

BD51

Business Intelligence & Data Warehouse

Examen Final du 17/1/2011

Rédacteur : Christian FISCHER

Département Informatique

Automne 2011

Conditions de réalisation :

Documents autorisés : Supports de Cours et TP de BD51

Durée : 2h

Barème : Le sujet comprend 2 dossiers indépendants.

| | | |
|---------------------|--|---|
| Dossier 1 7 pts | 1. Requêtes SQL OLAP Rédiger les ordres SQL | 7 pts |
| Dossier 2 13 pts | Modélisation du Data Warehouse 1. Modèle dimensionnel 2. Partitionnement de table 3. Optimisation 4. Internationalisation 5. Architecture du système d'information décisionnel | 5 pts 2 pts 2 pts 1 pt 3pts |

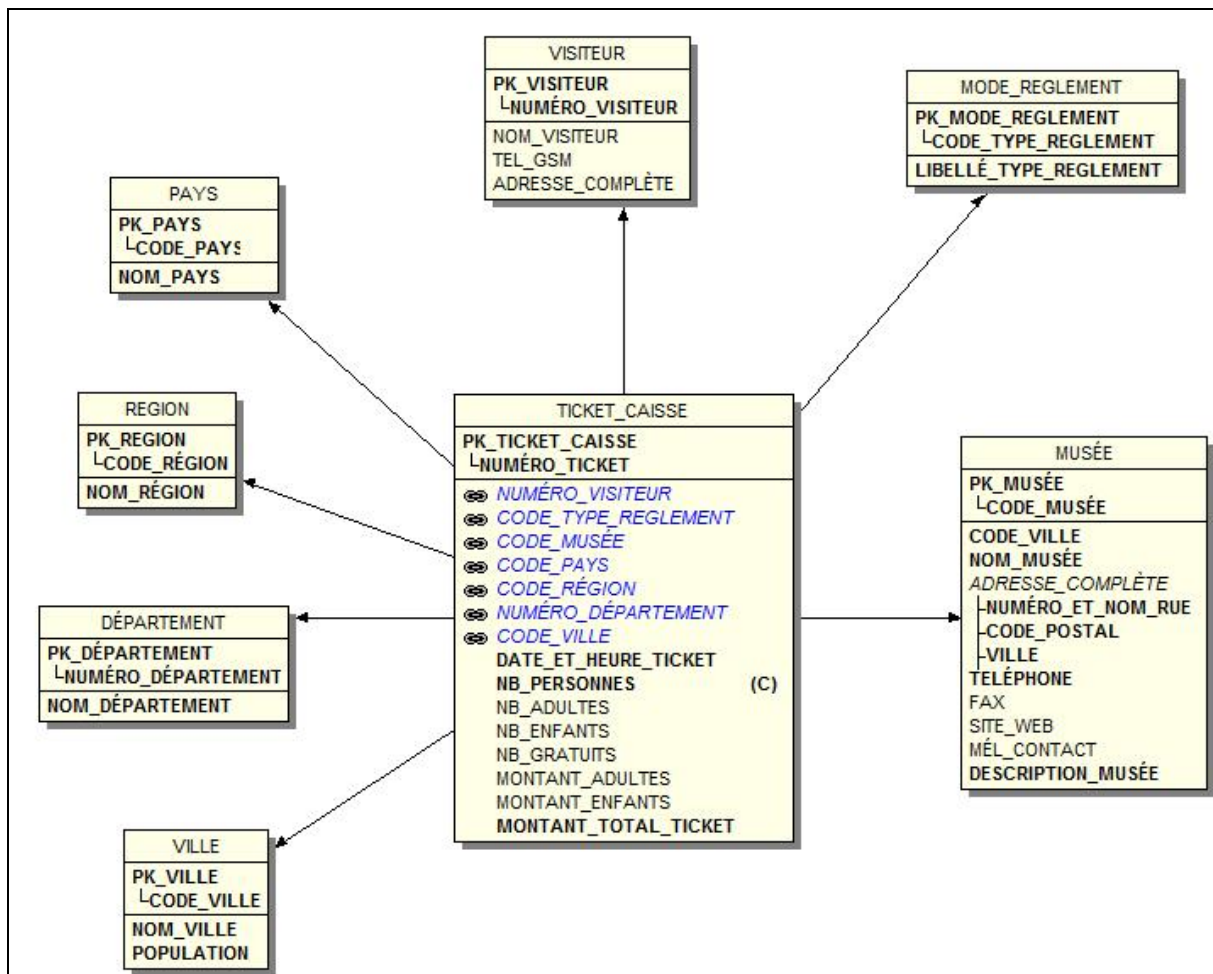
Dossier 1 : Interrogation du système d'information de la fréquentation des musées

Dossier 1 Première partie : Ecriture de requêtes SQL - OLAP

Le ministère de la culture a décidé de la mise en place du système d'information décisionnel et de pilotage de la fréquentation des musées.

