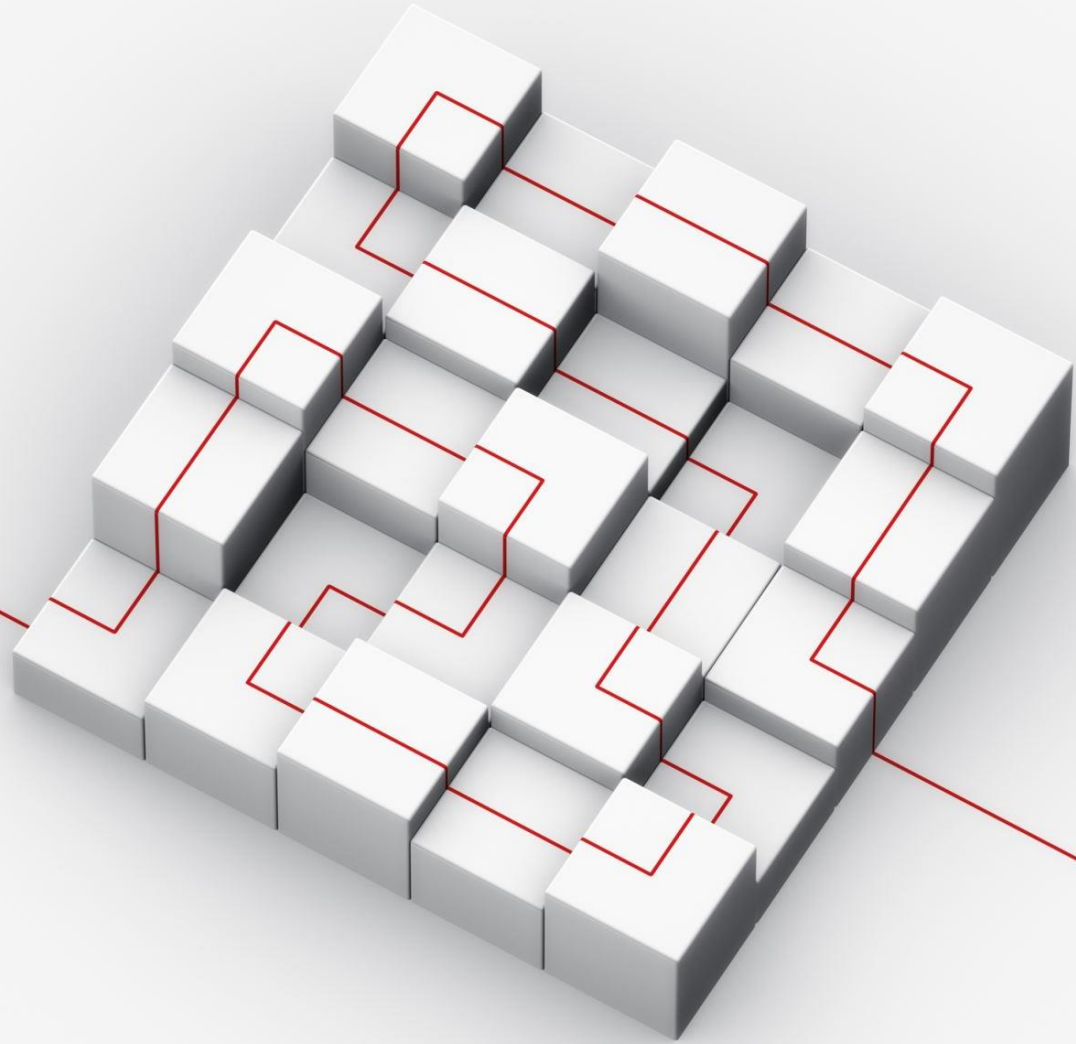




ENTREPÔT DE DONNÉES

Gabriel Goulet

Sébastien Borgne



Sommaire

- Contexte
- Schéma Base De Données
- Création des tables
- Définition des contraintes
- Trigger
- Table de faits
- Récupération & visualisation des KPIs
- Implémentation de rôles sur la base
- Masquage de données sensibles

Contexte:

- Organisation d'une base de données pour une plateforme d'ecommerce. Le but est de créer un data warehouse prêt pour un travail analytique.

