



Documentação Trabalho Prático I

Banco de Dados I

Alice Teles Lucena
Igor de Souza Lima
Matheus dos Santos Palheta
Nathália Rodrigues Machado dos Santos

Arquivos

Documentação: tp1_3.1.pdf

- Explicação da Escolha da Terceira Forma Normal (3FN)
- Diagrama do Esquema do Banco de Dados Relacional
- Dicionário de Dados

Script Python 1: tp1_3.2.py

- Criação do Esquema de Banco de Dados
- Extração dos Dados do Arquivos de Entrada
- Povoamento das Relações do BD com os dados extraídos

Script Python 2: tp1_3.3.py

- Execução das Consultas do Dashboard

Recursos Utilizados

Linguagem de Programação: Python

SGBD: PostgreSQL (versão 14.7)

Base de Dados: Amazon Product Co-purchasing Network Metadata

Plataforma de Hospedagem: GitHub

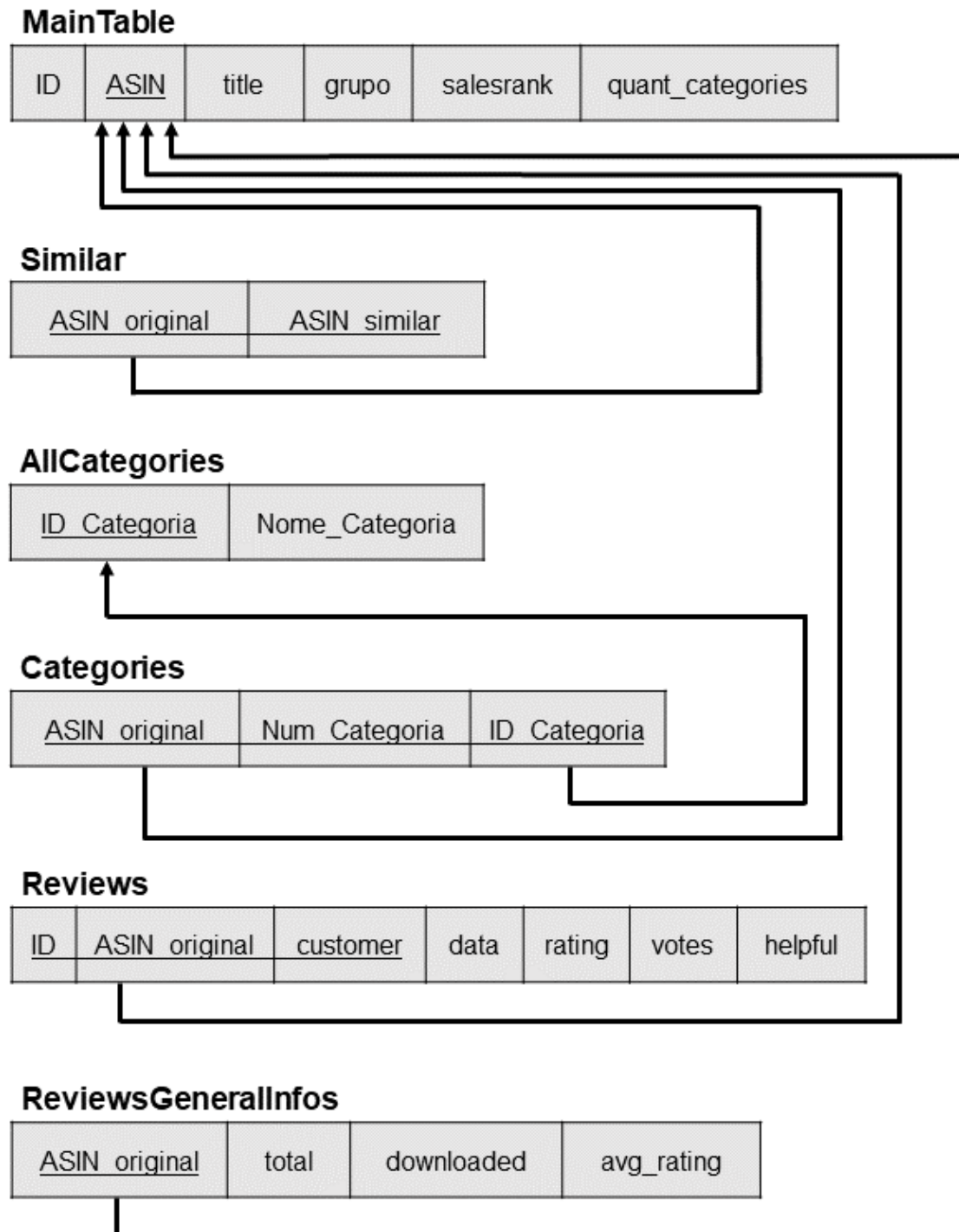
Sobre o Esquema do Banco de Dados

As regras aplicadas no nosso banco de dados seguem a **Terceira Forma Normal (3NF)**. Foi escolhida essa forma porque nosso principal objetivo era minimizar a redundância de dados e garantir a integridade deles sem se preocupar com possíveis anomalias de atualização que poderiam vir a ocorrer na Quarta Forma Normal (4NF) ou na Forma Normal de Boyce-Codd (BCNF).

O fato da 3NF já exigir que cada atributo dependa apenas da chave primária e não de outros atributos não chave elimina repetições e garante que não haja informações duplicadas em diferentes partes do banco de dados. Além disso, a 3NF facilita a manutenção do banco de dados, já que as atualizações são feitas em apenas uma relação. Por exemplo, uma relação que armazena informações sobre categorias do produto e outra que armazena reviews do produto, a 3NF garantirá que cada relação contenha apenas informações exclusivas e que as informações necessárias sejam obtidas por meio de referenciamentos entre as relações.

Por outro lado, a 4NF e a BCNF são mais restritivas, visto que impõem restrições adicionais nas dependências funcionais entre os atributos das relações, e podem exigir que as relações sejam divididas em várias relações menores para atender às suas restrições, o que poderia tornar nosso banco de dados mais complexo e difícil de manter.

Diagrama



Dicionário

Relação MainTable

Atributos

- **ID:** tipo inteiro | não pode ser nulo | é uma chave alternativa
- **ASIN:** tipo string (máximo de 10 caracteres) | **chave primária**
- **title:** tipo string (máximo de 100 caracteres)
