

# Relatório do TP1 de Banco de Dados I

Guilherme Dias Corrêa, Luiza da Costa Caxeixa e Sofia de Castro Sato

<sup>1</sup>Instituto de Computação (IComp) – Universidade Federal do Amazonas (UFAM)  
Av. Rodrigo Otávio, nº 6200, Coroadó I, Manaus – AM, 69080-900

{guilherme.correa, luiza.caxeixa, sofia.sato}@icomp.ufam.edu.br

## 1. Modelo Relacional e Justificativas de Projeto

A modelagem do banco de dados para o trabalho utilizou a técnica Bottom-Up, que se apresentou a mais adequada, uma vez que já tínhamos os dados de entrada. Iniciamos a modelagem identificando os atributos e as relações já existentes nos dados brutos, e prosseguimos para definir as tabelas de forma que os dados de entrada permanecessem consistentes no modelo relacional.

Através da identificação de atributos e relações, nós chegamos nas seguintes tabelas para o esquema relacional: **Produto**, **Similar**, **Review**, **Categoria** e **Categoria\_produto**.

E nessa identificação, conseguimos verificar quais atributos eram relacionados entre si, quais se repetiam e quais complementavam os outros. Dessa forma, foi possível detectar nossas chaves primárias, chaves estrangeiras e nossos relacionamentos entre as 5 tabelas. Junto à definição das chaves, identificamos quais atributos deveriam ser únicos e definimos nossas restrições de unicidade. Além disso, ao estruturar as chaves estrangeiras, optamos por utilizar a cláusula **ON DELETE CASCADE**, garantindo que exclusões em tabelas principais se refletissem automaticamente nas dependentes, mantendo a integridade referencial. Os domínios das tabelas vieram depois, com uma análise detalhada dos dados da entrada, mas, considerando que já tínhamos o que usar para povoar o banco de dados, não houveram problemas.

- **produto:** A tabela produto é a base do esquema e possui como identificador principal o **ASIN** (Amazon Standard Identification Number). Essa tabela não depende de outras, mas serve de referência para várias entidades do esquema, sendo o ponto central para as FKs.
- **produto\_similaridade:** A tabela produto\_similaridade modela os produtos que são apontados como similares. Ela utiliza duas chaves estrangeiras, ambas apontando para a tabela Produto: o *ASIN\_p* e o *ASIN\_s*, e sua chave primária é composta por esses dois campos. Para manter a consistência, aplicamos **ON DELETE CASCADE**, garantindo que, caso um produto seja removido da base, todos os vínculos de similaridade associados a ele sejam automaticamente removidos.
- **review:** A tabela review armazena as avaliações de cada produto. Ela contém um *Review\_id* como chave primária, pois analisando os dados, encontramos reviews idênticas, sendo necessário a criação de um id para diferenciá-las. Cada avaliação está relacionada com um ASIN (sendo uma FK) e um customer, mas por conta da normalização e do formato dos dados, esses atributos não entraram como chave primária. O **ON DELETE CASCADE** também foi aplicado, de forma que a exclusão de um produto remova automaticamente todas as suas avaliações, evitando registros órfãos.

- **categoria:** A tabela categoria representa as categorias hierárquicas dos produtos. Ela contém como chave primária o campo *Categoria\_id*, além de uma chave estrangeira *Pai\_id*, que referencia a própria tabela categoria. Esse autorrelacionamento permite modelar a hierarquia de forma flexível, com qualquer categoria podendo ter subcategorias. O uso de **ON DELETE CASCADE** neste caso garante que, ao excluir uma categoria, todas as suas subcategorias também sejam removidas automaticamente.
- **produto\_categoria:** Essa tabela faz a ligação entre as relações **produto** e **categoria**. Sua chave primária é composta pelo par (*ASIN*, *Categoria\_id*), e ambos os atributos são chaves estrangeiras, referenciando produto e categoria, respectivamente. Também foi definido **ON DELETE CASCADE** nas duas FKs, garantindo que, caso um produto ou uma categoria seja excluído, os vínculos correspondentes sejam automaticamente apagados, preservando a integridade do banco.

## 1.1. Normalização

- **1FN:** Nos dados da entrada, é possível ver algumas informações agrupadas nos campos similares, categorias e reviews. Considerando a regra da 1FN, em que os atributos devem ser atômicos, esses dados foram partidos em um formato que seja condizente com a forma normal, separando assim os produtos similares e as reviews em tabelas separadas do produto, e organizando as categorias em uma tabela própria.
- **2FN:** Continuando a respeitar a 1FN, passamos a observar se os atributos não chaves tinham dependências parciais com a chave. Na tabela **produto** todos os atributos não chave dependem apenas do ASIN (pk). Na tabela **produto\_similaridade** e **produto\_categoria**, os 2 atributos são pk. Nas tabelas **categoria** e **review**, a pk é um id, e todos os atributos dependem apenas dele.
- **3FN:** Por fim, foi verificado se havia dependência transitiva, isto é, se algum atributo não chave dependia de outro atributo não chave. Em todas as tabelas, os atributos não chave estão ligados às suas chaves primárias, sem relações de dependência indireta. Por exemplo, em **produto**, o título e o grupo dependem apenas do ASIN; em **review**, os atributos Review\_date, Rating, Votes e Helpful dependem apenas do id da review. Já em **categoria**, o relacionamento hierárquico é feito por meio de uma chave estrangeira (*Pai\_id*), mas esta ainda depende exclusivamente da chave primária da própria tabela, não caracterizando violação da forma normal.