A partir des données à disposition:

- Définition des KPIs
- Conception BDD
- Table de faits
- Récupération & visualisation des KPIs
- Sécurisation de la base de données



Outils utilisés

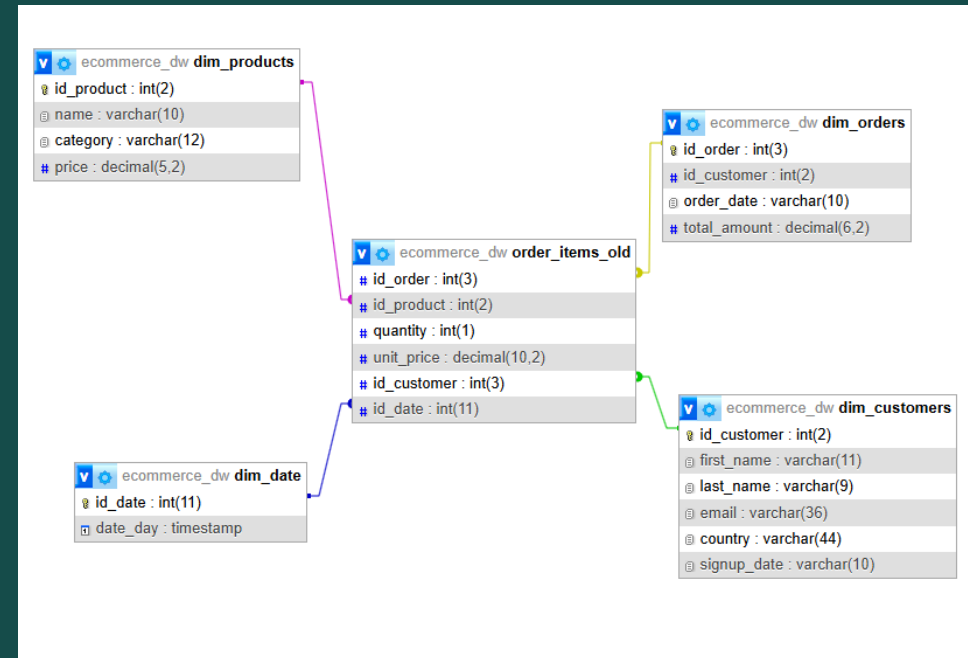
- Tableau
- PhpMyAdmin
- SQL
- MariaDB

Import des données & création des tables

Fichiers de données fournis



Base de données



Définition des contraintes & clés

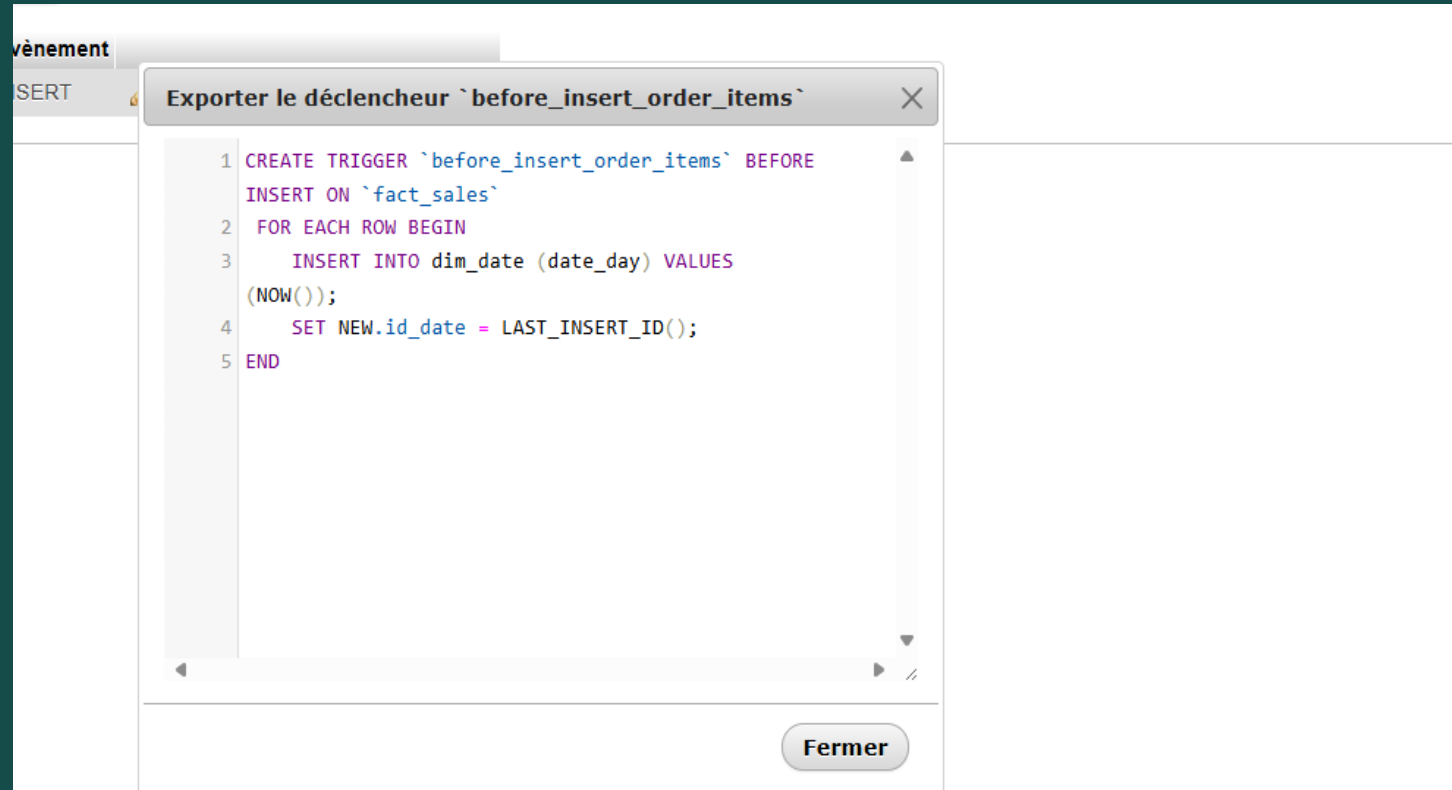
Structure de table		Vue relationnelle							
#	Nom	Type	Interclassement	Attributs	Null	Valeur par défaut	Commentaires	Extra	Action
<input type="checkbox"/>	1 id_customer	int(2)			Non	Aucun(e)		AUTO_INCREMENT	Modifier Supprimer Plus
<input type="checkbox"/>	2 first_name	varchar(11)	utf8_general_ci		Oui	NULL			Modifier Supprimer Plus
<input type="checkbox"/>	3 last_name	varchar(9)	utf8_general_ci		Oui	NULL			Modifier Supprimer Plus
<input type="checkbox"/>	4 email	varchar(36)	utf8_general_ci		Oui	NULL			Modifier Supprimer Plus
<input type="checkbox"/>	5 country	varchar(44)	utf8_general_ci		Oui	NULL			Modifier Supprimer Plus
<input type="checkbox"/>	6 signup_date	varchar(10)	utf8_general_ci		Oui	NULL			Modifier Supprimer Plus

```
1 ALTER TABLE fact_sales
2 ADD CONSTRAINT fk_customer
3 FOREIGN KEY (id_customer)
4 REFERENCES dim_customers(id_customer);
```

Trigger pour la dimension date

Automatise la gestion des dates :

- Insère automatiquement la date actuelle dans dim_date
- Lie cette date à chaque entrée dans la table de faits



