

Projeto Final: Design e Implementação de Banco de Dados



Tempo estimado necessário: 90 minutos

Parabéns! Você concluiu os módulos. Agora é hora de colocar suas habilidades à prova. Leia o cenário abaixo.

Cenário

Neste cenário, você foi recentemente contratado como Engenheiro de Dados por uma rede de cafeterias com sede em Nova York que busca expandir nacionalmente abrindo várias franquias. Eles querem otimizar as operações e reformular sua infraestrutura de dados como parte do processo de expansão.

Seu trabalho é projetar seus sistemas de banco de dados relacionais para melhorar a eficiência operacional e facilitar a tomada de decisões baseadas em dados por parte de seus executivos.

Atualmente, os dados deles residem em vários sistemas: software de contabilidade, bancos de dados de fornecedores, sistemas de ponto de venda (POS) e até mesmo planilhas. Você irá revisar os dados em todos esses sistemas e projetar um banco de dados central para armazenar todos os dados. Em seguida, você criará os objetos do banco de dados e os carregará com dados de origem. Finalmente, você criará subconjuntos de dados que seus parceiros de negócios precisam, exportará esses dados e os carregará em bancos de dados de preparação usando vários SGBDs.

Software utilizado neste projeto

Neste projeto, você usará [PostgreSQL Database](#), [MySQL](#). Estes são todos sistemas de gerenciamento de banco de dados relacionais (RDBMS) projetados para armazenar, manipular e recuperar dados de forma eficiente.



Dados usados neste projeto

Neste projeto, você trabalhará com um subconjunto de dados do [Coffee shop sample data](#).

Você usará uma versão modificada dos dados para o projeto, então, para ter sucesso, baixe os arquivos vinculados quando solicitado nas instruções. Você não precisa usar nenhum dado da fonte.

No seu cenário, você trabalhará com dados das seguintes fontes:

- Informações da equipe mantidas em uma planilha na sede (HQ)
- Informações dos pontos de venda mantidas em uma planilha na HQ
- Dados de vendas gerados como um arquivo CSV do sistema de PDV nos pontos de venda
- Dados de clientes gerados como um arquivo CSV de um sistema de gerenciamento de relacionamento com o cliente personalizado
- Informações sobre produtos mantidas em uma planilha exportada do banco de dados do seu fornecedor

Objetivos

Após concluir este laboratório, você será capaz de:

- Identificar entidades
- Identificar atributos
- Criar um diagrama de relacionamento de entidades (ERD) usando a ferramenta ERD do pgAdmin
- Normalizar tabelas
- Definir chaves e relacionamentos
- Criar objetos de banco de dados gerando e executando o script SQL da ferramenta ERD
- Criar uma visão e exportar os dados
- Criar uma visão materializada e exportar os dados
- Importar dados em um banco de dados MySQL usando a ferramenta GUI phpMyAdmin

Tarefa 1: Identificar entidades

O primeiro passo ao projetar um novo banco de dados é revisar os dados existentes e identificar as entidades para o seu novo sistema.

1. A imagem a seguir mostra dados de exemplo de cada fonte com a qual você trabalhará para projetar seu novo banco de dados central. Revise a imagem e identifique as entidades que você planeja criar.

staff					
staff_id	first_name	last_name	position	start_date	location
1	Sue	Tindale	CFO	08/03/2001	HQ
2	Ian	Tindale	CEO	3/8/2001	HQ
3	Marny	Hermione	Roaster	10/24/2007	WH
4	Chelsea	Claudia	Roaster	3/7/2003	WH
5	Alec	Isadora	Roaster	2/4/2008	WH
6	Xena	Rahim	Store Manager	7/24/2016	3
7	Kelsey	Cameron	Coffee Wrangler	10/18/2003	3
8	Hamilton	Emi	Coffee Wrangler	9/2/2005	3
9	Caldwell	Veda	Coffee Wrangler	9/9/2013	3
10	Ima	Winifred	Coffee Wrangler	10/12/2016	3

sales_outlet						
sales_outlet_id	sales_outlet_type	address	city	telephone	postal_code	manager
2	warehouse	164-14 Jamaica Ave	Jamaica	972-871-0402	11432	
3	retail	32-20 Broadway	Long Island City	777-718-3190	11106	6
4	retail	604 Union Street	Brooklyn	619-347-5193	11215	11
5	retail	100 Church Street	New York	343-212-5151	10007	16

