**TP : Conception et manipulation d’une base de données à partir d’un jeu de données**

**Objectif :**

L'objectif de ce TP est de vous familiariser avec la conception d'une base de données relationnelle à partir d'un jeu de données réel trouvé sur internet. Vous allez créer une base de données, la peupler avec des données réelles, et effectuer des requêtes SQL afin de manipuler et d'analyser ces données. Vous devrez appliquer les principes de conception de bases de données relationnelles et respecter les contraintes d'intégrité.

**Étapes du TP :**

1. **Choix du jeu de données :**
   * Vous devez choisir un jeu de données en ligne. Celui-ci peut provenir de sources comme [Kaggle](https://www.kaggle.com/), OpenDataFrance, ou toute autre source fiable.
   * Le jeu de données choisi doit être suffisamment détaillé pour permettre la création d'une base de données avec plusieurs tables et relations (au minimum trois tables liées par une clé étrangère).
2. **Analyse du jeu de données :**
   * Examinez les données disponibles pour identifier les entités (par exemple : "Personne", "Produit", "Commande", etc.) et les relations entre elles.
   * Pour chaque entité, identifiez les attributs pertinents (par exemple, pour une entité "Personne", les attributs pourraient être : Nom, Prénom, Email, etc.).
   * Déterminez les types de données (entiers, chaînes, dates, etc.) pour chaque attribut.
3. **Modélisation de la base de données :**

À l’aide de l’outil open source looping téléchargeable [ici](https://www.looping-mcd.fr/)

* + Créez un diagramme entité-association (DEA) pour représenter les entités, leurs attributs et leurs relations.
  + Définissez les clés primaires et étrangères.
  + Prévoyez les contraintes d'intégrité :
    - **Intégrité des entités** : Assurez-vous que chaque enregistrement a une clé primaire unique.
    - **Intégrité référentielle** : Les clés étrangères doivent être cohérentes et faire référence à des clés primaires existantes dans d’autres tables.
    - **Intégrité des données** : Appliquez des contraintes sur les valeurs (par exemple, non nullité, valeurs uniques, etc.).

1. **Création de la base de données :**
   * À l’aide de MariaDB (MySQL), créez la base de données et les tables correspondantes.
   * Implémentez les clés primaires, les clés étrangères et les contraintes d’intégrité.
2. **Importation des données :**
   * Téléchargez le jeu de données choisi et préparez-le au format approprié (CSV, SQL etc.).
   * Importez les données dans vos tables en utilisant des outils d'importation (phpMyAdmin) de votre SGBD ou des scripts .slq .
3. **Requêtes SQL (Interrogation de la base de données créée) :**
   * **Sélection de données** : Utilisez la commande SELECT pour interroger la base de données et afficher les données.
   * **Agrégation** : Utilisez GROUP BY pour effectuer des agrégations (par exemple, somme, moyenne) sur les données.
   * **Jointures** : Effectuez des jointures (INNER JOIN, LEFT JOIN, etc.) pour combiner plusieurs tables et obtenir des informations complètes.
   * **Filtres et conditions** : Utilisez WHERE, HAVING, ORDER BY, et d'autres conditions pour affiner les résultats.
   * Exemples de requêtes :
     + Afficher les informations des utilisateurs ayant passé plus de 5 commandes.
     + Trouver le produit le plus acheté par catégorie.
     + Calculer le revenu total généré par chaque vendeur.
4. **Analyse des résultats :**
   * Analysez les résultats des requêtes exécutées et tirez des conclusions pertinentes à partir des données.
   * Présentez les résultats sous forme de tableau ou de graphique (si nécessaire).
5. **Rapport final :**
   * Réalisez une présentation (power point) détaillant chaque étape de la conception de votre base de données, la modélisation, l’importation des données et les requêtes SQL réalisées.
   * Incluez des captures d'écran de vos requêtes et des résultats obtenus.

**Critères d’évaluation :**

* **Qualité de la modélisation de la base de données** : Adéquation entre les entités et les tables, bonnes pratiques de conception (normalisation, clés primaires et étrangères).
* **Respect des contraintes d’intégrité** : Les contraintes doivent être bien appliquées pour assurer la cohérence des données.
* **Justesse des requêtes SQL** : Les requêtes doivent être correctement formulées et donner les bons résultats.
* **Clarté du rapport** : Le rapport doit être bien structuré et détailler le processus de manière compréhensible.