

Trabalho 1 - Banco de dados

Alunos: Allan Carvalho de Aguiar 22153696
Girlana Souza 22152263
Kethelen do Carmo Souza 22152253

1. Esquema do Banco de Dados Relacional

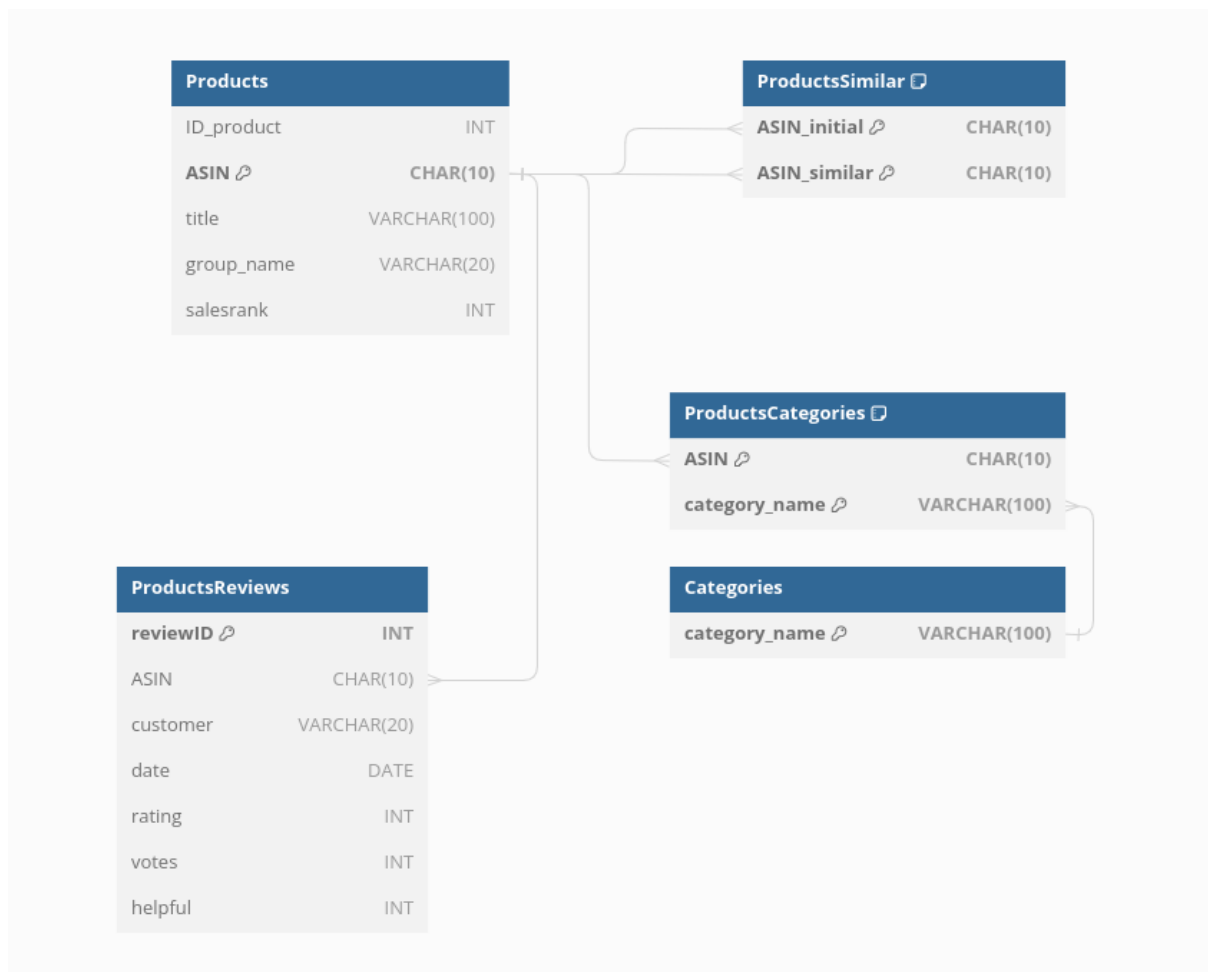


Figura 1: Esquema do banco de dados

2. Dicionário de dados

2.1. Tabelas ou Relações

Tabela 1: Products			
Descrição: Tabela de informações sobre os produtos da Amazon			
Atributos	Domínio	Restrições de Integridade	Descrição
ID_product	INT	Unique	Identificador único do produto
ASIN	CHAR(10)	PRIMARY KEY	Código ASIN do produto
title	VARCHAR(500)	-	Título do produto
group_name	VARCHAR(20)	-	Grupo ao qual o produto pertence.
salesrank	INT	-	Classificação de vendas do produto.

Tabela 2: Categories			
Descrição: Tabela de categorias.			
Atributos	Domínio	Restrições de Integridade	Descrição
category_name	VARCHAR(100)	PRIMARY KEY	Nome da categoria.

Tabela 3: ProductsCategories			
Descrição: Conexão entre produtos e categorias.			
Atributos	Domínio	Restrições de Integridade	Descrição
ASIN	CHAR(10)	PRIMARY KEY, FOREIGN KEY Products(ASIN)	Código ASIN do produto.
category_name	VARCHAR(250)	PRIMARY KEY,	Nome da categoria. Obs: category_name é uma chave estrangeira que

		FOREIGN KEY Categories(category_name)	referencia o atributo category_name da tabela Categories.
--	--	------------------------------------------	-----------------------------------------------------------

Tabela 4: ProductsReviews

Descrição: Tabela das avaliações dos produtos.

Atributos	Domínio	Restrições de Integridade	Descrição
reviewID	INT	PRIMARY KEY	Identificador único da review
ASIN	CHAR(10)	FOREIGN KEY Products(ASIN)	Código ASIN do produto.
customer	VARCHAR(20)	-	Identificador único do cliente.