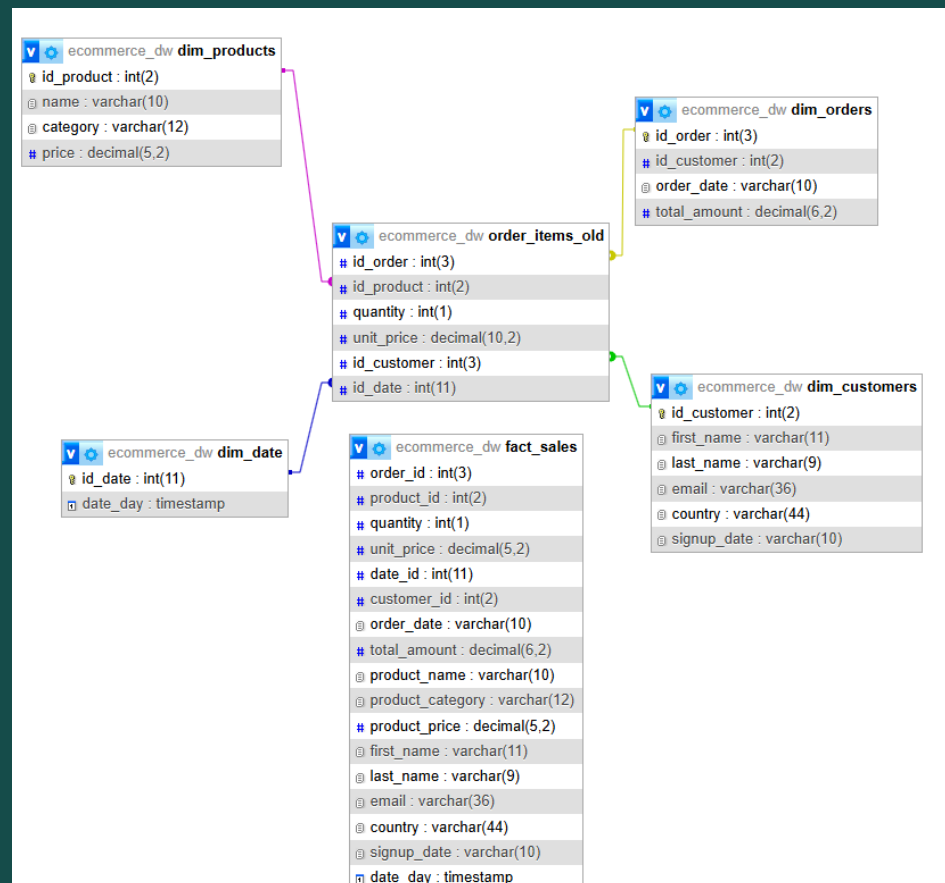
The screenshot shows a database management interface with a dialog box titled "Exporter le déclencheur `before_insert_order_items`". The dialog contains a text area with the following SQL code:

```
1 CREATE TRIGGER `before_insert_order_items` BEFORE
2 INSERT ON `fact_sales`
3   FOR EACH ROW BEGIN
4     INSERT INTO dim_date (date_day) VALUES
5     (NOW());
6     SET NEW.id_date = LAST_INSERT_ID();
7 END
```

At the bottom right of the dialog is a button labeled "Fermer".

Création de la table de fait fact_sales

Aggrège toute les données grâce à un trigger à l'insertion
Utile pour analytique



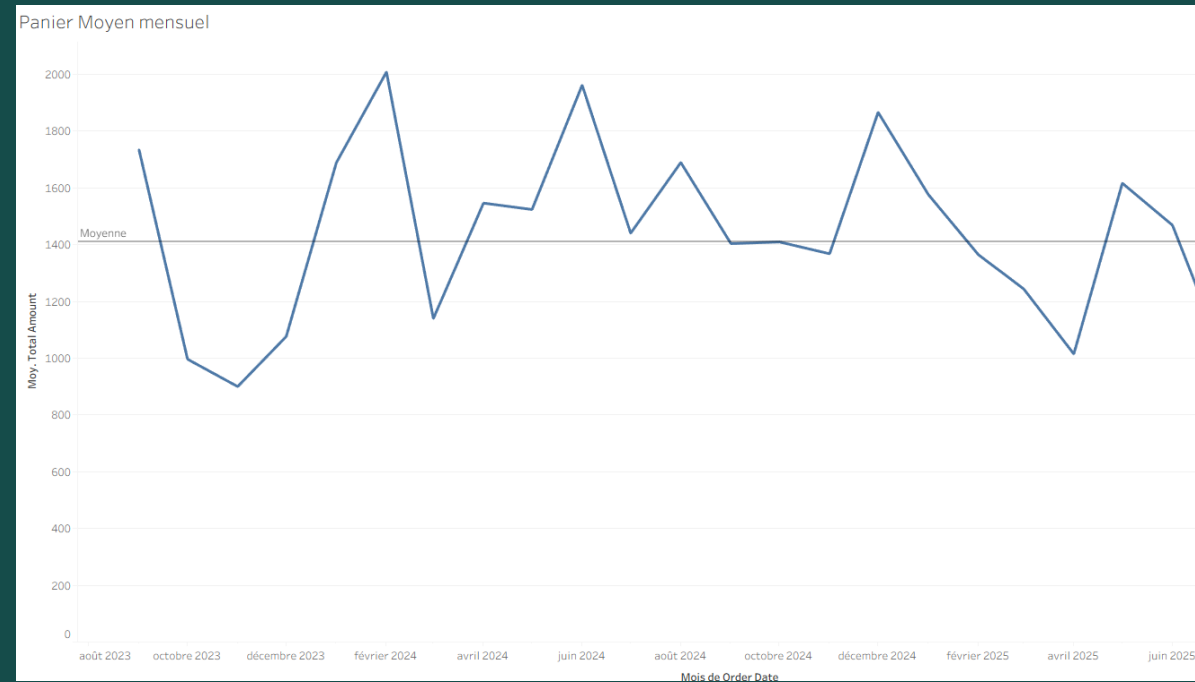
ecommerce_dw fact_sales
order_id : int(3)
product_id : int(2)
quantity : int(1)
unit_price : decimal(5,2)
date_id : int(11)
customer_id : int(2)
order_date : varchar(10)
total_amount : decimal(6,2)
product_name : varchar(10)
product_category : varchar(12)
product_price : decimal(5,2)
first_name : varchar(11)
last_name : varchar(9)
email : varchar(36)
country : varchar(44)
signup_date : varchar(10)
date_day : timestamp

