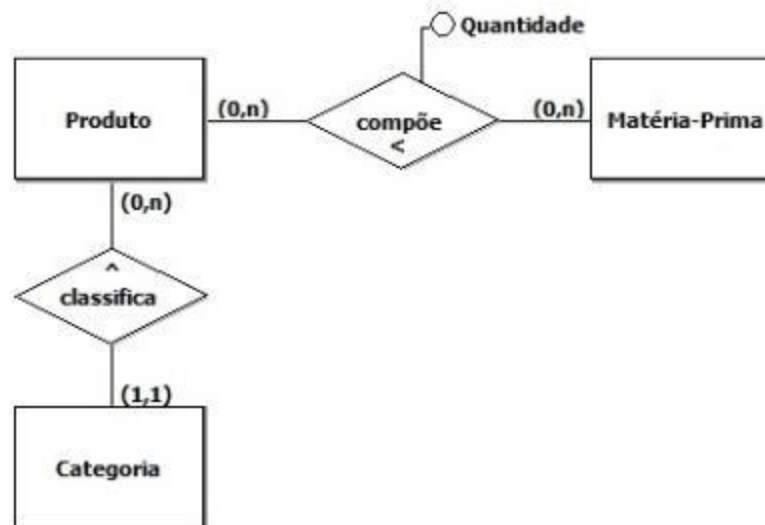


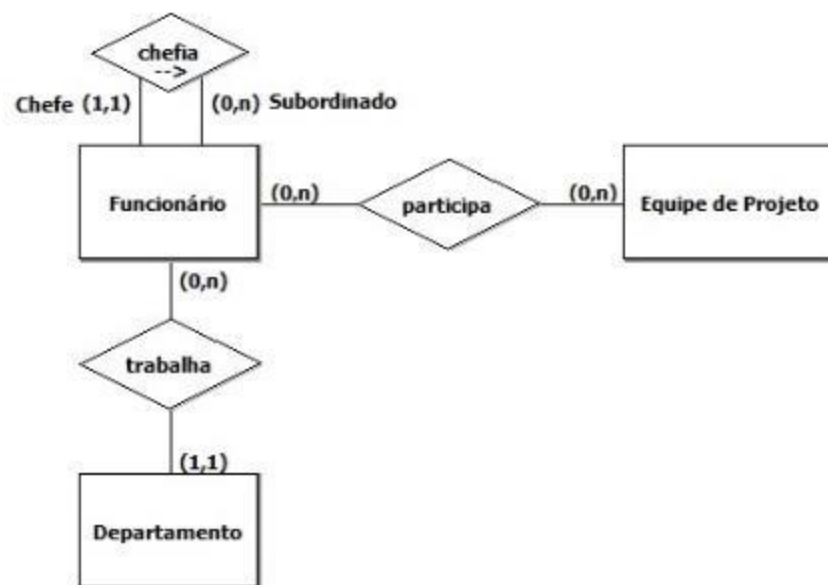
■ AGENDA

- Exercícios da Aula Anterior;
- Dicionário de Dados;
- Exemplo de Dicionário de Dados;
- Tipos de Dados;
- Sugestões Sobre Padrões de Nomenclatura;
- Exercício.

EXERCÍCIOS DA AULA ANTERIOR







FUNCIONARIO		
P *	COD_FUNCIONARIO	Integer
F	COD_CHEFE	Integer
	NOME	VARCHAR (255)
🔑 FUNCIONARIO PK (COD_FUNCIONARIO)		

FUNCIONARIO_EQUIPE		
PF *	COD_FUNCIONARIO	Integer
PF *	COD_EQUIPE	Integer
	HORAS	Integer
🔑 FUNCIONARIO_EQUIPE PK (COD_FUNCIONARIO, COD_EQUIPE)		

EQUIPE_PROJETO		
P *	COD_EQUIPE	Integer
	EQUIPE	VARCHAR (255)
🔑 EQUIPE_PROJETO PK (COD_EQUIPE)		