date	DATE	-	Data da avaliação.
rating	INT	-	Avaliação do produto.
votes	INT	-	Número de votos recebidos.
helpful	INT	-	Número de votos úteis.

Tabela 5: ProductsSimilar

Descrição: Relacionamento entre produtos semelhantes.

Atributos	Domínio	Restrições de Integridade	Descrição
ASIN_initial	CHAR(10)	PRIMARY KEY, FOREIGN KEY Products(ASIN)	Código ASIN do produto inicial.
ASIN_similar	CHAR(10)	PRIMARY KEY, FOREIGN KEY Products(ASIN)	Código ASIN do produto similar.

3. Definição de critérios

3.1. Critério da avaliação positiva

-Uma avaliação positiva é uma avaliação com o valor a partir de 3

3.2. Critério da avaliação útil

-Uma avaliação útil se a quantidade no valor helpful for igual ou maior que a metade dos votos

3.3. Critério da avaliação útil e positiva

- Uma avaliação útil e positiva possui um valor de avaliação a partir de 3 e pelo menos 1 no campo helpful

4. Explicando a consulta (e) “Listar os 10 produtos com a maior média de avaliações úteis positivas por produto”

```
SELECT Products.ID_product, Products.ASIN, Products.title,  
Products.group_name, Products.salesrank, Top.average_rating  
FROM Products  
JOIN(  
    SELECT ASIN, ROUND(avg(rating),2) as average_rating  
    FROM ProductsReviews WHERE helpful >= (votes/2)  
    AND rating >=3  
    GROUP BY ASIN  
    ORDER BY average_rating DESC LIMIT 10  
)  
AS Top  
ON Products.ASIN = Top.ASIN;
```

A condição `helpful >= (votes/2)` delimita que a avaliação é considerada útil se o número de valores em “helpful” for igual ou superior a metade do número total dos votos em “votes”. Isso implica que mais da metade dos votos devem ser marcados como úteis para a avaliação ser considerada como útil.

