

Exercice de Collecte, Stockage et Analyse de Données

Vous êtes chargé(e) de développer un système automatique pour collecter des données de commandes générées par un script Python appelé “orders_generator.py”. Ce script génère des fichiers au format JSON nommés “orders.json”, chaque fichier contenant 10 commandes. Votre objectif est de mettre en place un système qui récupère ces commandes, les stocke dans une base de données de votre choix, puis analyse, anonymise et stocker les données dans une nouvelle table propre pour des études ultérieures.

Tâches à réaliser :

1. Collecte des données :

- Mettez en place un système automatisé pour récupérer les fichiers “orders.json” générés par “orders_generator.py”.
- Assurez-vous que les nouveaux fichiers sont détectés dès qu'ils sont générés et que leurs données sont prêtes à être traitées.

2. Stockage des données :

- Choisissez une base de données adaptée pour stocker les commandes. Vous pouvez utiliser des solutions telles que MySQL, PostgreSQL, MongoDB, etc.
- Créez une structure de base de données pour stocker les informations des commandes de manière organisée.

3. Analyse des données :

- Implémentez des scripts pour analyser les données stockées. Cela peut inclure des analyses statistiques, des visualisations, etc.
- Identifiez les tendances, les habitudes d'achat, les produits les plus populaires, etc.

4. Traitement et anonymisation des données :

- Assurez-vous de respecter la confidentialité des utilisateurs en anonymisant les données sensibles.
- Remplacez les informations personnelles telles que les noms, adresses, etc., par des identifiants uniques ou des données fictives.
- Enlever aussi toute information que vous jugerez inutile

5. Documentation :

- Fournissez une documentation expliquant le fonctionnement de votre système, y compris les instructions d'installation, de configuration et d'utilisation.
- Expliquez les choix de conception, y compris la sélection de la base de données et les méthodes d'anonymisation des données.