Récupération des KPIs en SQL

- KPIs:
 - Produits les plus vendus
 - Taux de conversion
 - Chiffre d'affaire par mois
 - Évolution du panier moyen
 - Évolution des ventes / catégories / mois

Évolution du panier moyen SQL vs Tableau

mois	panier_moyen
2023-09	1734.70
2023-10	996.67
2023-11	900.23
2023-12	1076.55
2024-01	1689.05
2024-02	2008.73
2024-03	1139.90
2024-04	1546.71
2024-05	1524.13
2024-06	1962.34
2024-07	1440.72



KPIs sur Tableau

- Importation des données via CSV
- Etablir les relations

Quelle est la différence entre les relations et les jointures ? [En savoir plus](#)

orders.csv Opérateur customers.csv

Id Customer = # Id Customer (Cu

⊕ Ajouter des champs supplémentaires

> Options de performance

Supprimer une relation

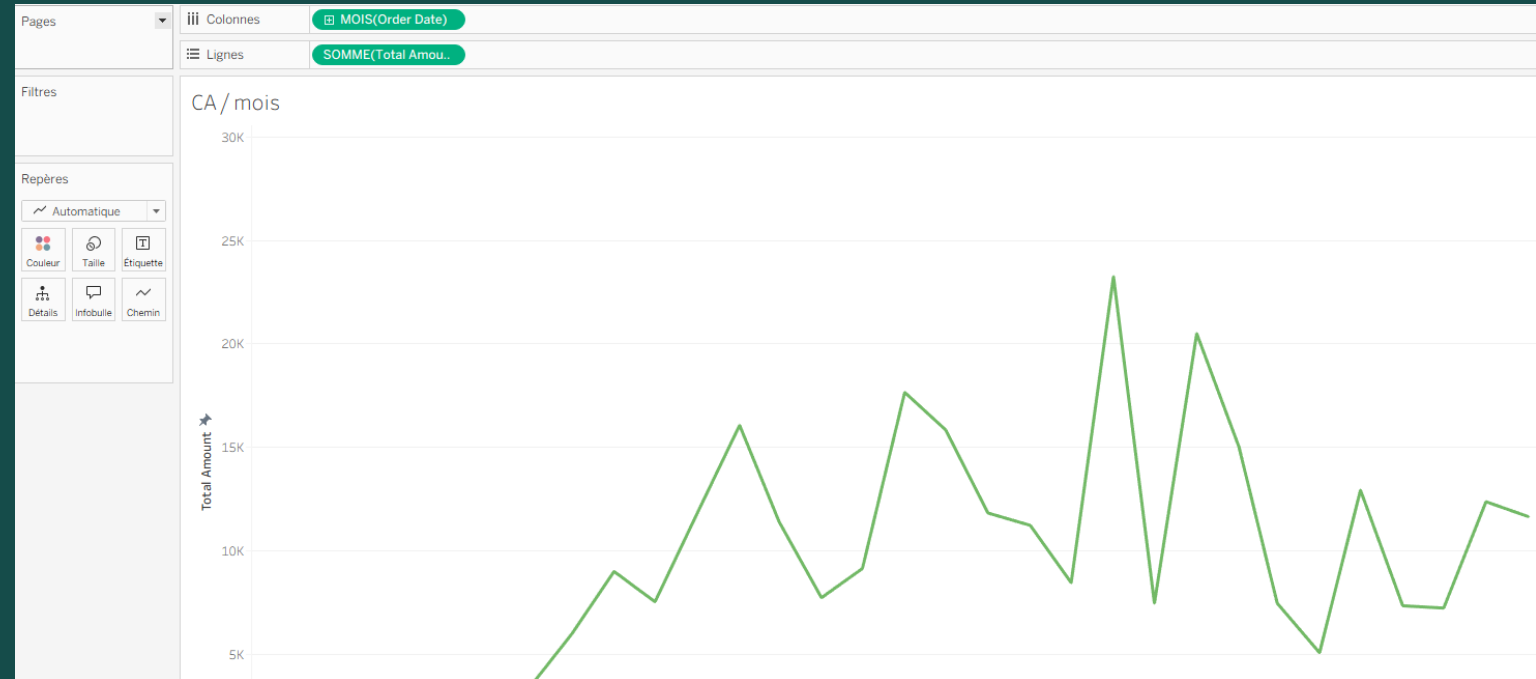
#	products.csv	products.csv	products.csv	products.csv
	Id-Product	Name	Category	Price
1		Produit_1	Maison	29,580
2		Produit_2	Maison	319,070
3		Produit_3	Électronique	420,220
4		Produit_4	Maison	40,300
5		Produit_5	Électronique	155,480
6		Produit_6	Électronique	214,160
7		Produit_7	Électronique	8,100
8		Produit_8	Vêtements	339,050



