

Projet Final : Conception et Mise en Œuvre de Base de Données



Temps estimé nécessaire : 90 minutes

Félicitations ! Vous avez terminé les modules. Il est maintenant temps de mettre vos compétences à l'épreuve. Lisez le scénario ci-dessous.

Scénario

Dans ce scénario, vous avez récemment été embauché en tant qu'ingénieur des données par une chaîne de cafés basée à New York, qui cherche à se développer au niveau national en ouvrant plusieurs emplacements en franchise. Ils souhaitent rationaliser leurs opérations et rénover leur infrastructure de données dans le cadre de leur processus d'expansion.

Votre mission est de concevoir leurs systèmes de bases de données relationnelles pour améliorer l'efficacité opérationnelle et faciliter la prise de décisions basée sur les données pour leurs dirigeants.

Actuellement, leurs données résident dans plusieurs systèmes : logiciels de comptabilité, bases de données de fournisseurs, systèmes de point de vente (POS) et même des tableurs. Vous examinerez les données dans tous ces systèmes et concevrez une base de données centrale pour regrouper toutes les données. Vous créerez ensuite les objets de la base de données et les chargerez avec les données sources. Enfin, vous créerez des sous-ensembles de données nécessaires à vos partenaires commerciaux, les exporterez et les chargerez dans des bases de données de staging en utilisant plusieurs SGBDR.

Logiciels utilisés dans ce projet

Dans ce projet, vous utiliserez PostgreSQL Database, MySQL. Ce sont tous des systèmes de gestion de bases de données relationnelles (SGBDR) conçus pour stocker, manipuler et récupérer les données de manière efficace.



Données utilisées dans ce projet

Dans ce projet, vous travaillerez avec un sous-ensemble de données provenant des [données d'exemple de café](#).

Vous utiliserez une version modifiée des données pour le projet, donc pour réussir dans le projet, téléchargez les fichiers liés lorsque cela vous est demandé dans les instructions. Vous n'avez pas besoin d'utiliser des données provenant de la source.

Dans votre scénario, vous travaillerez avec des données provenant des sources suivantes :

- Informations sur le personnel conservées dans un tableur au siège (HQ)
- Informations sur les points de vente conservées dans un tableur au HQ
- Données de vente exportées sous forme de fichier CSV depuis le système de point de vente dans les points de vente
- Données clients exportées sous forme de fichier CSV depuis un système de gestion de la relation client personnalisé
- Informations sur les produits maintenues dans un tableur exporté de la base de données de votre fournisseur

Objectifs

Après avoir terminé ce laboratoire, vous serez en mesure de :

- Identifier des entités
- Identifier des attributs
- Créer un diagramme de relation d'entités (ERD) en utilisant l'outil ERD de pgAdmin
- Normaliser des tables
- Définir des clés et des relations
- Créer des objets de base de données en générant et en exécutant le script SQL à partir de l'outil ERD
- Créer une vue et exporter les données
- Créer une vue matérialisée et exporter les données
- Importer des données dans une base de données MySQL en utilisant l'outil GUI phpMyAdmin

Tâche 1 : Identifier les entités

La première étape lors de la conception d’une nouvelle base de données consiste à examiner les données existantes et à identifier les entités pour votre nouveau système.

1. L'image suivante montre des données d'exemple provenant de chaque source avec laquelle vous allez travailler pour concevoir votre nouvelle base de données centrale. Examinez l'image et identifiez les entités que vous prévoyez de créer.

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

- Remarque : Vous pouvez télécharger une copie de cette image ou l’ouvrir dans un autre onglet de votre navigateur pour référence ultérieure dans le laboratoire.

2. Dressez une liste des entités que vous avez identifiées. Prenez une capture d’écran et enregistrez-la sous le nom Task1.jpg ou Task1.png.

Tâche 2 : Identifier les attributs

Dans cette tâche, vous allez identifier les attributs de l’une des entités que vous prévoyez de créer.

1. En utilisant les informations des données d'exemple dans l'image de la Tâche 1, identifiez les attributs de l'entité qui vont stocker les données de transaction de vente.
2. Dressez une liste des attributs de transaction de vente que vous avez identifiés. Prenez une capture d’écran et enregistrez-la sous le nom Task2.jpg ou Task2.png.

Tâche 3 : Créer un ERD

Maintenant que vous avez défini certains de vos attributs et entités, vous pouvez déterminer les tables et colonnes pour ceux-ci et créer un diagramme entité-relation (ERD).

1. Ouvrez un nouveau terminal depuis l’IDE Cloud côte à côte.
2. Utilisez le bouton ci-dessous pour démarrer une session de service PostgreSQL dans l’IDE Cloud.

