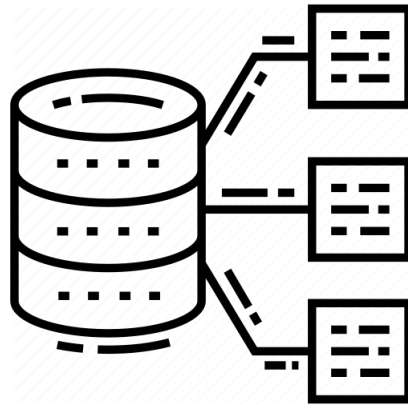




DEPARTEMENT INFORMATIQUE
DE LA FACULTE DES SCIENCES

Quentin Yeche (21520370), Yanis Allouch (21708237)

Rapport du Mini-Projet G°15 : Sanofi



HMIN122M — Entrepôts de données et Big-Data

Référent: Federico Ulliana et Anne-Muriel Chifolleau

2020

Table des matières

1	Une analyse complète du cas considéré	3
1.1	Quels sont les objectifs de l'entreprise (ou institution) considérée dans votre sujet ?	3
1.2	Quelle est sa position sur le marché ?	3
1.3	Quels services ou produits propose-t-elle ?	4
1.4	Quelles sont ses formes de revenu ?	4
1.5	Quelles informations seraient utiles pour la prise de décision au sein de l'entreprise ?	4
1.6	Action / opérations	5
2	La conception	7
2.1	Identifiez les deux actions / opérations les plus importantes à analyser.	7
2.2	Pour chaque actions / opérations, concevez un data-mart indépendant (c'est à dire, un modèle en étoile)	7
2.3	Définissez les dimensions nécessaires aux modèles.	9
2.4	Est-il possible de répondre aux traitements que vous avez indiqué avec le modèle que vous avez mis en place ?	11
2.5	Pour tester la pertinence de votre modèle, donner un exemple d'instance de l'entrepôt de données.	14
2.6	Estimez la taille des tables de l'entrepôt (en terme du nombre de lignes) sur 12 mois.	17
3	L'implémentation et le requêtage	19
3.1	Implantez en ORACLE les tables de faits et des dimensions prévues par votre modèle.	19
3.2	Proposez 10 requêtes analytiques correspondant aux traitements que vous avez indiqués.	21
3.3	Donnez l'ensemble des vues matérialisées permettant de répondre à l'ensemble de vos requêtes.	22
4	Annexe	23
	Références	27

Introduction

L'objectif du mini-projet est de concevoir, implémenter et interroger un entrepôt de donnée¹ sous SQL Oracle d'une entreprise réelle.

Travail demandé pour un groupe de 2 ou 3 personnes :

- Analyse des traitements (22 octobre, pour ce rendu considérez seulement les questions 1-9)
- Rapport final contenant analyse, implémentation, interrogation + transparents (7 novembre)
- Exposé oral de 10 minutes (5-6 novembre)

C'est dans le cadre de l'exploitation et l'application de nos connaissances acquises dans l'UE HMIN122M Entrepôt de Données et Big Data que nous réalisons ce mini-projet sur l'enseigne française Sanofi.

Voir les slides de cours référencés [Ull20b], [Ull20c], [Ull20d], [Ull20a], [Ull20e].

Il ne faudrait surtout pas oublier les excellentes explications de Kimball et Ross qu'on peut retrouver à travers leur livre [KR13] qui nous a été recommandé par et ainsi que Mr Ulliana F. qui nous a éclairer sur des détails de conception d'entrepôt de donnée.

Enfin, nous avons complété notre lecture par le livre de Christian S. Jensen et al, [JPT10] aussi recommander et utiliser comme ressource pédagogique.

Nous apportons à l'attention du lecteur que vous pouvez facilement retrouver la ressource citée grâce au lien incrusté au mot qui est [coloriés](#) pour vous diriger immédiatement sur la page qui la référence.

1. Traduction française du terme *data-warehouse*

1 Une analyse complète du cas considéré

Sanofi est une entreprise pharmaceutique française multinationale. Elle fait partie des leaders mondiaux du domaine, occupant régulièrement le haut du classement en matière de revenus annuels. Sa présence est globale. Depuis 2020 Sanofi est structuré en quatre GBU (Global Business Unit) : Médecine Générale, Médecine de Spécialité, Vaccins, et Santé Grand Public.

Note importante : Nous utiliserons souvent par la suite le terme "médicament" de manière générale pour désigner les produits proposés par Sanofi. Cependant certains produits, tels que les compléments alimentaires, ne sont pas des médicaments à proprement parler.