Création des visualisations des KPIs

CA / Mois

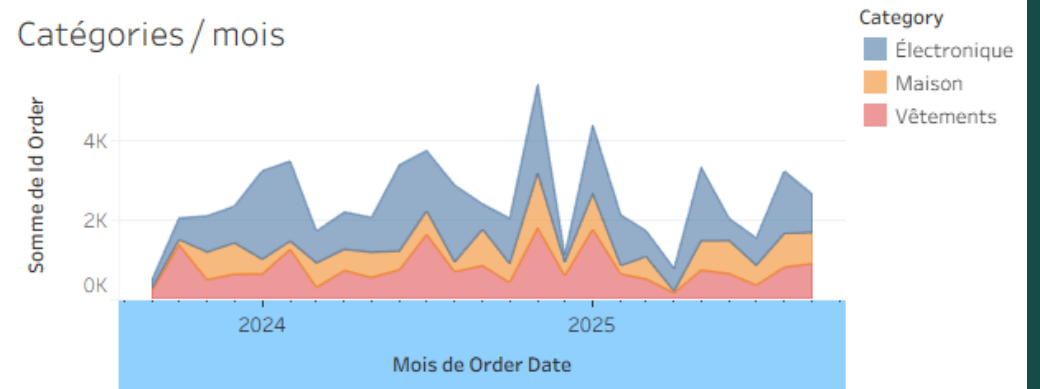
Choix de la visualisation
Définition mesures &
dimensions



Création du dashboard tableau

- Aggrège les différentes visualisation
- Aide au pilotage
- Décision basées sur les données

Catégories / mois



Classement des ventes

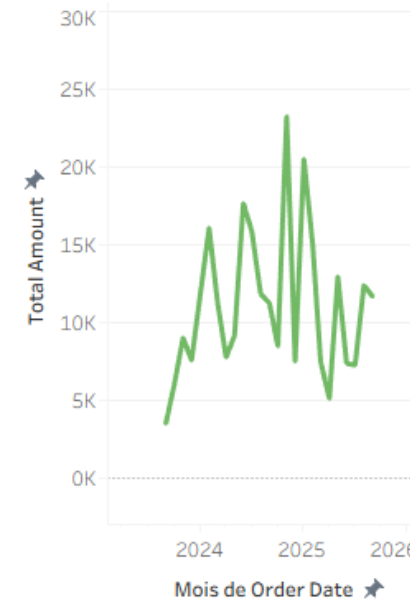
Name	
Produit_8	4689
Produit_10	4208
Produit_15	4116
Produit_19	3500
Produit_20	3337