## 2. Diagrama do Esquema Relacional

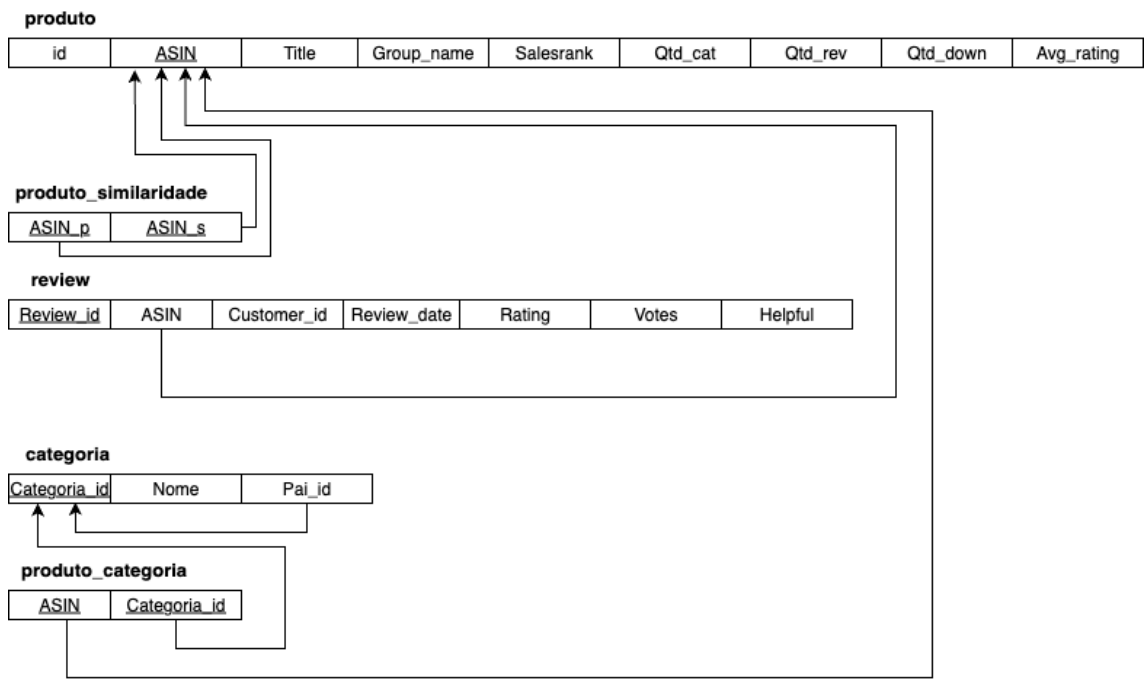


Figura 1. Esquema Relacional do Banco de Dados

## 3. Dicionário de Dados

### 3.1. Tabela Produto

| Atributo   | Tipo        | Chave/Restrição | Descrição                            |
|------------|-------------|-----------------|--------------------------------------|
| id         | SERIAL      | Único           | Identificador numérico interno.      |
| ASIN       | VARCHAR(20) | PK              | Identificador único do produto.      |
| Title      | TEXT        | NOT NULL        | Nome do produto.                     |
| Group_name | VARCHAR(50) | NOT NULL        | Grupo do produto.                    |
| Salesrank  | INTEGER     | -               | Ranking de vendas.                   |
| Qtd_cat    | INTEGER     | -               | Quantidade de categorias associadas. |
| Qtd_rev    | INTEGER     | -               | Quantidade de reviews.               |
| Qtd_down   | INTEGER     | -               | Quantidade de reviews baixadas.      |
| Avg_rating | FLOAT       | -               | Média das avaliações.                |

### 3.2. Tabela Categoria

| Atributo | Tipo | Chave/Restrição | Descrição |
|----------|------|-----------------|-----------|
|----------|------|-----------------|-----------|

|              |      |  |  |
|--------------|------|--|--|
| Categoria_id | INT  | PK   | Identificador da categoria.                  |
| Nome         | TEXT | NOT NULL   | Nome da categoria.                           |
| Pai_id       | INT  | FK → Categoria(Categoria_id),<br>ON DELETE CASCADE | Identificador da categoria pai (hierarquia). |

### 3.3. Tabela Produto\_categoria

| Atributo     | Tipo        | Chave/Restrição  | Descrição                   |
|--------------|-------------|--|-----------------------------|
| ASIN         | VARCHAR(20) | PK, FK → Produto(ASIN),<br>ON DELETE CASCADE           | Identificador do produto.   |
| Categoria_id | INT         | PK, FK → Categoria(Categoria_id),<br>ON DELETE CASCADE | Identificador da categoria. |

### 3.4. Tabela Produto\_similaridade

| Atributo | Tipo        | Chave/Restrição                              | Descrição   |
|----------|-------------|--|---|
| ASIN_p   | VARCHAR(20) | PK, FK → Produto(ASIN),<br>ON DELETE CASCADE | Produto principal.                                  |
| ASIN_s   | VARCHAR(20) | PK, FK → Produto(ASIN),<br>ON DELETE CASCADE | Produto similar.                                    |
| -        | -           | CHECK (ASIN_p <> ASIN_s)                     | Garante que um produto não seja similar a si mesmo. |

### 3.5. Tabela Review

| Atributo    | Tipo        | Chave/Restrição                          | Descrição                          |
|-------------|-------------|--|------------------------------------|
| Review_id   | SERIAL      | PK                                       | Identificador único da review.     |
| ASIN        | VARCHAR(20) | FK → Produto(ASIN),<br>ON DELETE CASCADE | Identificador do produto avaliado. |
| Customer_id | VARCHAR(20) | -  | Identificador do cliente.          |
| Review_date | DATE        | NOT NULL                                 | Data da avaliação.                 |

|         |     |   |   |
|---------|-----|---|---|
| Rating  | INT | - | Nota da avaliação (1 a 5).              |
| Votes   | INT | - | Número de votos recebidos.              |
| Helpful | INT | - | Quantidade de votos considerados úteis. |