

Visualisation des données en streaming à l'aide

de Tableau

Nom : Ilhem BAZIZ - Ould Taled Dyhia

Cours : Architecture Streaming Data

Date de rendu : 20 avril 2025

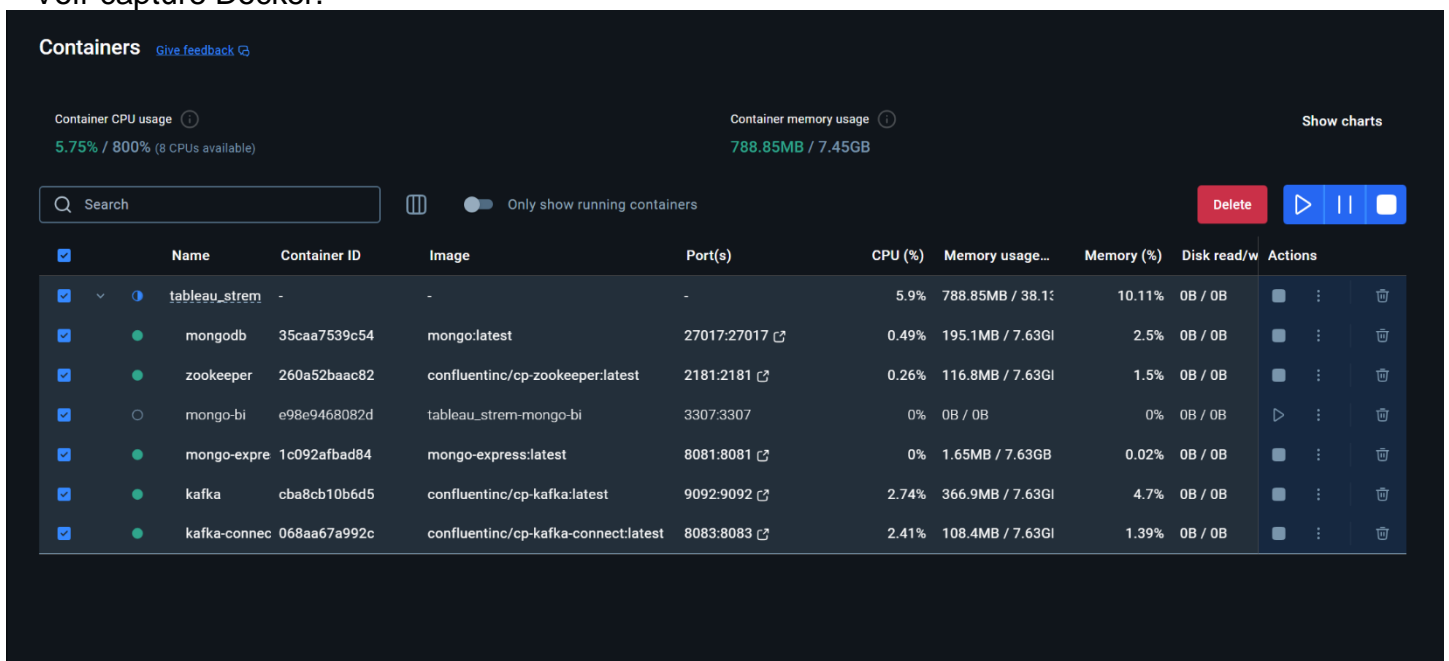
➤ Objectif du projet

- L'objectif de ce projet est de mettre en oeuvre une architecture de streaming de données en temps réel en utilisant Kafka comme producteur de données, MongoDB comme base de stockage NoSQL côté consommateur, et Tableau comme outil de visualisation. Le but final est de permettre l'analyse dynamique et interactive des ventes d'une entreprise fictive.

➤ Etapes de réalisation

1. Mise en place de l'environnement

- Création d'un compte Tableau Desktop.
- Déploiement via Docker : kafka, zookeeper, mongodb, mongo-express, mongo-bi, tableau_strem.
- Voir capture Docker.