sales_transaction								
transaction_id	transaction_date	transaction_time	sales_outlet_id	staff_id	customer_id	product_id	quantity	price
1	27/04/2019	09:53:55	8	42	0	38	2	3.75
1	27/04/2019	09:53:55	8	42	0	84	1	0.8
2	27/04/2019	08:00:34	8	42	0	51	2	3
3	27/04/2019	09:04:58	8	42	0	33	1	3.5
4	27/04/2019	08:48:32	8	42	8232	27	1	3.5
5	27/04/2019	09:21:40	8	45	8223	24	1	3

customer						
customer_id	customer_name	customer_email	customer_since	customer_card_number	birthdate	gender
3001	Kelly Key	Venus@adipiscing.edu	04/01/2017	908-424-2890	29/05/1950	M
3002	Clark Schroeder	Nora@fames.gov	07/01/2017	032-732-6308	30/07/1950	M
3003	Elvis Cardenas	Brianna@tellus.edu	10/01/2017	459-375-9187	30/09/1950	M
3004	Rafael Estes	Ina@non.gov	13/01/2017	576-640-9226	01/12/1950	M
3005	Colin Lynn	Dale@Integer.com	15/01/2017	344-674-6569	01/02/1951	M

product					
product_id	product_category	product_type	product_name	description	price
1	Coffee beans	Organic Beans	Brazilian - Organic	It's like Carnival in a cup. Clean and smooth.	18
2	Coffee beans	House blend Beans	Our Old Time Diner Blend	Our packed blend of beans that is reminiscent of the cup of coffee you used to get at a diner.	18
3	Coffee beans	Espresso Beans	Espresso Roast	Our house blend for a good espresso shot.	14.75
4	Coffee beans	Espresso Beans	Primo Espresso Roast	Our premium single source of hand roasted beans.	20.45
5	Coffee beans	Gourmet Beans	Columbian Medium Roast	A smooth cup of coffee any time of day.	15
6	Coffee beans	Gourmet Beans	Ethiopia	From the home of coffee.	21

- Nota: Você pode baixar uma cópia desta imagem ou abri-la em outra aba do navegador para referência posterior no laboratório.

2. Faça uma lista das entidades que você identificou. Tire uma captura de tela e salve-a como Task1.jpg ou Task1.png.

Tarefa 2: Identificar atributos

Nesta tarefa, você identificará os atributos de uma das entidades que planeja criar.

1. Usando as informações dos dados de exemplo na imagem da Tarefa 1, identifique os atributos da entidade que armazenarão os dados da transação de vendas.
2. Faça uma lista dos atributos da transação de vendas que você identificou. Tire uma captura de tela e salve-a como Task2.jpg ou Task2.png.

Tarefa 3: Criar um ERD

Agora que você definiu alguns de seus atributos e entidades, pode determinar as tabelas e colunas para eles e criar um diagrama de entidade-relacionamento (ERD).

1. Abra um novo terminal no Cloud IDE lado a lado.
2. Use o botão abaixo para iniciar uma sessão de serviço PostgreSQL no Cloud IDE.

[Open and Start PostgreSQL in IDE](#)

3. Use o link do pgAdmin para abrir o pgAdmin em uma nova aba no seu navegador.
4. Crie um novo banco de dados chamado `COFFEE`, visualize os esquemas no novo banco de dados `COFFEE` e, em seguida, inicie um novo projeto ERD.
5. Adicione uma tabela ao ERD para a entidade de transações de vendas usando as informações na tabela a seguir. Considere a convenção de nomenclatura a ser utilizada para que seus colegas possam entender seus dados e garanta que os nomes sejam válidos em outros SGBDs. Use os dados de exemplo mostrados na imagem da Tarefa 1 para determinar os tipos de dados apropriados para cada coluna.

Sales transaction
Transaction id
Date
Time
Sales outlet
Staff
Customer
Product
Quantity
Price

6. Tire uma captura de tela do seu ERD e salve-a como Task3A.png ou Task3A.jpg.
7. Adicione uma tabela ao ERD para a entidade de produto usando as informações na tabela a seguir.

