

Exercice 1 – Gestion des stocks à l'hôpital

L'hôpital possède un nombre important de produits stockés dans différents magasins internes à l'hôpital (pharmacie générale, magasin de compresses, etc.) provenant de fournisseurs variés et mis à disposition des services. Pour chaque fournisseur, l'administration dispose de différents éléments : nom, SIRET, adresse... Pour chaque produit, le magasinier dispose d'information sur le nom du produit, son fournisseur, son conditionnement, la taille du conditionnement ... Pour les services utilisant les produits, il y a le nom du service, le responsable du service, le nombre de lits du service, ... Un produit est demandé par un demandeur qui est salarié de l'hôpital et qui appartient à un service. Chaque soir le manager de chaque magasin doit faire le bilan de ses produits sortis pour les différents services médicaux sous forme d'un tableur, des stocks restants et des demandes non satisfaites et l'envoi à l'administration centrale afin qu'elle effectue les commandes et analyse les données globales de consommation des produits par service. Cela doit également permettre de savoir si la gestion des stocks est optimale, si des produits sont parfois en rupture (ce qui n'est pas pratique pour les patients) ou si les produits sont jetés car périmés.

A cette fin, l'hôpital vous contacte pour mettre en place un entrepôt de données pour pouvoir analyser plus facilement les données de la consommation de produits au sein de ses services.

Question 1 : Proposez un schéma en étoile pour cet entrepôt de données. Vous justifierez vos choix et préciserez les faits et les mesures.

Question 2 : L'hôpital dispose de 500 fournisseurs, accueille 1000 patients chaque jour, gère 25000 références de produits différents et possède 2 millions de produits en stock, possède 15 services, et 250 personnes y travaillent. Le mouvement journalier des produits est de 55%. L'hôpital désire s'équiper de entrepôt de données pour stocker les données sur 10 ans.

Estimez la taille de l'entrepôt de données.

Exercice 2 Une maladie très bizarre

Le centre hospitalier spécialisé dans la lutte contre les maladies tropicales s'intéresse de près à la population des patients qui souffrent d'une infection parasitaire nommée BI. Il s'avère que certains doivent être opérés en urgence car l'évolution de la maladie est rapide, tandis que d'autres peuvent attendre car l'évolution de la maladie est lente ou que l'antibiothérapie fonctionne.

Ils observent certaines données pour voir la typologie de leur patient.

	Age	Taux de transaminase	Résultat du scanner	Présence de kystes	Coagulation
P1	40	10	+	-	Bonne
P2	25	80	+	+	Mauvaise
P3	50	100	+	+	Très mauvaise
P4	17	50	-	-	Bonne

Question 1. Définissez formellement une distance permettant de considérer tous les attributs. Donnez la distance de P1 à P2 avec la distance précédemment définie.

Question 2. Les étudiants en médecine ont intégré des données supplémentaires sur d'autres patients et propose de vous donner la matrice de distance suivante pour vous aider :

	P5	P6	P7	P8
P5	0	10	40	50
P6		0	20	60
P7			0	30
P8				0

Déterminer une segmentation des patients P5 à P8 en utilisant l'algorithme **complete link**. Vous indiquerez les étapes de l'algorithme.

Exercice 3 Une nouvelle maladie

Un médecin de l'hôpital voudrait comprendre comment déterminer si des personnes ont une maladie qui n'avait pas encore été analysée. Il vient vous voir avec les données suivantes :

Apprentissage	N	T	R	F	Classe
D1	+	+	+	-	Oui (malade)
D2	+	+	-	-	Oui
D3	-	-	+	+	Oui
D4	+	-	-	-	Non (bonne santé)
D5	-	-	-	-	Non
D6	-	+	+	-	Non

N = Nez qui coule, T = Tousse, R = Rougeur sur la peau, F = Fièvre

Question 1. Expliquez au médecin à quel type de tâche de datamining il doit faire appel.

Question 2. Proposez lui un modèle avec l'algorithme de votre choix pour résoudre son problème.

Exercice 4 Le retour de la maladie très bizarre

L'équipe FIL du Centre hospitalier de "Contingence" a été chargée d'évaluer les performances diagnostiques du dosage des *Masters1* pour le diagnostic de la maladie parasitaire de l'exercice 2. Pour cela, l'équipe a dosé les *Masters1* chez tous les patients arrivés aux urgences du Centre hospitalier et souffrant de douleur de l'hypocondre droit et de fièvre revenant de zones où sévit le parasite, et a confronté le résultat de chaque dosage au diagnostic finalement retenu. Au bout de 3 mois, 109 patients ont été inclus dans l'étude. 78 d'entre eux ont eu un dosage des *Masters1* "positif", c'est-à-dire supérieur au seuil habituellement utilisé (0,5 mg/l). Le diagnostic d'infection parasitaire BI a finalement été retenu chez 45 patients après la réalisation d'autres examens considérés comme des examens de référence. Parmi ces 45 patients, 2 patients avaient un dosage des *Masters1* "négatif"

Question 1. Donnez la matrice de contingence de ces données.

Question 2. Selon cette étude, quelle est la spécificité du dosage des *Masters1*, au seuil de 0,5 mg/l, pour le diagnostic d'infection parasitaire BI?