The screenshot shows the Docker Desktop interface with a list of running containers. The containers are: tableau_strem, mongodb, zookeeper, mongo-bi, mongo-expre, kafka, and kafka-connec. The interface includes a search bar, a toggle for 'Only show running containers', and a 'Delete' button. The table below summarizes the data visible in the screenshot.

Name	Container ID	Image	Port(s)	CPU (%)	Memory usage...	Memory (%)	Disk read/w	Actions
tableau_strem	-	-	-	5.9%	788.85MB / 38.1%	10.11%	0B / 0B	[Stop] [Refresh] [Delete]
mongodb	35caa7539c54	mongo:latest	27017:27017	0.49%	195.1MB / 7.63Gi	2.5%	0B / 0B	[Stop] [Refresh] [Delete]
zookeeper	260a52baac82	confluentinc/cp-zookeeper:latest	2181:2181	0.26%	116.8MB / 7.63Gi	1.5%	0B / 0B	[Stop] [Refresh] [Delete]
mongo-bi	e98e9468082d	tableau_strem-mongo-bi	3307:3307	0%	0B / 0B	0%	0B / 0B	[Start] [Stop] [Refresh] [Delete]
mongo-expre	1c092afbad84	mongo-express:latest	8081:8081	0%	1.65MB / 7.63GB	0.02%	0B / 0B	[Stop] [Refresh] [Delete]
kafka	cba8cb10b6d5	confluentinc/cp-kafka:latest	9092:9092	2.74%	366.9MB / 7.63Gi	4.7%	0B / 0B	[Stop] [Refresh] [Delete]
kafka-connec	068aa67a992c	confluentinc/cp-kafka-connect:latest	8083:8083	2.41%	108.4MB / 7.63Gi	1.39%	0B / 0B	[Stop] [Refresh] [Delete]

2. Generation de donnees (producteur)

- Le script génère des données de ventes fictives et les envoie à un **topic Kafka** (ventes).
- Il choisit aléatoirement des produits, des catégories, des quantités, ainsi que des pays.
- Les informations incluent la **date**, le **montant total** de la vente, et la **quantité** achetée.
- Les données sont envoyées toutes les 2 secondes en utilisant un **producteur Kafka** pour simuler un flux de ventes en temps réel.

- Envoi de messages Kafka contenant : Date/heure, Montant, Produit, Region.

```
PS C:\Users\rymab\Desktop\Tableau_Strem> cd producer
PS C:\Users\rymab\Desktop\Tableau_Strem\producer> python producer.py
Envoi : {'date_heure': '2024-12-22 07:29:09', 'montant': 75.0, 'quantité': 5, 'produit': 'Accessoires téléphone', 'categorie': 'Téléphonie', 'region': 'Amérique', 'pays': 'États-Unis'}
Envoi : {'date_heure': '2024-12-18 23:14:36', 'montant': 160.0, 'quantité': 8, 'produit': 'Souris', 'categorie': 'Informatique', 'region': 'Asie', 'pays': 'Inde'}
Envoi : {'date_heure': '2023-03-01 18:29:26', 'montant': 2500.0, 'quantité': 5, 'produit': 'Téléphone', 'categorie': 'Téléphonie', 'region': 'Afrique', 'pays': 'Algérie'}
```

3. Consommation et stockage (MongoDB)

- Le script consomme les messages du topic Kafka (ventes) en utilisant un consommateur Kafka.
- Il se connecte à un serveur Kafka local pour récupérer les ventes envoyées en temps réel. Lorsqu'un message est reçu, il est désérialisé en format JSON.
- Ensuite, les données sont insérées dans une collection MongoDB (ventes_stream) dans la base de données streaming.
- Ce script permet de stocker les ventes en temps réel pour une analyse ultérieure.

```
PS C:\Users\rymab\Desktop\Tableau_Strem> cd consumer
PS C:\Users\rymab\Desktop\Tableau_Strem\consumer> python consumer.py
● Consommateur connecté à MongoDB. En attente des messages...
Vente reçue : {'date_heure': '2023-10-05 07:56:51', 'montant': 1000.0, 'quantité': 9, 'produit': 'Table', 'categorie': 'Maison', 'region': 'Asie', 'pays': 'Japon'}
Vente reçue : {'date_heure': '2023-06-29 22:55:09', 'montant': 500.0, 'quantité': 1, 'produit': 'Téléphone', 'categorie': 'Téléphonie', 'region': 'Asie', 'pays': 'Inde'}
Vente reçue : {'date_heure': '2023-03-03 10:24:34', 'montant': 60.0, 'quantité': 3, 'produit': 'Souris', 'categorie': 'Informatique', 'region': 'Amérique', 'pays': 'Mexique'}
```