1.1 Quels sont les objectifs de l'entreprise (ou institution) considérée dans votre sujet ?

Sanofi est investi dans presque toutes les étapes du cycle de vie de ses solutions de santé. Ce cycle peut être découpé de la façon suivante :

1. Recherche et développement de nouvelles substances actives et de nouveaux médicaments ;
2. Processus réglementaire nécessaire avant la commercialisation sur les marchés : tests cliniques, demandes administratives ;
3. Fabrication des solutions de santé ;
4. Vente aux pharmacies et aux grossistes répartiteurs ;
5. Marketing.

On voit donc que la vente aux patients ne fait pas partie de ses activités. En effet, en France et dans de nombreux autres pays, seuls les pharmaciens peuvent vendre les médicaments aux patients. Cela peut être dans des officines indépendantes ou des PUI (Pharmacie à Usage Intérieur : pharmacie dans un établissement de santé, tel qu'un hôpital ou un EHPAD).

Le marketing pharmaceutique et le processus d'approbation des nouveaux médicaments sont des activités hautement réglementées. Leurs spécificités dépendent aussi selon le pays. Ces aspects ne sont donc pas très adaptés à notre cas d'étude. De plus la R&D ne génère pas une quantité de données importantes. Pour Sanofi la quantité de médicaments simultanément en développement ne dépasse pas la centaine. Cependant certaines facettes du développement pourront être importantes dans d'autres aspects de l'activité. Par exemple le coût total de développement d'un médicament est une information importante lorsqu'il s'agit d'évaluer le succès d'un produit qui est en vente.

La stratégie dévoilée en décembre 2019 par le Directeur Général de Sanofi, Paul Hudson, fait état de ses objectifs stratégiques [Sanc] :

- **Se concentrer sur la croissance.** Sanofi liste les produits pour lesquels elle prévoit une croissance des ventes : Dupixent® et le secteur Vaccins ;
- **Accélérer l'innovation.** Sanofi cible certains de ses projets R&D comme étant les plus prometteurs. L'accélération du développement des produits porteurs permet logiquement une commercialisation plus rapide qui mène à des bénéfices plus importants ;
- **Accroître l'efficacité opérationnelle.** Cela concerne d'une part la gestion de priorité de la R&D mais également d'autre part la gestion opérationnelle : fabrication, vente et distribution des médicaments.

1.2 Quelle est sa position sur le marché ?

L'entreprise est un leader mondial [Sanb] de la santé avec plus de 100 000 collaborateurs représentant 142 nationalités. Sanofi est présent dans 100 pays et ses solutions de santé sont

disponibles dans 172 pays. C'est 73 sites industriels répartis dans 32 pays qui fournissent donc des solutions de santé pour plus de 90% du globe!

En 2019 les ventes s'élevaient à plus de 4,5 milliards d'unités (médicaments et produits de santé grand public), pour un chiffre d'affaires total de 36,126 milliards d'euros. Les contributions à ce CA par zone géographique sont les suivantes :

- Amérique du Nord 35% ;
- Europe 24,5% ;
- Amérique latine 7,5% ;
- Eurasie 3,5% ;
- Reste du monde 29,5%.

1.3 Quels services ou produits propose-t-elle ?

Sanofi est composé de quatre grandes entités commerciales qui démarquent le champ d'action de l'entreprise :

- Médecine Générale : diabète, cardio-vasculaire et produits de prescription établis ;
- Médecine de Spécialité : sclérose en plaques, neurologie, autres maladies inflammatoires et immunologie, maladies rares, oncologie et maladies hématologiques rare ;
- Vaccins ;
- Santé Grand Public : allergies, santé digestive, compléments alimentaire, antidouleurs, etc.

1.4 Quelles sont ses formes de revenu ?

Leur principale forme de revenus selon le chiffre d'affaires est attribuée à *l'activité pharmaceutique*. On note en particulier que les entités Médecine de Spécialités et Médecine Générale représentent à elles seules 76% du chiffre d'affaires total.

Plus précisément, le chiffre d'affaires est détaillé par produit (médicament) et par zone géographique. Ces données sont disponibles à la page 52 de leur rapport financier semestriel 2020 [Sana]. On peut d'ailleurs voir que Dupixent[®], mentionné dans la stratégie de décembre 2020, est effectivement un produit-phare pour l'entreprise puisqu'il génère à lui seul un chiffre d'affaires qui rivalise avec des sections entières de l'offre (Maladies rares, Oncologie, etc.).

Sanofi s'occupe aussi de la vente de médicaments qu'elle ne possède pas directement mais au travers d'une entreprise : VaxServe (A Sanofi Pasteur Company) comme le Eloctate et Alprolix dans le cadre des accords avec Swedish Orphan Biovitrum AB (SOBI). Là encore, nous ne nous intéresserons pas à cette partie de l'activité, puisqu'elle est cloisonnée du reste.

