

Banco de dados 1

Alcir Heber Castro Figueiredo; Pedro Lucas Bezerra Mendes;

Instituto de Computação – Universidade Federal do Amazonas (UFAM)

– 69080-900 – Amazonas – AM – Brasil

1. Arquivos

Arquivos utilizados para execução deste trabalho:

- **tp1_3.2.py** - Script python que realiza extração dos dados do arquivo de entrada, criação do banco de dados, a criação das tabelas e o povoamento do banco de dados.
- **tp1_3.3.py** - Script python que realiza as consultas do Dashboard.
- **setup.sh** - Script shell que realiza as instalações das dependências necessárias para execução dos arquivos tp1_3.2.py e tp1_3.3.py.

2. Esquema do banco de dados

Nosso banco de dados foi modelado para estar em conformidade com a 3ª Forma Normal (3FN). Isso implica que ele também atende aos requisitos da 1ª Forma Normal (1FN) e da 2ª Forma Normal (2FN), que serão explicados detalhadamente nas seções seguintes.

2.1. 1ª Forma Normal

Na Primeira Forma Normal todos os atributos de uma tabela devem ser atômicos, ou seja, a tabela não deve conter grupos repetidos e nem atributos com mais de um valor.

2.2. 2ª Forma Normal

Antes de mais nada, para estar na 2FN é preciso estar na 1FN. Além disso, todos os atributos não chaves da tabela devem depender unicamente da chave primária (não podendo depender apenas de parte dela).

2.3. 3ª Forma Normal

Para estar na 3FN, é preciso estar na 2FN. Além disso, os atributos não chave de uma tabela devem ser mutuamente independentes e dependentes unicamente e exclusivamente da chave primária (um atributo B é funcionalmente dependente de A se, e somente se, para cada valor de A só existe um valor de B).

2.4. Integridades

Nosso banco de dados foi estruturado para garantir o cumprimento de todas as restrições de integridade nativas. A restrição de domínio é assegurada, garantindo que todos os atributos assumem apenas valores válidos dentro de seus respectivos domínios. Seguimos a restrição de chave, de modo que não existem tuplas duplicadas para as chaves primárias ou colunas definidas como únicas. A restrição de entidade é cumprida, assegurando que as chaves primárias não possam conter valores nulos. Também garantimos a restrição de valores nulos em colunas apropriadas. Por fim, a integridade referencial é mantida através de chaves estrangeiras, garantindo que os relacionamentos entre as tabelas sejam corretamente respeitados.

2.5. Resultado Final

Com base nas informações fornecidas referente ao nosso banco de dados, temos a figura 1 que representa o nosso esquema construído.