A condição `rating >= 3` impõe que avaliação deve ser igual ou superior a 3 para ser considerada positiva.

5. Explicando a consulta (f) “Listar a 5 categorias de produto com a maior média de avaliações úteis positivas por produto”

```
SELECT ProductsCategories.category_name,  
ROUND(AVG(ProductsReviews.rating),2)
```

```
AS average_rating
FROM ProductsReviews
JOIN ProductsCategories
ON ProductsReviews.ASIN = ProductsCategories.ASIN
WHERE ProductsReviews.helpful >= (ProductsReviews.votes / 2)
AND ProductsReviews.rating >= 3
GROUP BY ProductsCategories.category_name
ORDER BY average_rating DESC
LIMIT 5;
```

Vamos detalhar os passos da consulta apenas para que a nossa interpretação seja considerada válida:

- Primeiramente, selecionamos o nome da categoria e a média arredondada das avaliações dos produtos nessa categoria, fizemos a junção dos dados das duas tabelas “ProductsReviews” e “ProductsCategories” para combinar as avaliações com as respectivas categorias.
- Realizamos a filtragem das avaliações úteis positivas, ou seja, consideramos que o voto útil (helpful) teria valor igual ou superior a metade da quantidade total dos votos e que a avaliação (rating) seria igual ou superior a 3, esses critérios juntos nos trazem “avaliações úteis e positivas”.
- Após realizado o agrupamento por categoria (todas as avaliações da mesma categoria, pertencem ao mesmo conjunto)
- Fizemos a ordenação em ordem decrescente, e limitamos as cinco maiores médias de avaliações.

Tabelas:

1.

Products

Descrição: Tabela de informações sobre os produtos da Amazon.

Atributos:

ID_product (INT, Unique): Identificador único do produto.

ASIN (CHAR(10), PRIMARY KEY): Código ASIN do produto.

title (VARCHAR(100)): Título do produto.

group_name (VARCHAR(20), FOREIGN KEY): Grupo ao qual o produto pertence.

salesrank (INT): Classificação de vendas do produto.

group_name é uma chave estrangeira que referencia o atributo group_name da tabela Groups.

2.

Groups:

Descrição: Tabela de grupos de produtos.

Atributos:

group_name (VARCHAR(20), PRIMARY KEY): Nome do grupo.

3.

Categories:

Descrição: Tabela de categorias.

Atributos:

category_name (VARCHAR(100), PRIMARY KEY): Nome da categoria.

4.

ProductsCategories:

Descrição: Relacionamento entre produtos e categorias.

Atributos:

ASIN (CHAR(10), PRIMARY KEY): Código ASIN do produto.

category_name (VARCHAR(100), PRIMARY KEY): Nome da categoria.

ASIN é uma chave estrangeira que referencia o atributo ASIN da tabela Products.

category_name é uma chave estrangeira que referencia o atributo category_name da tabela Categories.

5.

ProductsReviews:

Descrição: Tabela das avaliações dos produtos.

Atributos:

ASIN (CHAR(10), PRIMARY KEY): Código ASIN do produto.

customer (VARCHAR(20), PRIMARY KEY): Identificador único do cliente.

date (DATE): Data da avaliação.

rating (INT): Avaliação do produto.

votes (INT): Número de votos recebidos.

helpful (INT): Número de votos úteis.

ASIN é uma chave estrangeira que referencia o atributo ASIN da tabela Products.

6.

ProductsSimilar:

Descrição: Relacionamento entre produtos semelhantes.

Atributos:

ASIN_initial (CHAR(10), PRIMARY KEY): Código ASIN do produto inicial.

ASIN_similar (CHAR(10), PRIMARY KEY): Código ASIN do produto similar.

ASIN_initial é uma chave estrangeira que referencia o atributo ASIN da tabela Products.

ASIN_similar é uma chave estrangeira que referencia o atributo ASIN da tabela Products.