

Partie A : étude des fichiers sources

1. A partir des fichiers CSV fournis, proposez un schéma du système d'information associé. Vous pouvez utiliser le format que vous voulez (schéma Entité/Association, schéma relationnel, UML, etc.)
2. Proposez un script SQL permettant de créer ce schéma sous PostgreSQL
3. Proposez un script SQL permettant de charger, dans ce schéma, les données contenues dans les fichiers CSV fournis
4. Proposez une requête SQL, en donnant son résultat, permettant d'obtenir le modèle le plus vendu

Partie B : étude des ventes avec SQL

1. Proposez un schéma d'entrepôt de données avec les caractéristiques suivantes :
  - Fait : nombre de produits vendus
  - Dimensions : temps ; localisation avec pays, ville et magasin ; type avec catégorie et modèle
2. Proposez un script SQL permettant de créer cet entrepôt de données sous PostgreSQL
3. Intégrez des données dans cet entrepôt

Proposez des requêtes SQL, en donnant leur résultat, pour répondre aux interrogations suivantes :

4. Quel est le nombre total de produits vendus ainsi que le nombre de produits vendus par catégorie, par magasin et par année ; par catégorie et par magasin ; par catégorie ?
5. Quel est, pour chaque année, le classement des catégories de produits suivant le nombre de produits vendus ?
6. Pour chaque catégorie, quels sont les deux clients qui ont acheté le plus de produits ?
7. Quel est le nombre de catégories dont le nombre total de produits vendus est inférieur ou égale à 5000, est supérieur à 5000 et inférieur ou égale à 10000, est supérieur à 10000 et inférieur ou égale à 15000, est supérieur à 15000 ?
8. Quelle est l'évolution, mois après mois, de la quantité totale de produits vendus ?

Partie C : étude des ventes avec un programme

Pour cette partie, vous travaillerez directement sur les fichiers CSV fournis en utilisant le langage de programmation de votre choix

1. Proposez un programme permettant d'obtenir le modèle le plus vendu
2. Proposez des programmes permettant de répondre aux interrogations de la partie B (questions 4 à 8) en affichant leur résultat

Partie D : étude des bénéfices avec visualisation

1. En procédant comme dans la Partie B ou la Partie C, proposez des interrogations, en donnant leur résultat, permettant d'étudier les bénéfices par modèle plutôt que par catégorie
2. En vous appuyant sur les interrogations précédentes, proposez un tableau de bord (ensemble de graphiques) permettant de visualiser des indicateurs clés de performance (ICP) établis sur le bénéfice

Partie E : étude de la satisfaction client

1. Proposez un schéma d'entrepôt de données, permettant d'étudier la satisfaction des clients, avec les caractéristiques suivantes :
  - Fait 1 : la note de satisfaction client indiquée à chaque vente. Cette note va de 0 (pas satisfait) à 5 (très satisfait)
  - Fait 2 : le délai moyen de livraison en nombre de jours des produits d'une vente
  - Dimensions : temps ; client ; type avec catégorie et modèle
2. Intégrez des données dans cet entrepôt
3. Proposez des requêtes SQL sur cet entrepôt permettant d'étudier la satisfaction suivant les 2 faits énoncés

Partie F : discussion

Après cette étude, vous avez une certaine compréhension sur les fichiers sources CSV et le système d'information associé

1. Donnez des exemples d'autres études que l'on pourrait mener à partir de ce système d'information
2. Faites une analyse critique de ce système d'information en indiquant quels sont, selon vous, les points à améliorer en termes de modélisation. Illustrez vos propos avec des exemples