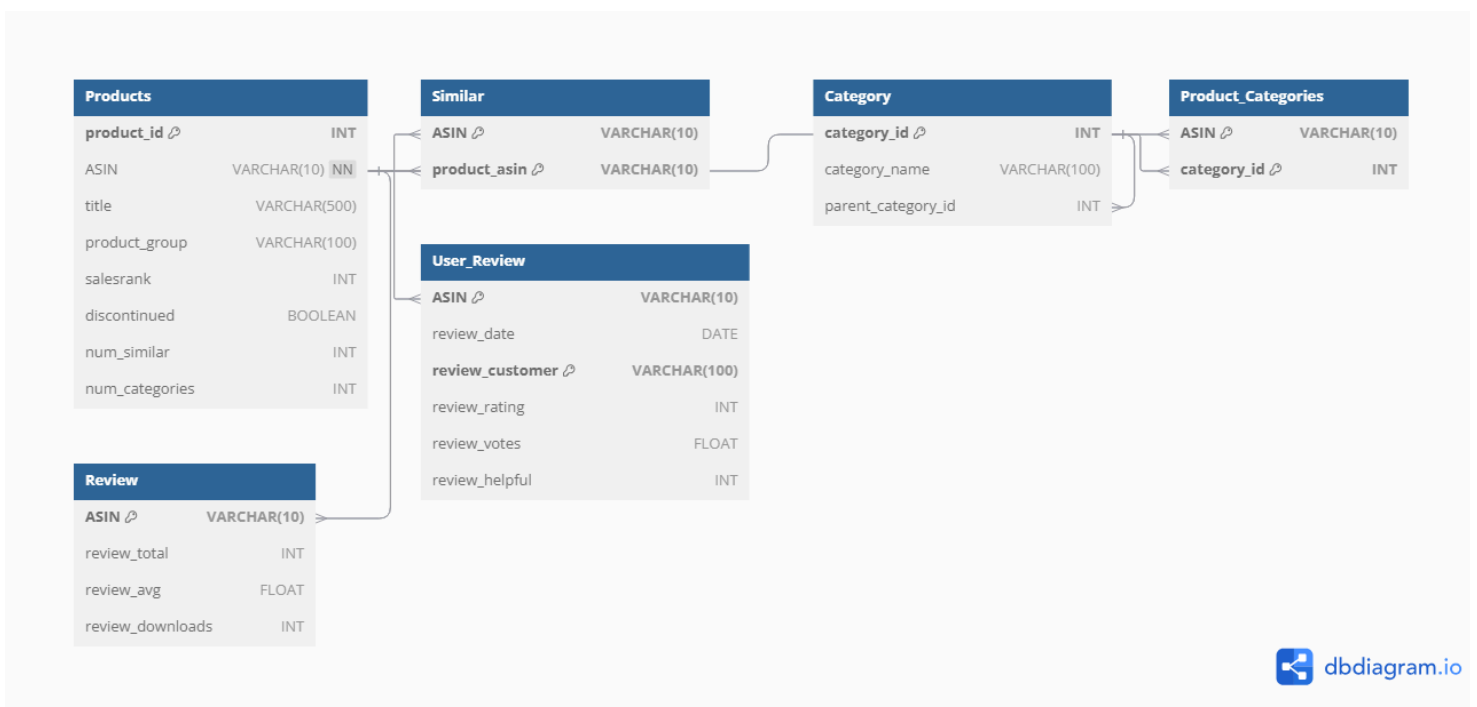


Figura 1: Esquema relacional do nosso banco de dados

3. Dicionário dos atributos

Tabela: Products

Atributo	Tipo de dados	Descrição	Restrição
product_id	INT	Identificador único do produto.	Chave primária (PRIMARY KEY)
ASIN	VARCHAR(10)	Código alfanumérico único do produto.	Único (UNIQUE), não nulo (NOT NULL)
title	VARCHAR(500)	Nome ou título do produto.	
product_group	VARCHAR(100)	Grupo ou categoria principal do produto.	
salesrank	INT	Classificação de vendas do produto.	
discontinued	BOOLEAN	Indica se o produto foi descontinuado.	
num_similar	INT	Número de produtos similares associados.	
num_categories	INT	Número de categorias associadas ao produto.	

Tabela: Similar

Atributo	Tipo de dados	Descrição	Restrição
ASIN	VARCHAR(10)	Código do produto original.	Chave primária composta (PRIMARY KEY (ASIN, product_asin)), Chave estrangeira (FOREIGN KEY) para Products(ASIN)
product_asin	VARCHAR(10)	Código do produto similar.	Chave primária composta (PRIMARY KEY (ASIN, product_asin)), Chave estrangeira (FOREIGN KEY) para Products(ASIN)

Tabela: Category

Atributo	Tipo de dados	Descrição	Restrição
category_id	INT	Identificador único da categoria.	Chave primária (PRIMARY KEY)
category_name	VARCHAR(100)	Nome da categoria.	
parent_category_id	INT	Identificador da categoria pai.	Chave estrangeira (FOREIGN KEY) para Category(category_id), pode ser nulo (ON DELETE SET NULL)

Tabela: Product_Categories

Atributo	Tipo de dados	Descrição	Restrição
ASIN	VARCHAR(10)	Código do produto.	Chave primária composta (PRIMARY KEY (ASIN, category_id)), Chave estrangeira (FOREIGN KEY) para Products(ASIN)
category_id	INT	Identificador da categoria associada ao produto.	Chave primária composta (PRIMARY KEY (ASIN, category_id)), Chave estrangeira (FOREIGN KEY) para Category(category_id)

Tabela: Review

Atributo	Tipo de dados	Descrição	Restrição
ASIN	VARCHAR(10)	Código do produto avaliado.	Chave primária (PRIMARY KEY), Chave estrangeira (FOREIGN KEY) para Products(ASIN)
review_total	INT	Total de reviews de um produto.	
review_avg	FLOAT	Média das avaliações do produto.	
review_downloads	INT	Número de downloads de reviews do produto.	

Tabela: User_Review

Atributo	Tipo de dados	Descrição	Restrição
ASIN	VARCHAR(10)	Código do produto avaliado.	Chave primária composta (PRIMARY KEY (ASIN, review_customer)), Chave estrangeira (FOREIGN KEY) para Products(ASIN)
review_date	DATE	Data da avaliação.	
review_customer	VARCHAR(100)	Nome do cliente que avaliou o produto.	Parte da chave primária (PRIMARY KEY (ASIN, review_customer))
review_rating	INT	Nota da avaliação (1 a 5).	Restrição de valor (CHECK review_rating BETWEEN 1 AND 5)
review_votes	FLOAT	Quantidade de votos recebidos pela avaliação.	
review_helpful	INT	Quantidade de votos úteis que a avaliação recebeu.	