DEPARTAMENTO		
P *	COD_DEPARTAMENTO	Integer
F *	COD_FUNCIONARIO	Integer
	DEPARTAMENTO	VARCHAR (255)
🔑 DEPARTAMENTO PK (COD_DEPARTAMENTO)		

DICIONÁRIO DE DADOS

I DICIONÁRIO DE DADOS

- O **Dicionário de Dados** (DD), Data Dictionary, **armazena** a **descrição** detalhada **das entidade e atributos** de um projeto de banco de dados, ou seja, **contém** os **metadados** do **projeto**. Em um DD completo, são informados também o volume de dados esperado para aquela entidade, a rotina para limpeza e descarte dos dados e o tempo de retenção dos dados, isto é, **por quanto tempo devem permanecer armazenados**.

I DICIONÁRIO DE DADOS

- A elaboração do **Dicionário de Dados faz parte do projeto da base de dados de um sistema**, da mesma forma que um Diagrama de Modelo de Dados;
- É uma documentação de projeto (geralmente com listas, tabelas e textos descritivos), contendo detalhes e definições sobre diversos componentes da base de dados;
- Obs.: Os **SGBDR's** também **possuem um “dicionário de dados”** ativo (catálogo), em sua arquitetura, **que contém os metadados técnicos** necessários para seu gerenciamento interno. Este tipo de dicionário é atualizado automaticamente e está disponível apenas para consulta (pelo administrador do BD e usuários, conforme seus privilégios).

I DICIONÁRIO DE DADOS

- Um “Dicionário de Dados”, conforme as fases (níveis) de modelagem, também pode descrever elementos dos modelos:
 - **Conceitual:** sistema (escopo), entidades, relacionamentos e atributos;
 - **Lógico:** tabelas e colunas (para bases relacionais), domínios, restrições de integridade (constraints), etc;
 - **Físico:** nomenclatura conforme regras de sintaxe do SGBD escolhido, definições de armazenamento físico, volumetria, índices, etc;
- O diagrama do modelo de dados e o dicionário de dados são complementares entre si.

EXEMPLO DE DICIONÁRIO DE DADOS

EXEMPLO DE DICIONÁRIO DE DADOS

ENTIDADES	
Entidade	Nome da entidade.
Descrição	Texto que conceitua a entidade e os dados representados por ela, além de circunstâncias relevantes ao modelo de dados. Trata-se de descrever o conceito mais próximo da regra de negócios, sem preocupação com as características técnicas que serão implementadas no banco de dados.
Nome da Tabela	Nome que será atribuído à tabela.
Volume de Dados Esperados	<9.999.999> por <Meses, Anos>.
Tempo de Retenção	<99> anos, prazo para manutenção dos dados na base de dados.
Rotina de Limpeza	Texto com a explicação sobre como será o processo de limpeza dos dados, bem como os critérios para eliminar os dados após o tempo de retenção.
ATRIBUTOS	
Atributo	Nome do atributo.
Nome do Campo	Nome do campo.
Tipo de Dado	Tipo de dado que será armazenado.
Tamanho	Quantidade de caracteres que serão armazenados.
Restrição	Regras que deverão ser implementadas na criação das tabelas.
Descrição	Breve descrição do atributo.

EXEMPLO ENTIDADE “MOTORISTA”

ENTIDADES	
Entidade	Motorista.
Descrição	Entidade responsável por armazenar os dados dos motoristas.
Nome da Tabela	TB_Motorista.
Volume de Dados Esperados	Carga inicial de 150 ocorrências (registros) e volume diário de 5 ocorrências.
Tempo de Retenção	Permanente.
Rotina de Limpeza	Não se aplica.