L'équipe chargée de l'entrepôt de données a récemment examiné la conception du système d'analyse actuel.

Le modèle dimensionnel en étoile de la base de données MUSEE_DW pour le suivi de la fréquentation est le suivant :



Travail à faire : Rédiger les requêtes SQL ci-dessous.

Utilisation des jointures ANSI, des alias de tables (première lettre de la table) et des alias de colonnes (nom entre parenthèse)

Les fonctions YEAR() et MONTH() sont utilisées pour extraire l'année et le mois à partir de la date et heure du ticket.

Requête 1 : (1 pt)

Calculer le nombre total de visites ainsi que le montant total des entrées par nom du musée et par année. (musée, année, fréquentation, ventes)

Seules les années 2010 et 2011 seront affichées.

Tous les musées doivent être affichés.

Le résultat sera trié par nom du musée et année

Requête 2 : (1 pt)

Classement des musées situés en France par fréquentation pour l'année 2011.

Affichage du nom du musée (musée), de la fréquentation totale (fréquentation), du rang général (rang national), du rang par région (rang régional).

Le résultat sera trié par rang national croissant.

Requête 3 : (1 pt)

Afficher les musées avec la fréquentation mensuelle l'année 2011 (Nom du musée, Mois, Fréquentation mensuelle).

Une ligne calculant la fréquentation totale annuelle par musée sera ajoutée.

Une ligne calculant la fréquentation totale annuelle pour tous les musées sera ajoutée.

Le total général sera la dernière ligne du résultat.

Le résultat sera trié par musée et mois, suivi du total annuel.

Requête 4 : (1 pt)

Calculer le nombre total de visites (fréquentation) ainsi que le montant total des entrées (ventes) pour chaque pays et année.

Seules les trois dernières années, plus l'année en cours (déterminée par la date d'exécution de la requête) doivent être retournées.

Une ligne de totalisation pour chaque pays et une ligne de totalisation pour chaque année doivent être ajoutées dans le résultat.

L'utilisation d'une clause Grouping Sets + la décodification des valeurs nulles est préconisée.

Le niveau de groupe est calculé à l'aide de la fonction grouping_id().

Le résultat sera trié par pays et année.

Les colonnes seront organisées ainsi : pays, année, fréquentation, ventes, niveau de groupe

Requête 5 : (1 pt)

Afficher les visiteurs (nom, adresse complète, nombre de musée visités) ayant fréquentés le plus de musées différents en 2011 et situés en Franche-Comté.

Seuls les 10 visiteurs ayant la fréquentation la plus élevée seront affichés.

Le résultat sera trié de manière décroissante sur le nombre de musée visités.

Requête 6 : (2 pts)

Afficher le nom du musée (musée), le montant total des tickets de chaque musée (ventes), le montant moyen du ticket pour tous les musées (ticket moyen tous musées), le montant moyen du ticket de chaque musée (ticket moyen du musée), l'écart entre le montant moyen général et le montant moyen de chaque musée (écart), le rang global (sans trou en cas d'égalité) calculé sur le montant moyen du ticket par musée (rang par ticket moyen).

Seuls les tickets de l'année 2011 seront pris en compte.