1.5 Quelles informations seraient utiles pour la prise de décision au sein de l'entreprise ?

Bill Inmon définit dans son livre [JPT10] page 14, section 2.6, un entrepôt de données comme un "*subject oriented, integrated, time variant, non-volatile collection of data in support of management's decision making process.*"². Sanofi a besoin de pouvoir accéder aux données qui se concentrent sur la gestion opérationnelle de ses produits. Cette gestion opérationnelle inclut la fabrication, la gestion des stocks et la vente des produits.

2. sujet orienté, intégré, dépendant du temps, collection de données non volatiles pour l'aide à la décision

1.6 Action / opérations

Indiquez les actions / opérations (e.g., ventes, livraisons) à tracer pour récupérer ces informations.

Voici les actions et opérations à tracer pour récupérer ces informations :

- Vente d'un produit à une pharmacie ou à un grossiste répartiteur.
- Mouvement de stock de médicaments. Livraison entrante depuis une usine ou un autre entrepôt, livraison sortante pour livrer un client.
- Gestion des coûts de production. Coût d'opération d'une usine pour produire un médicament, coûts de transit et de livraison, coûts relatifs aux durées de stockage, pertes dues à la péremption de médicaments invendus.

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.

1. Vente

- Quels sont les chiffres d'affaires liés à chacun des produits vendus dans une zone géographique ?
- Quels sont les volumes de ventes liés à chacun des produits pour une période de l'année ?
- Quels sont les pays qui génèrent les plus gros chiffres d'affaires ?
- Quels sont les produits les plus populaires pour les pharmacies indépendantes ? Pour les répartiteurs grossistes ?
- Quelle est le chiffre d'affaires relatif aux produits stockés dans chaque entrepôt sur une même période ? Quels sont les entrepôts les plus rentables ?
- Quelle est l'évolution des chiffres d'affaires par produit par rapport à une période précédente ?
- Quel est le nombre de produits vendus par maladie, par tranche temporelle et par pays ?

2. Stock

- Quels sont les produits dont les stocks connaissent de fortes variations (des produits dont l'approvisionnement est fragile) ?
- Quels produits, relativement à leur demande moyenne, sont en pénurie ou proches de l'être dans une zone géographique ?
- Quel est, pour chaque produit, le temps moyen passé dans un entrepôt pour une unité ? Quel est donc le coût de stockage relatif à ce produit ?
- Quels sont les entrepôts ayant le plus de produits en stock à un jour donné ?

3. Production

- Quelle est la moyenne de livraison par jour, semaine, mois et par ville, région, pays, continent ?
- Quels sont les taux d'incidence de défauts par produit ?
- Quel est, pour chaque produit, le coût moyen associé à la production d'une unité ?
- Quels sont les coûts totaux associés à chacune des sources de coûts, fabrication, transit, stockage, livraison, pertes ?
- Quel est le bénéfice moyen par région ?
- Quels prestataires ont causées le plus de pertes en nombre d'occurrences ?

Ordonnez les actions par ordre d'importance / rentabilité potentielle

On définit l'ordre tel que le suivant :

1. Production
2. Stock

3. Vente

On espère, à travers les traitements précédemment proposés, fournir les outils pour que Sanofi puisse gérer avec une plus grande finesse ses stocks et ses coûts de production, et analyser ses chiffres de vente.

Cela dit, nous citons la page 35 du chapitre 3 de la référence [JPT10] « *Due to the dynamic nature of reality, the modeled reality as well as the uses of the data warehouse change over time.* »³ qui nous rappelle que notre modèle est destiné à évoluer pour suivre les besoins de l'entreprise.

3. Étant donnée la nature évolutive du monde réel, la modélisation ainsi que les utilisations de l'entrepôt de données s'adaptent dans le temps.

2 La conception

2.1 Identifiez les deux actions / opérations les plus importantes à analyser.

Les deux actions importantes à analyser sont la production et les stocks. Les ventes ne génèrent pas la même quantité de données (puisque la majorité du chiffre d'affaires vient de grosses ventes à des grossistes répartiteurs). De plus les réglementations encadrant la vente des médicaments limitent beaucoup le type de décisions que l'on pourrait prendre pour augmenter les ventes (e.g. prospection de nouveaux clients, marketing).

Pour la partie production, une mesure se dégage assez naturellement : le coût associé à chaque étape de vie du médicament : fabrication, transit, stockage, livraison, pertes dues aux défauts ou à la péremption, etc. Pour pouvoir optimiser sa marge, Sanofi a donc besoin de connaître le coût de chaque étape savoir où concentrer ses efforts.

Après notre entretien avec Mr Ulliana le 23 octobre 2020, nous adoptons donc les deux tables de faits suivantes :

1. Table de fait concernant la production. Cette table de fait est similaire à l'exemple d'Amazon qu'on peut retrouver dans les slides [U1120d] 28 à 36. On peut suivre un produit au long de son cycle de vie, depuis sa fabrication jusqu'à sa vente ou inscription en tant que perte. Puisque le nombre de références proposées par Sanofi est très faible par rapport à celles proposées par Amazon le volume de données ne sera pas exorbitant.
2. Table de faits sur les stocks. Cette table permet d'effectuer un suivi sur les quantités de médicaments dans chaque entrepôt ou par zone géographique.

