

Examen de BI & Big Data

Durée 2 heures

Exercice 1

Pour ses besoins d'analyses, une entreprise pharmaceutique a élaboré le schéma de la base de données relationnelle suivante :

- **Médicament**(NumMédicament, Nom, Type, Famille). On supposera que le type de médicaments est plus spécifique que la famille. Par exemple, le type est 'Antibiotique anaérobie' et la famille 'Antibiotique'.
- **Patient** (NumPatient, Ville, Pays, Age, StatutImmunitaire, StatutMaladiesChroniques). *StatutImmunitaire* peut avoir les valeurs ('très bon', 'bon', 'déficient', 'très déficient') et *StatutMaladiesChroniques* peut avoir les valeurs ('aucune', 'une maladie', 'une famille de maladie', 'poly chronique').
- **Date** (NumDate, jour, Semaine, Mois, Année)
- **Effet** (NumMédicament, NumPatient, NumDate, DuréeTraitementEnjours, Dose, Résultat). L'attribut *Dose* représente la dose journalière en milligrammes du traitement, le résultat est un nombre représentant le nombre de jours depuis le début du traitement jusqu'à la fin des symptômes.

I. Dans le schéma ci-dessus, déterminer :

1. La table de faits et les mesures des faits. En justifiant votre réponse.
2. Les tables de dimension.
3. Proposer une hiérarchisation en termes de niveau de granularité des tables de dimension.
4. Pour chaque mesure de la table de faits dire si elle est additive, semi-additive ou non additive en justifiant vos réponses.
5. Proposer une mini-dimension démographique de la table *Patient* et redessiner le schéma de l'entrepôt ainsi modifié.

II. Quelle est la taille de l'entrepôt en termes de nombre de faits puis en octets (pour cette question on ne considère pas la mini-dimension démographique) en sachant que :

- Le nombre d'années considérées est de 10.
- Le nombre de patients est de 1000000.
- 10% des patients sont concernés par une prise de médicaments par an.
- Le nombre moyen de médicaments par patient concerné est de 2.
- La durée moyenne d'une prise de médicaments est de 7 jours.
- Les champs de la table de faits sont tous codés sur 4 octets.

5/5

1/2

Exercice 2

Soit la base de données opérationnelles ci-dessous :

Etudiant(NuméroEtudiant, Nom, Prénom, Filière)

Enseignement(NuméroEnseignement, Libellé, Domaine, *NuméroEnseignant*)

Enseignant(NuméroEnseignant, Nom, Prénom, Département)

Evaluation(NuméroEtudiant, NuméroEnseignement, Note, date)

1. Dresser un schéma en étoile permettant de représenter les dimensions étudiantes, enseignant et enseignement et les indicateurs Moyenne des notes et nombre des étudiants ayant réussi (moyenne > 10).
2. Donner la taille de la base de faits de l'entrepôt de données en supposant que le nombre des étudiants est de 10000 et que chaque étudiant passe en moyenne 20 évaluations par an. On supposera également que l'on veuille garder des données dans l'entrepôt pour une période de 10 ans et que chaque attribut de la table de faits est codé sur 4 octets.

8/11

2/2