Open and Start PostgreSQL in IDE
3. Utilisez le lien pgAdmin pour ouvrir pgAdmin dans un nouvel onglet de votre navigateur.
4. Créez une nouvelle base de données nommée C0FFEE, consultez les schémas dans la nouvelle base de données C0FFEE, puis commencez un nouveau projet ERD.
5. Ajoutez une table à l’ERD pour l’entité des transactions de vente en utilisant les informations dans le tableau suivant. Considérez la convention de nommage à utiliser afin que vos collègues puissent comprendre vos données et assurez-vous que les noms sont valides dans d’autres SGBDR. Utilisez les données d’exemple montrées dans l’image de la Tâche 1 pour déterminer les types de données appropriés pour chaque colonne.

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

6. Prenez une capture d’écran de votre ERD et enregistrez-la sous le nom Task3A.png ou Task3A.jpg.
7. Ajoutez une table à l’ERD pour l’entité produit en utilisant les informations dans le tableau suivant.

Product
Product id
Category
Type
Product
Description
Price

8. Prenez une capture d’écran de votre ERD et enregistrez-la sous le nom Task3B.png ou Task3B.jpg.

Tâche 4 : Normaliser les tables

Lorsque vous examinez votre ERD, vous remarquez qu’il ne respecte pas la deuxième forme normale. Dans cette tâche, vous normaliserez certaines des tables dans la base de données.

1. Examinez les données dans la table des transactions de vente. Notez que la colonne d’identifiant de transaction ne contient pas de valeurs uniques car certaines transactions incluent plusieurs produits.
2. Déterminez quelles colonnes devraient être stockées dans une table séparée pour éliminer les lignes répétées et pour mettre cette table en deuxième forme normale.
3. Ajoutez une nouvelle table nommée sales_detail à l’ERD, définissez les colonnes de la nouvelle table et supprimez les colonnes déplacées de la table des transactions de vente, en laissant une colonne correspondante dans chacune des deux tables pour créer une relation entre elles plus tard.
4. Prenez une capture d’écran de votre ERD et enregistrez-la sous le nom Task4A.png ou Task4A.jpg.
5. Examinez les données dans la table des produits. Notez que les colonnes de catégorie de produit et de type de produit contiennent des données redondantes.
6. Déterminez quelles colonnes devraient être stockées dans une table séparée pour réduire les données redondantes et pour mettre cette table en deuxième forme normale.

7. Ajoutez une nouvelle table nommée `product_type` à l'ERD, définissez les colonnes de la nouvelle table et supprimez les colonnes déplacées de la table des produits, en laissant une colonne correspondante dans chacune des deux tables pour créer une relation entre elles plus tard.
8. Prenez une capture d'écran de votre ERD et enregistrez-la sous le nom `Task4B.png` ou `Task4B.jpg`.

Tâche 5 : Définir les clés et les relations

Après avoir normalisé vos tables, vous pouvez définir leurs clés primaires et les relations entre les tables dans votre diagramme de relation d'entité (ERD).

1. Identifiez une colonne appropriée dans chaque table pour être une clé primaire et créez les clés primaires dans les tables de votre diagramme de relation d'entité (ERD).
2. Prenez une capture d'écran de votre ERD et enregistrez-la sous le nom `Task5A.png` ou `Task5A.jpg`.
3. Identifiez les relations entre les paires de tables suivantes, puis créez les relations dans votre ERD :
 - `sales_detail` à `sales_transaction`
 - `sales_detail` à `product`
 - `product` à `product_type`
4. Prenez une capture d'écran de votre ERD et enregistrez-la sous le nom `Task5B.png` ou `Task5B.jpg`.

Tâche 6 : Créer des objets de base de données en générant et en exécutant le script SQL à partir de l'outil ERD

Maintenant que votre conception est terminée, vous allez générer un script SQL à partir de votre ERD, que vous pourrez utiliser pour créer votre schéma de base de données. Pour ce projet, vous utiliserez ensuite un script SQL donné pour vous assurer que vous pouvez charger les données d'exemple dans le schéma. Enfin, vous chargerez les données existantes provenant de diverses sources dans votre nouveau schéma de base de données.

1. Utilisez la fonctionnalité Générer SQL dans l'outil ERD pour créer un script SQL à partir de votre ERD.
2. Téléchargez le fichier suivant `GeneratedScript.sql` sur votre ordinateur local.
 - [GeneratedScript.sql](#)
3. Dans pgAdmin, ouvrez l'outil de requête, téléchargez et ouvrez le fichier `GeneratedScript.sql` depuis votre ordinateur local, puis exécutez le script pour créer les tables définies dans l'ERD. Vérifiez que les tables existent maintenant dans le schéma public de la base de données COFFEE.
4. Prenez une capture d'écran des tables affichées dans le panneau de visualisation arborescente sur le côté gauche de la page et enregistrez-la sous le nom `Task6A.png` ou `Task6A.jpg`.
5. Téléchargez le fichier suivant `CoffeeData.sql` sur votre ordinateur local.
 - [CoffeeData.sql](#)
6. Dans pgAdmin, ouvrez une autre instance de l'outil de requête, téléchargez et ouvrez le fichier `CoffeeData.sql` depuis votre ordinateur local, puis exécutez le script pour peupler les tables que vous venez de créer.
7. Dans pgAdmin, visualisez les 100 premières lignes de la table `sales_detail`.
8. Prenez une capture d'écran du panneau de sortie de données et enregistrez-la sous le nom `Task6B.png` ou `Task6B.jpg`.