4.





















- Kafka-Connect lit les topics Kafka et les insère dans MongoDB.

5. Connexion de Tableau a MongoDB

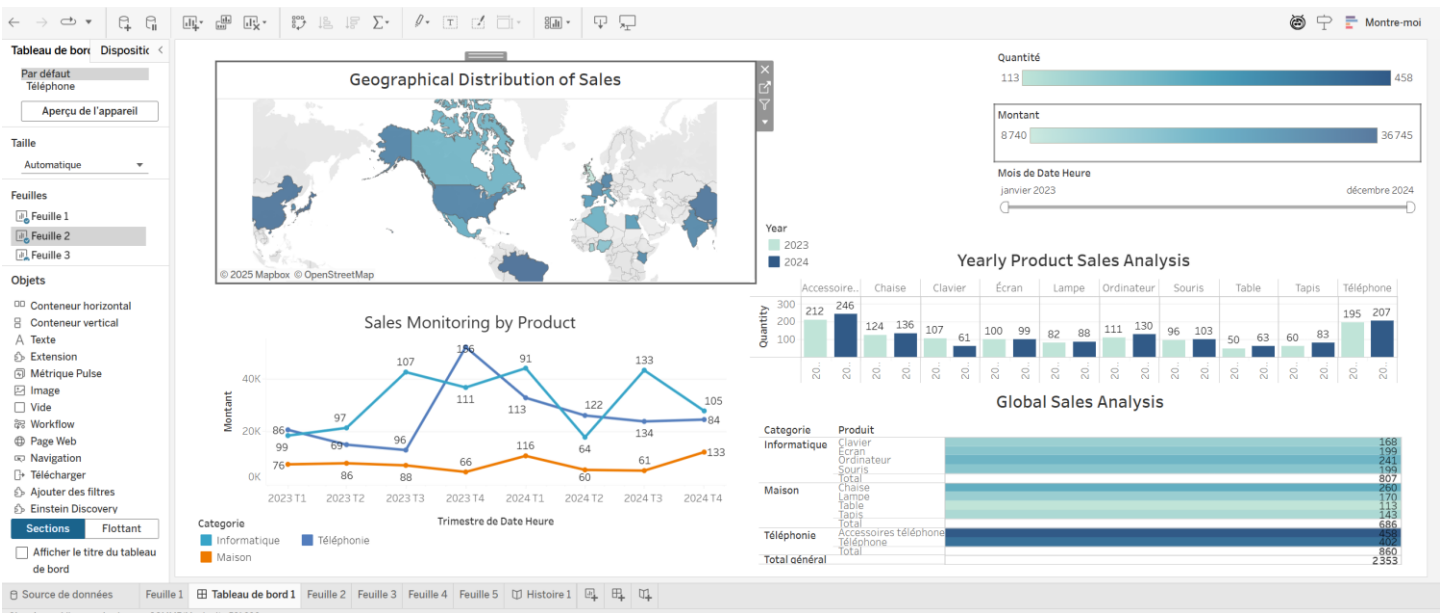
- Connexion via MongoDB BI Connector : Pour permettre la visualisation des données issues de MongoDB dans Tableau, une connexion a été établie à l'aide du **MongoDB BI Connector**. Ce connecteur permet d'exposer les données NoSQL de MongoDB dans un format relationnel compatible avec Tableau (type MySQL).
- Dans ce projet, le **BI Connector a été déployé via Docker**, et Tableau a été configuré pour se connecter au port du connecteur (localhost:3307) en utilisant un connecteur MySQL. Une fois la connexion établie, les collections MongoDB sont accessibles comme des tables classiques, ce qui permet de créer des visualisations dynamiques dans Tableau.