EXEMPLO ATRIBUTOS DO “MOTORISTA”

ATRIBUTOS					
ATRIBUTO	NOME DO CAMPO	TIPO DE DADO	TAMANHO	RESTRIÇÃO	DESCRIÇÃO
Código do Motorista	Cod_Motorista	Numérico inteiro	5	Chave Primária	Registro de matrícula, atribuído quando é cadastrado na empresa e utilizado para identificar cada ocorrência.
Nome do Motorista	Nome_Motorista	Alfanumérico	30	Preenchimento Obrigatório	Nome completo e sem abreviações.
Data de Nascimento do Motorista	Dat_Nasc	Data	Padrão	Data Consistente	Data de nascimento do motorista.
CPF do Motorista	CPF	Numérico inteiro	11	Chave Única	Número do CPF do motorista, sem caracteres de formatação. Item deve ser único para cada ocorrência.
Sexo do Motorista	Sexo	Caractere	1	Aceitar apenas F ou M	O sexo representa o gênero do motorista, devendo ser representado por F (feminino) ou M (Masculino).

TIPOS DE DADOS

I TIPOS DE DADOS

- Os **tipos de dados seguem** um **padrão** do Instituto Nacional Americano de Padrões (American National Standard Institute, **ANSI**), entretanto **podem ocorrer variações** na capacidade de armazenamento e precisão **de** um **SGBD para outro**;
- Alguns tipos de dados definidos pelo padrão ANSI:
 - Alfanuméricos;
 - Numéricos;
 - Data e hora.

I DADOS ALFANUMÉRICOS

- O tipo de dado alfanumérico, também conhecido como literal, **string**, **texto** ou **caractere**, pode armazenar **letras**, **números** e **símbolos especiais**;
- Pode apresentar as seguintes variações:
 - Character(n) ou **Char(n)**: campo com tamanho fixo de n bytes;
 - Character Varying(n) ou **Varchar(n)**: campo com tamanho variável de até n bytes;
 - National Character(n) ou Nchar(n): campo com tamanho fixo de n bytes, que suporta um conjunto de caracteres internacionais (caracteres que não utilizam o alfabeto latino, assim como alfabeto chinês, árabe, grego e etc.);
 - National Character Varying(n) ou Nvarchar(n): campo com tamanho variável de até n bytes, que suporta um conjunto de caracteres internacionais.

CHAR X VARCHAR

- Exemplo de um campo com tamanho fixo, do tipo CHAR ou NCHAR, limitado a 20 posições, ou 20 bytes:

- `nome_matéria char(20)`

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
B	A	N	C	O		D	E		D	A	D	O	S						

- Exemplo de um campo com tamanho variável, do tipo VARCHAR ou NVARCHAR, limitado a 20 posições, ou 20 bytes:

- `nome_matéria varchar(20)`

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
B	A	N	C	O		D	E		D	A	D	O	S						

I DADOS NUMÉRICOS

- O tipo de dado numérico **armazena números inteiros** ou **reais** e pode ser representado da seguinte forma:
 - **Integer**: armazena valores inteiros (aprox. 10 dígitos), sejam eles positivos ou negativos. Sua capacidade varia de acordo com o SGBD utilizado;
 - **Smallint**: possui a mesma função que o Integer (aprox. 5 dígitos), entretanto ocupa menos espaço para armazenamento dos dados.
- O tipo de dado numérico (também conhecido como ponto flutuante, float, devido à existência de casas decimais), seja ele positivo ou negativo, pode ser representado das seguintes formas:
 - **Numeric** (**precisão**, **escala**): onde a precisão determina a quantidade de dígitos significativos (aprox. 15 dígitos) e a escala determina a quantidade de casas decimais;
 - **Decimal** (**precisão**, **escala**): similar ao Numeric, entretanto, em alguns SGBDs, possibilita uma escala maior (aprox. 38 dígitos).

■ NUMERIC(PRECISÃO, ESCALA)

- Inserindo o valor 7.456.123,89:

PRECISÃO E ESCALA	
Numeric	7456123,89
Numeric(9,0)	7.456.124
Numeric(9,1)	7.456.123,9
Numeric(9,2)	7.456.123,89

I DADOS DE DATA E HORA

- Os tipos de dados de data e hora abrangem o armazenamento de data, horário ou data e horário. Existem outros tipos de dados que também possibilitam o armazenamento de fuso horário;
- São representados como:
 - **Date**: possibilita o armazenamento de datas, por exemplo, 01-04-2013;
 - **Time**: possibilita o armazenamento de hora, por exemplo, 11:45:27;
 - **Timestamp**: combina o armazenamento de data e hora, por exemplo, 01-04-2013 11:45:27.