Tâche 7 : Créer une vue et exporter les données

La société de paie externe a demandé une liste des employés et des lieux où ils travaillent. Cette liste ne doit pas inclure le PDG ou le CFO qui possède la société. Dans cette tâche, vous allez créer une vue dans votre base de données PostgreSQL qui retourne ces informations et exporter les résultats dans un fichier CSV.

1. Dans votre base de données COFFEE, créez une nouvelle vue nommée `staff_locations_view` en utilisant le SQL suivant :

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

- Affichez toutes les lignes retournées par la vue.
- Enregistrez les résultats de la requête dans un fichier nommé `staff_locations_view.csv` sur votre ordinateur local.
- Prenez une capture d'écran de la vue affichée dans le panneau de vue arborescente sur le côté gauche de la page avec les résultats dans le panneau de sortie de données, et enregistrez-la sous le nom `Task7.png` ou `Task7.jpg`.

Tâche 8 : Créer une vue matérialisée et exporter les données

Un consultant en marketing nécessite un accès à vos données produit dans sa base de données MySQL pour une campagne marketing. Vous allez créer une vue matérialisée dans votre base de données PostgreSQL qui renvoie ces informations et exporter les résultats dans un fichier CSV.

- Dans votre base de données `C0FFEE`, créez une nouvelle vue matérialisée nommée `product_info_m-view` en utilisant le SQL suivant :

```
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;
```

- Rafraîchissez la vue matérialisée avec les données.
- Affichez toutes les lignes retournées par la vue.
- Enregistrez les résultats de la requête dans un fichier nommé `product_info_m-view.csv` sur votre ordinateur local.
- Prenez une capture d'écran de la vue affichée dans le panneau de l'arborescence à gauche de la page avec les résultats dans le panneau de sortie des données, et enregistrez-la sous le nom `Task8.png` ou `Task8.jpg`.

Tâche 9 : Importer les données `staff_location` dans une base de données MySQL

La société de paie externe vous a demandé de télécharger les informations sur les emplacements du personnel dans leur base de données MySQL.

- Cliquez sur le bouton ci-dessous pour démarrer une instance MySQL dans le Cloud IDE.

Open and Start MySQL in IDE

- Ouvrez phpMyAdmin dans un nouvel onglet de votre navigateur.
- Dans phpMyAdmin, créez une nouvelle base de données nommée `STAFF_LOCATIONS`, puis importez les informations de localisation enregistrées dans le fichier `staff_locations_view.csv` que vous avez exporté de la vue que vous avez créée dans la Tâche 7.
- Explorez la nouvelle table, puis visualisez les données qu'elle contient.
- Prenez une capture d'écran du contenu de la nouvelle table et enregistrez-la sous le nom `Task9.png` ou `Task9.jpg`.

Tâche 10 : Importer les données `coffee_shop_products` dans une base de données MySQL

Le consultant en marketing vous a demandé de télécharger les informations sur les produits dans leur base de données MySQL.

- Dans phpMyAdmin, créez une nouvelle base de données nommée `coffee_shop_products`, puis importez les informations sur les produits enregistrées dans le fichier `product_info_m-view.csv` de votre vue matérialisée dans une nouvelle table de la base de données `coffee_shop_products`.
- Parcourez le contenu de la nouvelle table.
- Prenez une capture d'écran du contenu de la nouvelle table et enregistrez-la sous le nom `Task10.png` ou `Task10.jpg`.

Tâche 11[Optionnel] : Importer les données staff_location dans la base de données Db2

La société de paie externe vous a demandé de télécharger les informations de localisation du personnel dans leur base de données Db2.

1. Dans un nouvel onglet de votre navigateur, allez sur cloud.ibm.com/login, connectez-vous avec vos identifiants, puis ouvrez une console pour votre instance Db2 sur le Cloud que vous avez créée plus tôt dans ce cours.
2. Utilisez la fonction Charger des données pour charger une nouvelle table nommée STAFF_LOCATIONS avec les informations de localisation du personnel enregistrées dans le fichier staff_locations_view.csv que vous avez exporté de la vue que vous avez créée dans la Tâche 7.
3. Explorez la nouvelle table puis visualisez les données qu'elle contient.
4. Prenez une capture d'écran du contenu de la nouvelle table et enregistrez-la sous le nom Task9.png ou Task9.jpg.

Récapitulatif

Dans ce projet, vous avez créé les objets et les vues que vous avez exportés dans une base de données MySQL.

Author(s)

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

© IBM Corporation. Tous droits réservés.