%

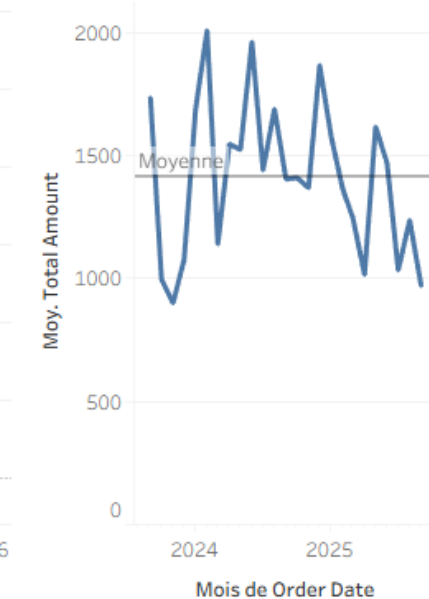
Conversi
on

98,00

CA / mois



Panier Moyen mensuel



Implémentation des rôles sur la base de données

```
MariaDB [(none)]>  
MariaDB [(none)]> CREATE USER 'analyst'@'%' IDENTIFIED BY 'password';  
Query OK, 0 rows affected (0.002 sec)  
  
MariaDB [(none)]> GRANT SELECT ON ecommerce_dw.view_customers TO 'analyst'@'%';  
Query OK, 0 rows affected (0.002 sec)  
  
MariaDB [(none)]> FLUSH PRIVILEGES;  
Query OK, 0 rows affected (0.001 sec)  
  
MariaDB [(none)]>
```

- Création d'un rôle
- Définition des droits
- Appliquer les changements

Création d'un masquage & vue 'sécurisée'

```
# mysql -u analyst -p
Enter password: *****
Welcome to the MariaDB monitor.  Commands end with ; or \g.
Your MariaDB connection id is 1461
Server version: 10.4.32-MariaDB mariadb.org binary distribution

Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.

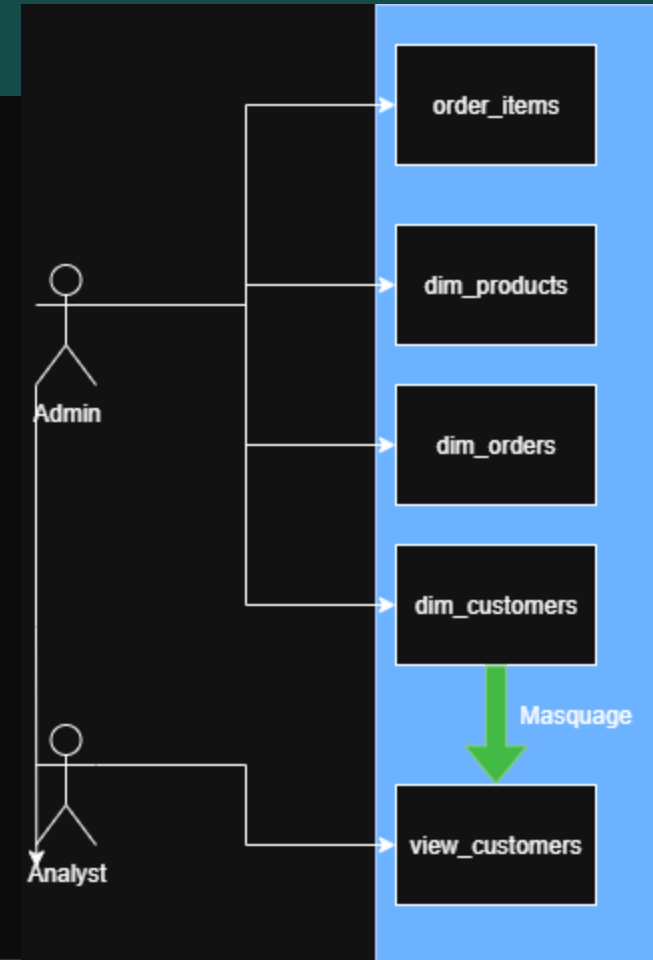
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

MariaDB [(none)]> select * from dim_customers;
ERROR 1046 (3D000): No database selected
MariaDB [(none)]> use ecommerce_dw
Database changed
MariaDB [ecommerce_dw]> select * from dim_customers;
ERROR 1142 (42000): SELECT command denied to user 'analyst'@'localhost' for table 'ecommerce_dw`.`dim_customers`
MariaDB [ecommerce_dw]> select * from view_customers limit 5;
```

