

Projet

Travail demandé pour un groupe de 4 étudiants et à remettre avant les examens semestriels:

1. Rapport final contenant analyse, implémentation, interrogation + Présentation en Powerpoint ou Latex.
2. Exposé oral de 10 minutes.

La composition du groupe de travail peut concerner différents groupes de la promotion.

L'objectif de ce projet est de concevoir, implémenter et interroger un entrepôt de données. Le choix du sujet est libre, **mais** cela doit être lié aux thèmes suivants. Il est bien évidemment possible de proposer d'autres sujets - dont la réalisation d'un entrepôt de données est pertinente.

- Réseaux Sociaux (Instagram, Twitter, Facebook, LinkedIn...).
- Journalisme Numérique (LeMonde, ...).
- Vidéo ou musique à la demande (Youtube, Netflix, Amazon, Spotify...).
- Commerce électronique (Amazon...).
- Expéditions et livraisons (DHL...).
- Sites d'annonces commerciales (LeBonCoin...).
- Infrastructures et services : télécommunications, eau, électricité (Orange, EDF, SFR ...).
- Réseaux de transport : aérien, terrestre, nautique (SNCF, AirAlgérie, ...).
- Jeux vidéos en-ligne (Poker...).
- Gestion de ressources humaines, recrutements, inscriptions (Université, Entreprise, ...).
- Chaînes de magasins ou cinémas (Carrefour, Gaumont...).
- Agence de location.
- Industrie (secteur automobile, pharmaceutique, ...).

Merci d'indiquer votre groupe [Ici](#).

Les traitements.

La mise en place d'un entrepôt de données dépend exclusivement des traitements permettant à une entreprise d'améliorer sa position dans le marché (e.g., la vente d'un produit) ou encore la logistique et la gestion de son fonctionnement interne (e.g., les inventaires, la chaîne de production). Il est donc impératif de ne pas stocker dans un entrepôt toutes les données disponibles dans les différentes sources (sauf pour des cas rares), et de l'alimenter seulement avec les données nécessaires aux requêtes analytiques. En outre, vu le coût de mise en place d'un système d'aide à la décision, il est obligatoire de modéliser avec une grande précision les actions / opérations dont l'analyse est la plus rentable (e.g., augmenter les ventes en proposant des promotions de produit pour une population ciblée).

L'analyse des traitements aura un poids très important dans l'évaluation du projet.

En cas de manque d'inspiration, allez à la recherche d'éléments pour le cas à étudier dans le Web ou encore dans les cours ou le modèle qui vous a été donné.

Répondre aux questions suivantes.

1. Réaliser une analyse complète du cas considéré (au moins 1 page). Quels sont les objectifs de l'entreprise (ou institution) considérée dans votre sujet ? Quelle est sa position sur le marché ? Quels services ou produits propose-t-elle ? Quelles sont ses formes de revenu ? Quelles informations seraient utiles pour la prise de décision au sein de l'entreprise ?
2. Indiquez les actions / opérations (e.g., ventes, livraisons) à tracer pour récupérer ces informations.
3. Pour chaque action / opération, proposez au moins trois traitements possibles (i.e., requêtes analytiques) permettant d'aider à la prise de décision sur le sujet.
4. Ordonnez les actions par ordre d'importance / rentabilité potentielle (e.g., augmentation des ventes vs. Utilisation optimale de l'espace de stockage dans le magasin).

La conception de l'entrepôt dépendra exclusivement des traitements que vous avez indiqués.

5. Identifiez les deux actions / opérations les plus importantes à analyser.
6. Pour chaque actions / opérations, concevez un data-mart indépendant (c'est à dire, un modèle en étoile)
 - a. Pour l'action / opération la plus importante, prévoyez une modélisation détaillée dans la table des faits.
 - b. Pour la deuxième action, prévoyez une modélisation moins détaillée (snapshot, transactions, record-update)
 - c. Pour chaque action / opération, donnez la liste des faits (ou mesures) à enregistrer dans l'entrepôt.
 - d. Dites s'il s'agit d'une mesure additive, semi-additive, ou non-additive.
7. Définissez les dimensions nécessaires aux modèles. Pour chaque action, indiquez au moins 5 dimensions.
 - a. Pour chaque dimension, prévoir une dizaine d'attributs.
8. Est-il possible de répondre aux traitements que vous avez indiqués avec le modèle que vous avez mis en place ? Expliquez pourquoi et comment.
9. Pour tester la pertinence de votre modèle, donner un exemple d'instance de l'entrepôt de données (2 ou 3 lignes par table suffisent).
10. Estimez la taille des tables de l'entrepôt (en terme du nombre de lignes) sur 12 mois. Est-ce que cette taille est raisonnable ? Justifiez votre réponse.

Attention : On ne vous demande pas de faire le modèle conceptuel des données (on se focalise directement sur le modèle logique de l'entrepôt).

L'implémentation et le requêtage

11. Implantez en ORACLE les tables de faits et des dimensions prévues par votre modèle. (**En cas de problème avec ORACLE utiliser SQL Live ou installer POSTGRES : www.postgresql.org/download**)
Utilisez des vues virtuelles pour les dimensions partagées.
12. Proposez 10 requêtes analytiques correspondant aux traitements que vous avez indiqués.
13. Donnez l'ensemble des vues matérialisées permettant de répondre à l'ensemble de vos requêtes.