SUGESTÕES SOBRE PADRÕES DE NOMENCLATURA

REGRAS PARA NOMES DE OBJETOS

- Os **nomes** de **objetos** no SGBD Oracle devem respeitar algumas **regras**:
 - De no máximo **128 bytes** para tabelas (bancos **30 bytes**);
 - **Palavras reservadas** (como SELECT, CREATE, UPDATE, TABLE, etc) não podem ser usadas como nomes de objetos;
 - Todos os nomes devem **começar com** uma **letra** de A a Z;
 - Os nomes podem conter letras, números, sublinhado (_), cifrão (\$) ou o símbolo de hash (#);
 - Letras minúsculas são convertidas para letras maiúsculas automaticamente durante a criação do objeto.

Saiba mais: [Documentação Oracle](#)

CASE SENSITIVE PARA NOMES DE OBJETOS

- **Normalmente**, nomes de **objetos não são “case sensitive”** na sintaxe SQL, ou seja, ignora-se a diferença entre minúsculas e maiúsculas (sendo que os nomes são catalogados em maiúsculas no dicionário de dados do SGBD);
- Exemplo:
 - `SELECT * FROM TB_CLIENTES = SELECT * FROM tb_clientes (case insensitive);`
 - `SELECT * FROM TB_CLIENTES ≠ SELECT * FROM tb_clientes (case sensitive).`



CONVENÇÕES DE NOMES DE OBJETOS

- **Prefixos** (sugeridos) para cada tipo de **objeto**:

PREFIXO	TIPO DE OBJETO	COMPOSIÇÃO
TB	Tabela	TB_<nome>
VW	View (visão)	VW_<nome>
IDX	Índices (index)	IDX_<mnemônico_da_tabela>_<nome>
SQ	Sequences	SQ_<mnemônico_da_tabela>
TR	Triggers	TR_<mnemônico_da_tabela>_<nome>
SP	Stored Procedures	SP_<nome>
FC	Functions	FC_<nome>

- **Mnemônicos**: formas **abreviadas** para identificar as tabelas.

CONVENÇÕES DE NOMES DE CONSTRAINT

- **Prefixos (sugeridos)** para cada tipo de **constraint**:

PREFIXO	CONSTRAINT	COMPOSIÇÃO
NN	Not Null	NN_<mnemônico_da_tabela>_<mnemônico_da_coluna>
PK	Primary Key	PK_<mnemônico_da_tabela>
FK	Foreign Key	FK_<mnemônico_tabela_filha>_<mnemônico_tabela_mãe>_<complemento_opcional>
UK	Unique Key	UK_<mnemônico_da_tabela>_<complemento_opcional>
CK	Check	CK_<nome_da_coluna> ou CK_<mnemônico_da_tabela>_<complemento_opcional>

- **Mnemônicos**: formas **abreviadas** para identificar as tabelas.

CONVENÇÕES DE NOMES DE COLUNAS

- **Prefixos (sugeridos)** para cada “classe” de **coluna**:

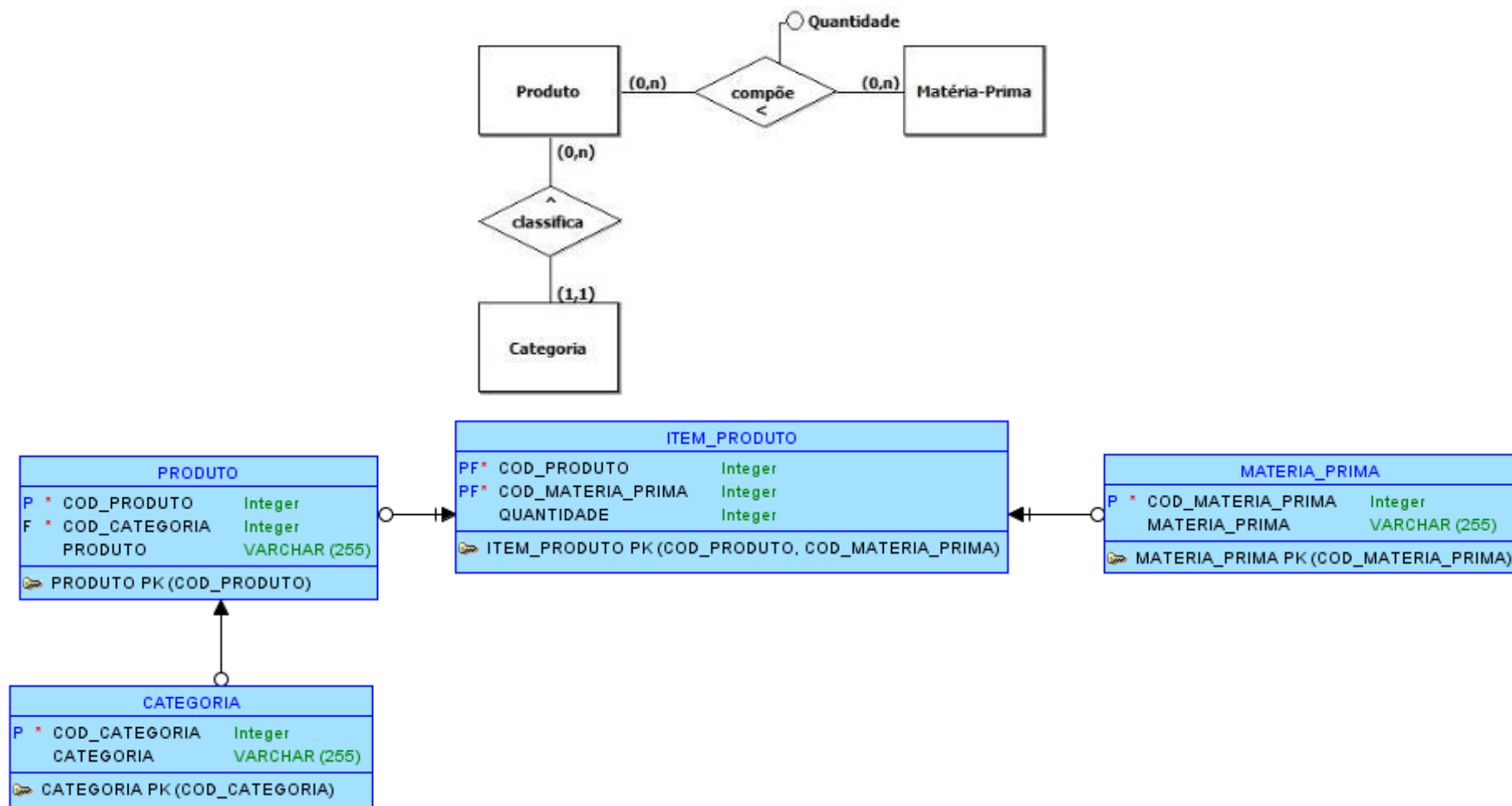
Prefixo	Classe / Natureza
CD	Código
SQ	Sequência
NR	Número
TP	Indicador de Tipo
ST	Situação (status)
SG	Sigla
DS	Descrição
NM	Nome

Prefixo	Classe / Natureza
DT	Data
DH	Data e Horário
VL	Valor
QT	Quantidade
PC	Percentual
MD	Medida
OB	Objeto Binário
TX	Texto (extenso)

I EXERCÍCIO

1. Criar um dicionário de dados para o Modelo criado na aula anterior.
2. Quais regras de negócios são relevantes para serem armazenadas e onde ficarão salvas?
3. Há alguma padronização de nomenclatura de desenvolvimento? Se sim, onde isto poderá ser encontrado?

MODELO CONCEITUAL E LÓGICO



Copyright © 2024 Prof. Tadeu Kanashiro e Prof. André Santos

Todos direitos reservados. Reprodução ou divulgação total ou parcial deste documento é expressamente proibido sem o consentimento formal, por escrito, do Professor (autor).