- **grupo:** tipo string (máximo de 20 caracteres)
- **salesrank:** tipo inteiro
- **quant_categories:** tipo inteiro

Relação Similar

Atributos

- **ASIN_original:** tipo string (máximo de 10 caracteres) | **chave primária composta**
- **ASIN_similar:** tipo string (máximo de 10 caracteres) | **chave primária composta**

Restrições de Integridade Referencial

- **ASIN_original** é uma chave estrangeira que referencia a chave primária ASIN na relação MainTable.

Outras Restrições

- **ASIN_original** não pode ter o mesmo valor que seu **ASIN_similar**

Relação Categories

Atributos

- **ASIN_original:** tipo string (máximo de 10 caracteres) | **chave primária composta**
- **Num_Categoria:** tipo inteiro | **chave primária composta**
- **ID_Categoria:** tipo inteiro | **chave primária composta**

Restrições de Integridade Referencial

- **ASIN_original** é uma chave estrangeira que referencia a chave primária ASIN na relação *MainTable*
- **ID_Categoria** é uma chave estrangeira que referencia a chave primária ID_Categoria na relação *AllCategories*

Outras Restrições

- **Num_Categoria** não pode ter valores maiores do que o atributo `quant_categories` da relação *MainTable*, contando que ambos os atributos possuem o mesmo ASIN

Relação AllCategories

Atributos

- **ID_Categoria**: tipo inteiro | **chave primária**
- **Nome_Categoria**: tipo string (máximo de 100 caracteres) | não pode ser nulo

Relação Reviews

Atributos

- **ID**: tipo inteiro | se auto incrementa | **chave primária composta**
- **ASIN_original**: tipo string (máximo de 10 caracteres) | **chave primária composta**
- **customer**: tipo string (máximo de 20 caracteres) | **chave primária composta**
- **data**: tipo date
- **rating**: tipo inteiro
- **votes**: tipo inteiro
- **helpful**: tipo inteiro

Restrições de Integridade Referencial

- **ASIN_original** é uma chave estrangeira que referencia a chave primária ASIN na relação *MainTable*

Relação ReviewsGeneralInfos

Atributos

- **ASIN_original**: tipo string (máximo de 10 caracteres) | **chave primária**
- **total**: tipo inteiro
- **downloaded**: tipo inteiro
- **avg_rating**: tipo decimal com 2 dígitos sendo que 1 deles deve estar na parte decimal

Restrições de Integridade Referencial

- **ASIN_original** é uma chave estrangeira que referencia a chave primária ASIN na relação *MainTable*