id_customer	first_name	last_name	country	signup_date	masked_email
1	Margaret	Holmes	United States Virgin Islands	2024-10-30	m***@hopkins.com
2	Christopher	Obrien	Anguilla	2025-08-12	n***@mcfarland.com
3	Michael	Walker	Italy	2023-09-20	l***@thomas.com
4	Rebecca	Williams	Mongolia	2023-09-21	p***@yahoo.com
5	Tammy	Nguyen	Argentina	2024-03-17	a***@gmail.com

```
5 rows in set (0.001 sec)

MariaDB [ecommerce_dw]>
```





- Questions & Réponses



Fin & annexes



Vente / catégorie / mois :

```
---  
### 6. Commandes par mois & par catégories  
```SQL  
CREATE OR REPLACE VIEW vue_commandes_par_mois_et_categorie AS
SELECT
 LEFT(dim_orders.order_date, 7) AS mois,
 dim_products.category AS category,
 COUNT(*) AS total_commandes
FROM dim_orders
JOIN order_items ON dim_orders.id_order = order_items.id_order
JOIN dim_products ON dim_products.id_product = order_items.id_product
GROUP BY mois, category
ORDER BY mois;

```

mois	category	total_commandes
2023-09	Électronique	2
2023-09	Maison	1
2023-09	Vêtements	2
2023-10	Électronique	5
2023-10	Maison	1
2023-10	Vêtements	8
2023-11	Électronique	10
2023-11	Maison	10
2023-11	Vêtements	6
2023-12	Électronique	8
2023-12	Maison	6
2023-12	Vêtements	5
2024-01	Électronique	18
2024-01	Maison	4

# Produits les plus vendus (5):

**\*\*Requête SQL :\*\***

```
```sql
DROP TABLE IF EXISTS top_ventes;
CREATE TABLE top_ventes AS
SELECT dp.id_product, dp.name, COUNT(*) AS total_ventes
FROM dim_products dp
JOIN order_items oi ON dp.id_product = oi.id_product
JOIN dim_orders o ON o.id_order = oi.id_order
GROUP BY dp.id_product, dp.name
ORDER BY total_ventes DESC
LIMIT 5;
```
```

Options supplémentaires

| id_product | name       | total_ventes |
|------------|------------|--------------|
| 15         | Produit_15 | 40           |
| 8          | Produit_8  | 39           |
| 10         | Produit_10 | 38           |
| 3          | Produit_3  | 38           |
| 19         | Produit_19 | 36           |

# Taux de conversion:

(% de clients ayant effectué un achat après l'ouverture d'un compte)

**\*\*Requête SQL :\*\***

```sql

DROP TABLE IF EXISTS taux_fideliation;

CREATE TABLE taux_fideliation AS

SELECT ROUND(100.0 * COUNT(DISTINCT o.id_customer) / (SELECT COUNT(*) FROM dim_customers), 2) AS pct
FROM dim_orders o;

```

☐ Tout afficher | M

Options supplémentaires

pct

98.00

# Évolution du chiffre d'affaire par mois:

**\*\*Requête SQL :\*\***

```
```sql
DROP TABLE IF EXISTS ca_mensuel;
CREATE TABLE ca_mensuel AS
SELECT LEFT(order_date, 7) AS mois, SUM(total_amount) AS chiffre_affaires
FROM dim_orders
GROUP BY mois
ORDER BY mois;
```
```

| mois    | chiffre_affaires |
|---------|------------------|
| 2023-09 | 3469.39          |
| 2023-10 | 5980.02          |
| 2023-11 | 9002.27          |
| 2023-12 | 7535.87          |
| 2024-01 | 11823.38         |
| 2024-02 | 16069.80         |
| 2024-03 | 11398.99         |
| 2024-04 | 7733.57          |
| 2024-05 | 9144.79          |
| 2024-06 | 17661.05         |