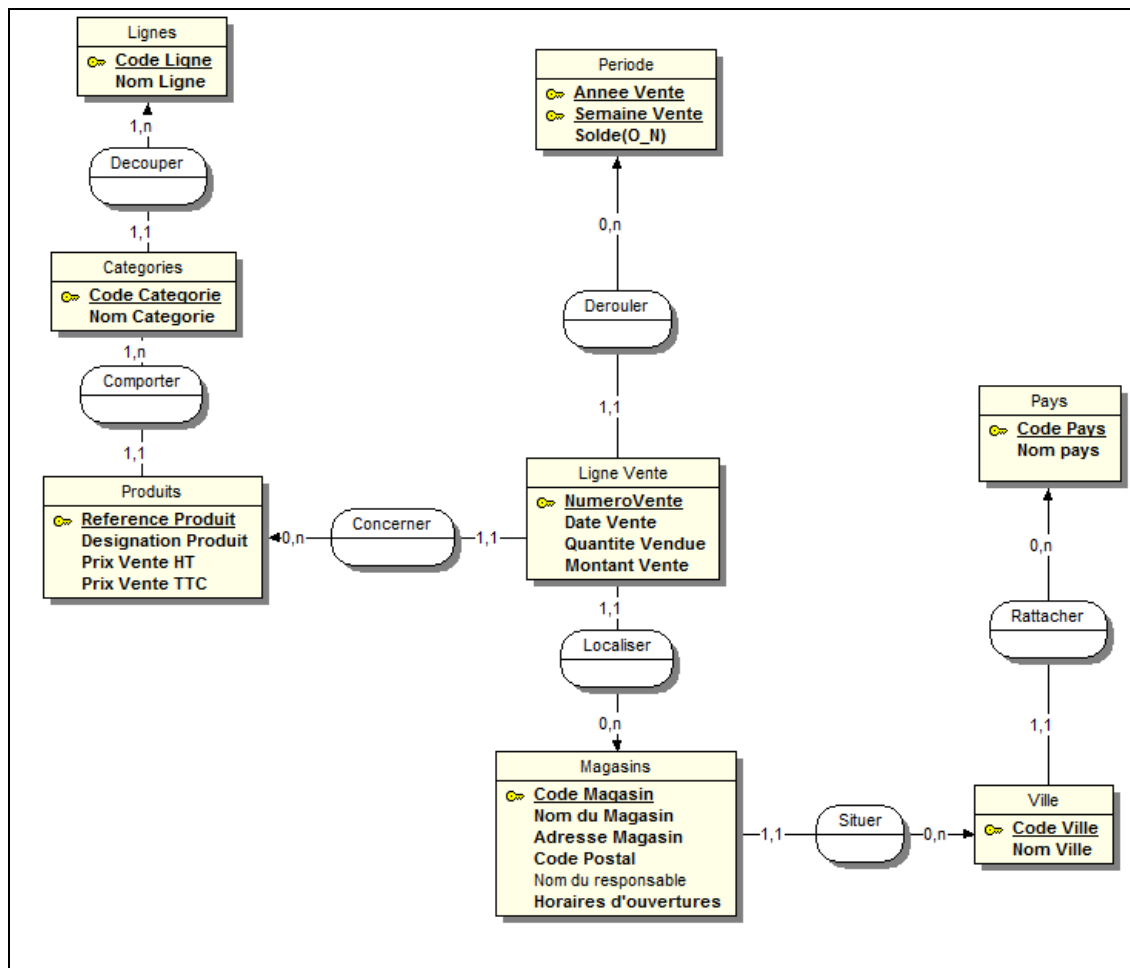
Le résultat sera trié par nom du musée.

Dossier 2 : Modèle dimensionnel de la gestion d'un chaîne de magasins

Contexte général

Une société de distribution commercialise ses produits à travers différents point de ventes (magasins) répartis dans les grandes villes européennes de la zone euro.

Modèle entité-association validé



Travail à faire

Q1. A partir du modèle entité-association présenté, vous devez fournir un modèle dimensionnel en étoile.

Dans le modèle fourni vous indiquerez quelles sont les tables de dimensions et les tables de faits.

Q2. La table «Ligne de Vente » contient les données depuis le 1er janvier 2008. Cette table doit être partitionnée par année civile jusqu'à fin 2013.

Rédiger l'ordre SQL de création de la table partitionnée en utilisant la syntaxe d'Oracle 10G.

Q3. Pour la restitution des données vous avez conçu un univers BusinessObjects à l'aide d'un modèle en étoile.

La plupart des rapports s'effectue sur l'axe chronologique Année et Mois.

1. Proposez une solution pour améliorer les performances de l'actualisation de ces rapports, sans modifier la structure existante (seuls des ajouts sont possibles).

2. Rédiger les ordres SQL associés à votre solution.

Q4. La société souhaite s'implanter dans des pays hors zone Euro (USA, Canada, Inde, Chine, Brésil, ...).

Quelle modification devez-vous mettre en œuvre sur la structure de votre entrepôt de données ?

Q5. Les directeurs commerciaux de chaque pays doivent établir des statistiques sur la rentabilité de chaque magasin. Un directeur n'ayant accès qu'aux données du pays dont il est responsable.

1. Proposez une architecture technique et logicielle permettant, à ces personnes n'ayant qu'une connaissance très limitée du langage SQL, de réaliser les tableaux de bord nécessaires à leur métier.

2. Comment pouvez-vous garantir la confidentialité de l'accès aux données d'un pays ?

3. Quelle modification est nécessaire sur la structure de votre entrepôt de données ?