

Documentação Trabalho Prático 1

Giovanna Andrade Santos - 22052563

José Mateus Córdova Rodrigues - 22052567

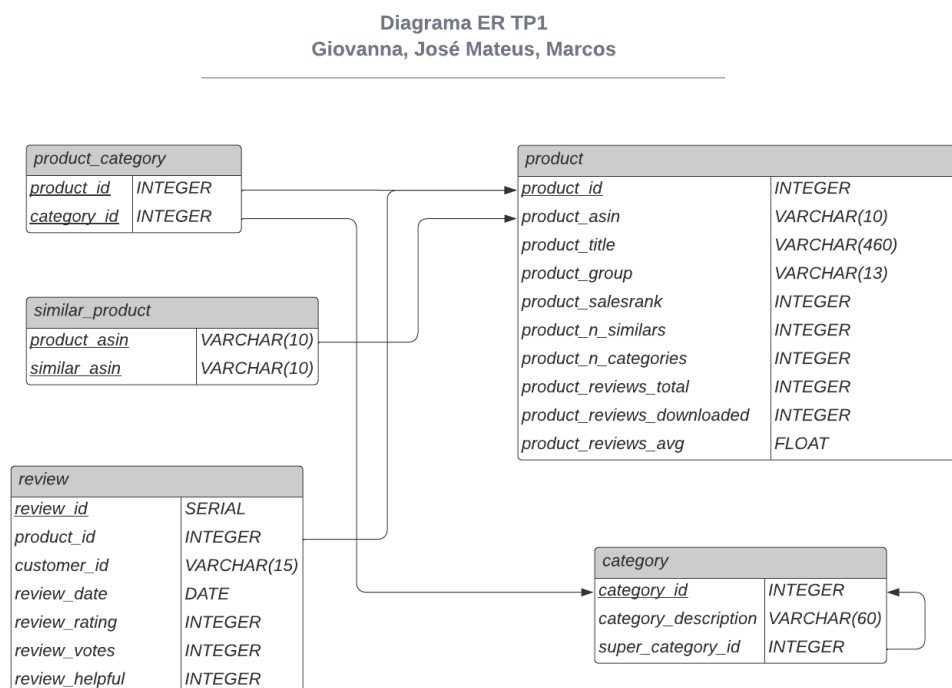
Marcos Avner Pimenta de Lima - 21852448

1. Introdução

Esse relatório apresenta detalhes sobre o primeiro trabalho prático da disciplina de Banco de Dados 1, incluindo aspectos do Banco de Dados formado a partir da base de dados “*Amazon product co-purchasing network metadata*” e construído através de scripts em python juntamente com o SGBD PostgreSQL. Serão apresentados: um diagrama do esquema relacional, um dicionário de dados contendo informações sobre os dados e relações do banco de dados e uma descrição da Forma Normal seguida pelo esquema.

2. Diagrama do Esquema Relacional

Segue abaixo o Diagrama do esquema relacional do Banco de Dados desenvolvido, apresentando as relações com seus atributos, domínios, chaves e referências entre relações:



3. Dicionário de Dados

Dicionário contendo todas as informações de todas as relações que compõem o Banco de Dados:

3.1 Descrição Geral das relações

Tabela	Nº de atributos	Referencia	Descrição
product	10	-	Tabela contendo as características principais de cada produto
product_category	2	product / category	Tabela que apresenta a categoria de um determinado produto
category	3	category	Contém os atributos de cada categoria
similar_product	2	product	Para cada produto, apresenta um produto similar
review	7	product	Apresenta as informações das revisões de cada produto

3.2 Dicionário de cada relação

Tabela	Atributo	Domínio	Restições	Valor Padrão	Descrição
product	product_id	INTEGER	PRIMARY KEY	-	ID do produto
	product_asin	VARCHAR	UNIQUE, NOT NULL	-	Código identificador do produto
	product_title	VARCHAR	-	-	Título do produto
	product_goup	VARCHAR	-	-	Grupo ao qual o produto pertence
	product_salesrank	INTEGER	-	-	Rank de vendas do produto
	product_n_categories	INTEGER	-	0	Número de

					categorias ao qual o produto pertence
	product_reviews_total	INTEGER	-	0	Total de reviews que o produto possui
	product_reviews_down loaded	INTEGER	-	0	Quantidade de reviews baixados para a base de dados
	product_reviews_avg	FLOAT	-	0.0	Média das notas do produto

Tabela	Atributo	Domínio	Restições	Valor Padrão	Descrição
similar_product	product_asin	VARCHAR	PRIMARY KEY, FOREIGN KEY (Referencia tabela product)	-	Código identificador do produto
	similar_asin	VARCHAR	PRIMARY KEY, *tem que ser diferente do product_asin	-	Código identificador do produto similar ao produto referenciado

**Esses atributos devem possuir valores diferentes, pois não há como um produto ser similar a ele mesmo.*

Obs: Anteriormente similar_asin era uma FOREIGN KEY que referenciava a tabela product, mas existem chaves de similares que constam na base mas não existem o respectivo produto. E como nossa base é uma amostra apenas, decidimos manter os dados sem a restrição para ter todas as informações no banco

Tabela	Atributo	Domínio	Restições	Valor Padrão	Descrição
category	category_id	INTEGER	PRIMARY KEY	-	ID da categoria
	category_description	VARCHAR	-	-	Descrição da categoria
	super_category_id	INTEGER	FOREIGN	-	ID da categoria

			KEY (referencia a própria tabela category)		“mãe” à categoria analizada
--	--	--	---	--	-----------------------------------

Tabela	Atributo	Domínio	Restições	Valor Padrão	Descrição
product_category	product_id	INTEGER	PRIMARY KEY, FOREIGN KEY (referncia a tabela product)	-	ID do produto
	category_id	INTEGER	PRIMARY KEY, FOREIGN KEY (referncia a tabela category)	-	ID da categoria

Tabela	Atributo	Domínio	Restições	Valor Padrão	Descrição
review	review_id	SERIAL	PRIMARY KEY	-	ID do review
	product_id	INTEGER	NOT NULL, FOREIGN KEY (referencia a tabela product)	-	ID do produto
	customer_id	VARCHAR	NOT NULL	-	ID do cliente que fez o review
	review_date	DATE	NOT NULL	-	Data em que o review foi feito
	review_rating	INTEGER	-	0	Nota atribuída ao produto no review
	review_votes	INTEGER	-	0	Quantidade de votos de um review

	review_helpful	INTEGER	-	0	Índice que indica quanto o review foi útil
--	----------------	---------	---	---	--

4. Forma Normal

O Banco de Dados elaborado neste trabalho está de acordo com as regras das Primeira, Segunda e Terceira Formas Normais. Já que para uma relação estar de acordo com uma Forma Normal específica, aquela precisa estar de acordo com as formas anteriores.

- **Primeira Forma Normal:** Todos os atributos de todas as relações são atômicos e não são compostos;
- **Segunda Forma Normal:** Todos os atributos não-chaves de todas as relações dependem funcionalmente e totalmente de atributos-chaves. Não há dependência parcial;
- **Terceira Forma Normal:** Nenhum atributo não-chave depende funcionalmente de outro atributo não-chave. Ou seja, todos os atributos não-chaves dependem somente do atributo chave.