

EXAMEN DE BASES DE DONNEES MULTIDIMENSIONNELLES ET NOMADES

Exercice 1. LA LECTURE EN FRANCE

L'Insee réalise une étude sur la lecture en France. L'objectif est de réaliser des tables comme suit :

	Femmes	Hommes	16-24 ans	25-39 ans	40-59 ans	60 ans ou plus	Ensemble
<i>en %</i>							
Part des personnes ayant lu un livre (hors bande dessinée) dans l'année							
Au moins un livre	71	53	58	60	65	62	62
<i>dont au moins six livres</i>	43	27	29	30	36	40	35

On souhaite également pouvoir suivre l'évolution de ces indicateurs d'année en année, et pouvoir en étudier les variations sur le territoire métropolitain, au niveau national, régional, ou bien départemental. Pour cela, des questionnaires sont envoyés aux sondés, et les résultats sont intégrés à un entrepôt de données. Sur ces questionnaires figurent les questions suivantes :

Questionnaire

- Dans quelle tranche d'âge vous situez vous ?
16-24 25-39 40-59 60 ou plus
- Vous êtes : un homme une femme
- Indiquer le code de votre département de résidence : _____
- Au cours de l'année passée combien de livres (hors bandes-dessinées) avez-vous lu : _____

Référence du questionnaire : XGFSWERCRSERE

Dans cet exercice, on va construire un modèle dimensionnel pour le stockage et l'exploitation des réponses à ce questionnaire. Ce modèle sera structuré autour d'une unique table de faits. Les tables de faits et de dimension utiliseront des champs parmi les suivants (en plus des clés primaires et étrangères) :

- Tranche d'âge/sexe/région/département/nombre de livres lus
- référence questionnaire
- jour/mois/année de réception du questionnaire rempli.

Question 1. Choisir le nom et le grain de la table de faits.

Question 2. Choisir les dimensions

Question 3. Choisir les mesures. Pour chaque mesure, on indiquera (et on justifiera) si elle est additive, semi-additive ou non-additive ?

Question 4. Compléter les tables de dimension, en proposant des hiérarchies dimensionnelles.

Question 5. Proposer un modèle conceptuel de données complet.

Question 6. (Bonus) Proposer une ou plusieurs requêtes sql permettant de construire la table donnée en exemple ci-dessus.

Exercice 2. STRATEGIE DE SONDAGE

Les questionnaires servant à l'étude ci-dessus peuvent être envoyés aux sondés de diverses manières : par courrier, par email, ou bien via des messages sur les réseaux sociaux (Facebook, LinkedIn, Myspace). On souhaite comparer l'efficacité relative de ces différentes méthodes en termes de taux de réponse ainsi qu'en termes de délai de réponse. Pour cela on va introduire dans notre entrepôt une seconde table de faits, enregistrant les envois et les réceptions de questionnaires, ainsi que le mode d'envoi de chaque questionnaire.

Question 1. Choisir le nom et le grain de la table de faits.

Question 2. Choisir les dimensions, en mettant en avant les dimensions conformes de l'entrepôt.

Question 3. Choisir les mesures, ou justifier leur absence.

Question 4. Compléter les tables de dimension, et proposer des hiérarchies dimensionnelles.

Question 5. Proposer un modèle conceptuel de données complet.

Question 6. (Bonus) Proposer une requête permettant de construire la table suivante :

Mode d'envoi	Volume envoyé	Taux de retour	Délai moyen de réponse (en jours)
email	1000000	5%	5
courier	20000	10%	15
Facebook	2000000	45%	6
LinkedIn	10000	60%	2
myspace	5	20%	200