Product
Product id
Category
Type
Product
Description
Price

8. Tire uma captura de tela do seu ERD e salve-a como Task3B.png ou Task3B.jpg.

Tarefa 4: Normalizar tabelas

Ao revisar seu ERD, você nota que ele não está em conformidade com a segunda forma normal. Nesta tarefa, você irá normalizar algumas das tabelas dentro do banco de dados.

1. Revise os dados na tabela de transações de vendas. Observe que a coluna de id da transação não contém valores únicos porque algumas transações incluem múltiplos produtos.
2. Determine quais colunas devem ser armazenadas em uma tabela separada para remover as linhas repetidas e colocar essa tabela na segunda forma normal.
3. Adicione uma nova tabela chamada `sales_detail` ao ERD, defina as colunas na nova tabela e exclua as colunas movidas da tabela de transações de vendas, mantendo uma coluna correspondente em cada uma das duas tabelas para criar um relacionamento entre elas mais tarde.
4. Tire uma captura de tela do seu ERD e salve-a como Task4A.png ou Task4A.jpg.
5. Revise os dados na tabela de produtos. Observe que as colunas de categoria de produto e tipo de produto contêm dados redundantes.
6. Determine quais colunas devem ser armazenadas em uma tabela separada para reduzir dados redundantes e colocar essa tabela na segunda forma normal.
7. Adicione uma nova tabela chamada `product_type` ao ERD, defina as colunas na nova tabela e exclua as colunas movidas da tabela de produtos, mantendo uma coluna correspondente em cada uma das duas tabelas para criar um relacionamento entre elas mais tarde.
8. Tire uma captura de tela do seu ERD e salve-a como Task4B.png ou Task4B.jpg.

Tarefa 5: Definir chaves e relacionamentos

Após normalizar suas tabelas, você pode definir suas chaves primárias e os relacionamentos entre as tabelas no seu diagrama de entidade-relacionamento (ERD).

1. Identifique uma coluna apropriada em cada tabela para ser a chave primária e crie as chaves primárias nas tabelas do seu diagrama de entidade-relacionamento (ERD).
2. Faça uma captura de tela do seu ERD e salve como Task5A.png ou Task5A.jpg.
3. Identifique os relacionamentos entre os seguintes pares de tabelas e, em seguida, crie os relacionamentos no seu ERD:

- sales_detail para sales_transaction
- sales_detail para product
- product para product_type

4. Faça uma captura de tela do seu ERD e salve como Task5B.png ou Task5B.jpg.

Tarefa 6: Criar objetos de banco de dados gerando e executando o script SQL da ferramenta ERD

Agora que seu design está completo, você gerará um script SQL a partir do seu ERD, que poderá usar para criar seu esquema de banco de dados. Para este projeto, você usará um script SQL fornecido para garantir que consiga carregar os dados de exemplo no esquema. Finalmente, você carregará os dados existentes de várias fontes em seu novo esquema de banco de dados.

1. Use a funcionalidade Gerar SQL na ferramenta ERD para criar um script SQL a partir do seu ERD.
2. Baixe o seguinte arquivo `GeneratedScript.sql` para o seu computador local.
 - [GeneratedScript.sql](#)
3. No pgAdmin, abra a ferramenta de consulta, faça o upload e abra o arquivo `GeneratedScript.sql` do seu computador local e, em seguida, execute o script para criar as tabelas definidas no ERD. Verifique se as tabelas existem no esquema público do banco de dados COFFEE agora.
4. Faça uma captura de tela das tabelas mostradas no painel de visualização em árvore no lado esquerdo da página e salve-a como Task6A.png ou Task6A.jpg.
5. Baixe o seguinte arquivo `CoffeeData.sql` para o seu computador local.
 - [CoffeeData.sql](#)
6. No pgAdmin, abra outra instância da ferramenta de consulta, faça o upload e abra o arquivo `CoffeeData.sql` do seu computador local e, em seguida, execute o script para preencher as tabelas que você acabou de criar.
7. No pgAdmin, visualize as primeiras 100 linhas da tabela `sales_detail`.
8. Faça uma captura de tela do painel de Saída de Dados e salve-a como Task6B.png ou Task6B.jpg.

Tarefa 7: Criar uma visualização e exportar os dados

A empresa de folha de pagamento externa solicitou uma lista de funcionários e os locais onde trabalham. Esta lista não deve incluir o CEO ou o CFO que possui a empresa. Nesta tarefa, você criará uma visualização em seu banco de dados PostgreSQL que retorna essas informações e exportará os resultados para um arquivo CSV.

