

## Etude de cas

### Gestion des Données

#### Exercice 1: Mise en Œuvre de Mesures de Sécurité des Données

- **Objectif :** Sécuriser les données sensibles en utilisant le chiffrement et les contrôles d'accès.
- **Source des Données :** Utilisez un jeu de données simulé d'informations clients disponible sur Mockaroo, qui peut générer des ensembles de données avec des noms, des emails, des numéros de téléphone, et d'autres informations personnelles.
- **Tâches :**
  1. Chiffrer le jeu de données en se basant sur le tuto\_chiffrement.ipynb ou en utilisant des bibliothèques Python comme `cryptography` ou `PyCrypto`.
  2. Mettre en place un contrôle d'accès basé sur les rôles (RBAC) expliqué en pièce jointe **RBAC.pdf** pour restreindre l'accès aux champs sensibles dans le jeu de données, tels que les emails et les numéros de téléphone/prénom/ carte de crédit selon la base des données que t'as générer.
  3. Développer un document de politique qui décrit comment le chiffrement et les contrôles d'accès sont mis en œuvre pour protéger la confidentialité des données.
- **Résultat Attendu :**
  - Une version chiffrée du jeu de données avec des contrôles d'accès restreints.

- Un document de politique expliquant les méthodes de chiffrage et les mécanismes de contrôle d'accès utilisés.

## Exercice 2: Analyse des Tendances de Produits dans les Ventes

### Objectif

Analyser les données de ventes pour identifier les tendances de popularité des produits et les périodes de forte demande. Cet exercice vise à dégager des insights sur les produits les plus vendus et les variations saisonnières des ventes.

### Données

Téléchargez le fichier `data_vente.csv`, qui contient les informations suivantes :

- **Order ID** : numéro de la commande
- **Product** : nom du produit
- **Quantity Ordered** : quantité vendue
- **Price Each** : prix unitaire
- **Order Date** : date de la commande
- **Purchase Address** : adresse de vente

**Python** : Téléchargez le fichier `data_vente.csv` et importez les données en utilisant la bibliothèque `pandas`.

**MySQL** : Importez les données dans MySQL et créez la table `SalesData`. ou charger les données du fichier CSV dans la table,

### Questions

#### 1. Quels sont les produits les plus populaires ?

- **Python** : Utilisez la bibliothèque `pandas` pour compter les occurrences de chaque produit.

- **MySQL** : Utilisez une requête SQL pour grouper par produit et trier par quantité vendue.
2. Quel est le meilleur mois de vente pour chaque produit ?
- **Python** : Extrayez le mois de la colonne *Order Date* et calculez les ventes totales pour chaque produit par mois.
  - **MySQL** : Ajoutez une colonne *Mois* et regroupez par mois et produit.
3. Quels produits ont une demande saisonnière ?
- **Python** : Utilisez un graphique pour visualiser les ventes par mois et identifier les produits avec une demande saisonnière.
  - **MySQL** : Agrégez les ventes par mois et produit pour détecter les variations significatives.
4. Quel est le revenu total généré par chaque produit ?
- **Python** : Ajoutez une colonne calculée *Total Sales* (prix unitaire  $\times$  quantité) et calculez le revenu total par produit.
  - **MySQL** : Ajoutez une colonne calculée pour *Total Sales* et regroupez par produit.

Ces questions permettent de pratiquer des concepts fondamentaux en analyse de données tout en découvrant des informations utiles sur les tendances de ventes dans l'entreprise.

## Exercice 3 : Politique d'archivage et de destruction des données

- **Contexte**

Une entreprise de services financiers possède un système de gestion des données clients qui contient des informations sensibles, notamment des données personnelles, des informations bancaires, et des historiques de transactions. L'entreprise doit se conformer aux réglementations de protection des données (par exemple, RGPD) et gérer efficacement l'archivage et la destruction des données pour minimiser les risques de sécurité et de non-conformité.

- **Objectifs**

- Élaborer une politique d'archivage et de destruction des données.
- Identifier les bonnes pratiques en matière de gestion du cycle de vie des données.
- Proposer un plan de mise en œuvre pour assurer la conformité réglementaire et la sécurité des données.

- **Tâches**

- Étape 1 : Identification des données

- Listez les types de données que l'entreprise devrait conserver à long terme et celles qui peuvent être archivées ou détruites. Expliquez les critères qui guideraient ces décisions (ex. : valeur réglementaire, utilité commerciale, etc.).

- Étape 2 : Archivage

- Proposez un processus d'archivage des données. Incluez les méthodes de stockage (ex. : bases de données archivées, stockage sécurisé), les niveaux de sécurité requis, et la fréquence d'archivage.
- Comment garantir que seules les personnes autorisées ont accès aux archives?

- Étape 3 : Destruction des données

- Décrivez le processus de destruction sécurisé des données. Quelles méthodes recommanderiez-vous (ex. : suppression logicielle, déchiquetage physique pour les supports physiques)?
- Identifiez des outils ou technologies pouvant faciliter la destruction sécurisée et conforme des données sensibles.

- Étape 4 : Conformité et audit

- Comment l'entreprise peut-elle s'assurer que l'archivage et la destruction des données sont réalisés conformément aux réglementations (comme le RGPD) ? Proposez un plan d'audit et de surveillance pour vérifier la conformité.

- Étape 5 : Formation et sensibilisation
  - Définissez un plan de formation pour sensibiliser les employés aux politiques d'archivage et de destruction des données. Quels éléments de formation seraient nécessaires?
- Livrable attendu Rédigez un document décrivant la politique d'archivage et de destruction des données pour l'entreprise, comprenant les étapes, les rôles et responsabilités, et les contrôles de conformité.
- Questions de réflexion
  - Quels sont les risques potentiels pour l'entreprise si l'archivage et la destruction des données ne sont pas effectués correctement?
  - Comment le non-respect des politiques d'archivage et de destruction des données peut-il impacter la confiance des clients et la réputation de l'entreprise?

## Exercice : Analyse de Conformité des Régulations pour le Partage des Données

- **Objectif :** Assurer la conformité avec les réglementations internationales de protection des données lors du partage des données.
- **Source des Données :** Utilisez un jeu de données de santé disponible sur HealthData.gov, tel que le "Chronic Condition Data Warehouse", qui contient des informations de santé sensibles.
- **Tâches :**
  1. Examiner le jeu de données et identifier les éléments de données sensibles qui nécessitent une protection sous des réglementations comme le RGPD, HIPAA (pour les données de santé), ou CCPA.

2. Proposer des techniques d'anonymisation (par exemple, masquage des données, pseudonymisation) pour protéger les informations sensibles avant le partage du jeu de données.
3. Développer une liste de vérification de conformité qui inclut les étapes pour s'assurer que le partage des données est conforme aux lois de protection des données pertinentes.

- **Résultat Attendu :**

- Une version anonymisée du jeu de données prête pour un partage conforme.
- Une liste de vérification de conformité qui garantit que toutes les protections nécessaires sont en place pour le partage des données.