

DataWarehouse : Devoir

Exercice 1 : Technique ETL

Le fichier « Devoir.xls » contient des enregistrements de tests médicaux appliqués sur 35 patients. Ce fichier est constitué de trois feuilles. Chaque feuille représente les résultats de Fréquence cardiaque (FC) ou Tension artérielle (TA) obtenus par ces tests.

L'objectif de ce travail est d'effectuer quelques actions sur ces enregistrements afin d'obtenir le format convenable afin d'être chargés dans des destinations spécifique

Travail à faire

- Créer un projet SSIS afin de réaliser l'objectif précisé.
- Cette solution SSIS doit respecter les consignes suivantes :
- La solution assurer les objectifs suivants :
 - **Tâche 1 :**
 1. Lire le contenu à partir des trois feuilles du fichier source « Devoir.xls ».
 2. Effectuer les transformations nécessaires qui vont vous permettre de réaliser les modifications suivantes :
 - Dans la première feuille excel « Test1 », ajouter une colonne "RéponseVagaleT1" qui permet de calculer la réponse vagale (RV) en se basant sur les champs FCmin et FCmax. Cette mesure est calculée en utilisant la formule suivante :

$$RV = (FCmax - FCmin) / FCmin * 100$$
 - Dans la deuxième feuille excel « Test2 », ajouter deux colonne "RéponseVagaleT2" et « RéponseAlphaT2 » qui permettent de calculer la réponse vagale et la réponse périphérique alpha en utilisant les formules suivantes :

$$RéponseVagaleT2 = (FCmax - FCmin) / FCmin * 100$$

$$RéponseAlphaT2 = (TAmix - TAmin) / TAmin * 100$$
 - Dans la troisième feuille excel « Test3 », ajouter deux colonne "RéponseBetaT3" et « RéponseAlphaT3 » qui permettent de calculer réponse centrale alpha et beta en utilisant les formules suivantes :

$$RéponseBetaT3 = (FCmax - FCmin) / FCmin * 100$$

$$RéponseAlphaT3 = (TAmix - TAmin) / TAmin * 100$$
 3. Charger le résultat dans une seule table SQL Server. Le contenu chargé doit être représenté dans la table comme suit :

ID	FC min T1	FC max T1	Répons eVagal eT1	FC min T2	FC max T2	Répons eVagal eT2	TA min T2	TA max T2	Répons eAlpha T2	FC min T3	FC max T3	Répo nseBe taT3	TA min T3	TA max T3	Répons eAlpha T3

- **Tâche 2 :**
 1. Extraire le contenu que vous venez de charger dans la table SQL Server.
 2. Appliquer les transformations convenables qui vont vous permettre de remplacer le contenu des colonnes ajoutées dans la première tâche. Il s'agit de remplacer les valeurs numériques par des valeurs discrètes en respectant les conditions suivantes :

- Pour la colonne "RéponseVagaleT1" :
 - Si RéponseVagaleT1<28 alors Faible
 - Si 28<=RéponseVagaleT1<32 alors Normal
 - Sinon alors Elevé
 - Pour les autres colonnes "RéponseVagaleT2", « RéponseAlphaT2 », "RéponseBetaT3" et « RéponseAlphaT3 » :
 - Si [Valeur]<8 alors Faible
 - Si 8<=[Valeur]<12 alors Normal
 - Sinon alors Elevé
3. Charger le contenu transformé dans les deux destinations suivantes :
- Une autre table SQL Server
 - Un fichier texte

Exercice 2 : Modélisation multidimensionnelle

Etude de cas : compagnie d'assurance

Une compagnie d'assurance de biens (automobile, immobilier, responsabilité civile) possède une application transactionnelle de production permet de gérer les polices (contrats) de ses clients ainsi que les sinistres (accidents) déclarés par ces clients.

Gestion des polices

Pour gérer les polices, les agents d'assurance peuvent effectuer les transactions suivantes :

- Créer, mettre à jour ou supprimer une police d'assurance
- Créer, mettre à jour ou supprimer un risque (pour une police donnée)
- Créer, mettre à jour ou supprimer des biens assurés (voiture, maison) sur un risque
- Chiffrer ou refuser le risque
- Valider ou refuser la police

On enregistre dans ces transactions un grand nombre d'informations, et notamment : date d'écriture (date de la transaction), date d'effet (date de début d'assurance), client (personne(s) privée(s), personne morale), opérateur (employé, agent: chiffrage, vérificateur : validation), risque (produit vendu par la compagnie d'assurance), couverture (description des biens assurés), police (numéro de police, « note » de la police ou du risque,...) , transaction (code transaction).

Gestion des sinistres

Pour gérer les sinistres déclarés par les clients, les agents d'assurance ont à leur disposition les transactions suivantes :

- Créer, mettre à jour ou supprimer une déclaration de sinistre
- Créer, mettre à jour ou supprimer une expertise
- Créer, mettre à jour ou supprimer des paiements
- Clore le sinistre

Ces transactions comportent notamment : date d'écriture (date de la transaction), date d'effet (date de déclaration), client, opérateur, risque, biens sinistrés, police, les tiers impliqués dans le sinistre, les montants financiers (limites, déjà payé, reste à payer, ...), code transaction.

Conception

A partir de cette application transactionnelle, on veut créer un Datawarehouse permettant de répondre aux questions suivantes :

- Pour chaque bien assuré, on veut connaître le montant de la prime (somme annuelle payée par le client pour assurer le bien) associée au bien assuré, et le nombre de transactions du mois pour ce bien.
- De même on veut pouvoir sortir des tableaux de bord par sinistre avec le total payé dans le mois et le total reçu dans le mois pour ce sinistre.

Travail à faire

1. Faire le schéma en étoile d'un Datamart « **Police** » ne prenant pas en compte les sinistres.
 - a. Identifier explicitement les faits et les attributs des dimensions.
 - b. Identifier les hiérarchies des dimensions.
 - c. Identifier le grain du Datamart « **Police** ».
2. De même, faire le schéma en étoile d'un Datamart « **Sinistre** ».
 - a. Identifier explicitement les faits et les attributs des dimensions.
 - b. Identifier les hiérarchies des dimensions.
 - c. Identifier le grain du Datamart « **Sinistre** ».
3. Faire un seul Datawarehouse de ces deux Datamarts.