1. No seu banco de dados COFFEE, crie uma nova visualização chamada `staff_locations_view` usando o seguinte SQL:

```
SELECT staff.staff_id,  
       staff.first_name,  
       staff.last_name,  
       staff.location  
FROM staff  
WHERE "position" NOT IN ('CEO', 'CFO');
```

2. Visualize todas as linhas retornadas pela visualização.
3. Salve os resultados da consulta em um arquivo chamado `staff_locations_view.csv` em seu computador local.
4. Tire uma captura de tela da visualização mostrada no painel de visualização em árvore no lado esquerdo da página com os resultados no painel de Saída de Dados, e salve-a como Task7.png ou Task7.jpg.

Tarefa 8: Criar uma visão materializada e exportar os dados

Um consultor de marketing requer acesso aos dados do seu produto em seu banco de dados MySQL para uma campanha de marketing. Você criará uma visão materializada em seu banco de dados PostgreSQL que retorna essas informações e exportará os resultados para um arquivo CSV.

1. No seu banco de dados COFFEE, crie uma nova visão materializada chamada `product_info_m-view` usando o seguinte SQL:

```
SELECT product.product_name, product.description, product_type.product_category  
FROM product  
JOIN product_type  
ON product.product_type_id = product_type.product_type_id;
```

2. Atualize a visão materializada com os dados.
3. Veja todas as linhas retornadas pela visão.
4. Salve os resultados da consulta em um arquivo chamado `product_info_m-view.csv` no seu computador local.
5. Tire uma captura de tela da visão mostrada no painel de visualização em árvore no lado esquerdo da página com os resultados no painel de Saída de Dados e salve-a como `Task8.png` ou `Task8.jpg`.

Tarefa 9: Importar dados de `staff_location` para um banco de dados MySQL

A empresa de folha de pagamento externa pediu que você carregasse as informações de localização da equipe em seu banco de dados MySQL.

1. Clique no botão abaixo para iniciar uma instância MySQL no Cloud IDE.

Open and Start MySQL in IDE

2. Abra o phpMyAdmin em uma nova aba no seu navegador.
3. No phpMyAdmin, crie um novo banco de dados chamado `STAFF_LOCATIONS`, e depois importe as informações de localização salvas no arquivo `staff_locations_view.csv` que você exportou da visualização que criou na Tarefa 7.
4. Explore a nova tabela e depois visualize os dados nela.
5. Faça uma captura de tela do conteúdo da nova tabela e salve-a como `Task9.png` ou `Task9.jpg`.

Tarefa 10: Importar dados de `coffee_shop_products` para um banco de dados MySQL

O consultor de marketing pediu que você carregasse as informações dos produtos para o banco de dados MySQL deles.

1. No phpMyAdmin, crie um novo banco de dados chamado `coffee_shop_products` e, em seguida, importe as informações dos produtos salvas no arquivo `product_info_m-view.csv` de sua visão materializada para uma nova tabela no banco de dados `coffee_shop_products`.
2. Navegue pelo conteúdo da nova tabela.
3. Faça uma captura de tela do conteúdo da nova tabela e salve-a como `Task10.png` ou `Task10.jpg`.

Tarefa 11[Opcional]: Importar dados de `staff_location` para o banco de dados Db2

A empresa externa de folha de pagamento pediu que você carregasse as informações de localização da equipe em seu banco de dados Db2.

1. Em uma nova aba do navegador, acesse cloud.ibm.com/login, faça login usando suas credenciais e, em seguida, abra um console para sua instância Db2 na Nuvem que você criou anteriormente neste curso.
2. Use o recurso Carregar Dados para carregar uma nova tabela chamada `STAFF_LOCATIONS` com as informações de localização da equipe salvas no arquivo `staff_locations_view.csv` que você exportou da visualização que criou na Tarefa 7.
3. Explore a nova tabela e, em seguida, visualize os dados nela.
4. Tire uma captura de tela do conteúdo da nova tabela e salve-a como `Task9.png` ou `Task9.jpg`.

Recapitulação

Neste projeto, você criou os objetos e visualizações que exportou em um banco de dados MySQL.

Autor(es)

- [Lin Joyner](#)
- [Pratiksha Verma](#)

© IBM Corporation. Todos os direitos reservados.