🗑️ Delete all 422 documents retrieved

<< < 1 2 3 4 5 > >>

_id	date_heure	montant	quantité	produit	categorie	region	pays
 	68054eeffb7f6a52385f342d	2024-09-19 22:29:34	135	9	Accessoires téléphone	Téléphonie	Afrique Algérie
 	68054ef1fb7f6a52385f342e	2024-07-11 17:26:31	120	6	Souris	Informatique	Asie Inde
 	68054ef3fb7f6a52385f342f	2024-09-05 14:51:55	300	4	Chaise	Maison	Asie Corée du Sud
 	68054ef5fb7f6a52385f3430	2023-10-08 13:09:58	45	3	Accessoires téléphone	Téléphonie	Amérique États-Unis
 	68054ef7fb7f6a52385f3431	2024-06-09 04:29:54	60	3	Souris	Informatique	Asie Chine
 	68054ef9fb7f6a52385f3432	2024-10-23 08:48:55	200	10	Souris	Informatique	Europe Espagne
 	68054efbfb7f6a52385f3433	2024-12-19 20:08:35	405	9	Lampe	Maison	Europe Espagne
 	68054efd7fb7f6a52385f3434	2023-09-10 02:10:05	3000	6	Téléphone	Téléphonie	Amérique Brésil
 	68054efffb7f6a52385f3435	2024-05-04 20:04:54	105	7	Accessoires téléphone	Téléphonie	Amérique Argentine
 	68054f01fb7f6a52385f3436	2023-05-10 06:01:21	90	6	Accessoires téléphone	Téléphonie	Asie Chine

6. Visualisations produites



Storytelling – Dashboard de Vente

- Objectif : analyser les ventes de 2023 à 2024 selon les produits, les catégories, les périodes et les régions géographiques.
- La carte montre une forte concentration des ventes en Amérique du Nord, Europe, et Asie (États-Unis, France, Corée du Sud, Afrique du Sud).
- Les ventes de produits comme **Accessoires téléphone**, **Téléphone** et **Ordinateur** ont augmenté en 2024 par rapport à 2023.
- Le **Clavier** et quelques autres produits ont connu une baisse de ventes entre 2023 et 2024.
- La catégorie **Téléphonie** est globalement la plus performante en volume de ventes.
- Le suivi trimestriel montre un **pic de ventes au 4^e trimestre 2023**, notamment dans la Téléphonie (effet fêtes ou promotions).
- En analyse globale, **Téléphonie (860 unités)** dépasse **Informatique (807 unités)** et **Maison (686 unités)**.
- Recommandations :
 - Renforcer les campagnes marketing autour des produits téléphonie.
 - Cibler les zones géographiques les plus actives (Amérique, Asie).
 - Exploiter les périodes fortes comme la fin d'année pour booster les ventes.

Fonctionnalités Interactives du Tableau de Bord

- Filtres dynamiques :
 - Filtrage des ventes par année (2023 / 2024)
 - Sélection par mois de la date grâce à un curseur temporel
 - Possibilité de filtrer par produit, catégorie, ou région
- Carte interactive :
 - Affichage des ventes par pays avec survol (tooltip) pour voir les chiffres
 - Zoom et déplacement sur la carte pour explorer les zones
 - Mise à jour automatique des visualisations :
 - Tous les graphiques se mettent à jour instantanément en fonction des filtres sélectionnés
 - Tooltip personnalisé .
 - Survol des barres et lignes pour voir les valeurs exactes (quantité, montant, etc.)
 - Comparaison entre année.
- Visualisation des différences entre 2023 et 2024 par produit grâce aux barres colorées par année
- Navigation fluide entre feuilles :
- Organisation des analyses sur plusieurs feuilles reliées à un tableau de bord unique