2.2 Pour chaque actions / opérations, concevez un data-mart indépendant (c'est à dire, un modèle en étoile)

Action / Opération - Productions

La construction du data-mart, voir Figure 1, doit prendre en compte tous les éléments cités précédemment. La table de fait Production est une table transactionnelle. Ses dimensions sont les suivantes :

1. Dates. Chaque fait enregistré a une date associée : date de production, dates de stockage, dates de vente, etc.
2. Lieux. Chaque fait enregistré a un lieu associé. Ce lieu peut être soit une usine où a lieu la fabrication, soit un entrepôt.
3. Produits. Une ligne par référence proposée par Sanofi. Cette dimension possède de nombreuses colonnes décrivant le produit
4. Prestataires. Pour certaines étapes de la vie d'un produit Sanofi peut être amenée à confier certaines tâches à des prestataires. Deux exemples naturels sont la livraison et le transit.

Elle a pour mesure

- semi-additive, la quantité de produit associée à l'opération. En effet on ne peut pas sommer les quantités de produit pour des opérations différentes puisqu'une même boîte de médicaments pourra être comptée plusieurs fois au cours de sa vie.
- additive, le coût total d'une opération de X^4 pour un lieu, un produit, un prestataire et une date. d'opération. Le chiffre d'affaires d'une vente est donc un coût négatif.

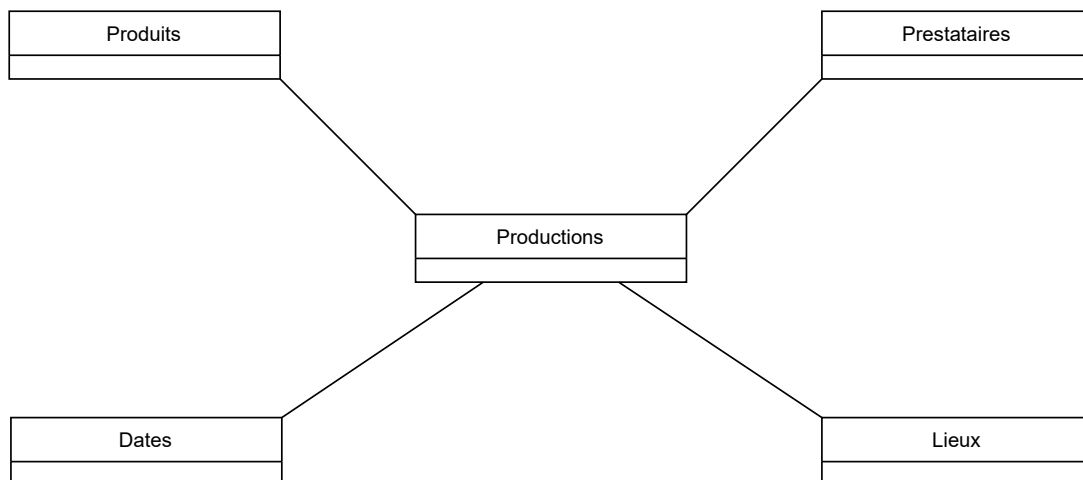


FIGURE 1 – Data-mart indépendant : Production

Table Productions

- | | |
|--|---|
| 1. id du produit (FK). | 7. flag type d'opération : |
| 2. id lieu de départ (FK). | {fabrication, transit, entrepôt, vente, |
| 3. id lieu d'arrivée (FK). | livraison, perte}. |
| 4. id prestataire (FK). | 8. mesure : quantité de produit (nombre |
| 5. id date de début de l'opération (FK). | de boîtes de médicaments). |
| 6. id date de fin de l'opération (FK). | 9. mesure : coût |

Action / Opération - Stocks

La construction du data-mart, voir Figure 2, doit prendre en compte tous les éléments cités précédemment. La table de fait Stocks est une table snapshot. Ses dimensions sont les suivantes :

1. Dates. Chaque fait enregistré a une date associée : date de production, dates de stockage, dates de vente, etc...
2. Produits. Une ligne par référence proposée par Sanofi. Cette dimension possède de nombreuses colonnes décrivant le produit.
3. Entrepôts. Possède beaucoup de colonnes, de sa désignation commerciale à ses spécificités techniques.

Comme le font remarquer Kimball & Ross au chapitre 4, page 113 [KR13]

« *The dimensions immediately fall out of this grain declaration : date, product, and store. This often happens with periodic snapshot fact tables where you cannot express the granularity in the context of a transaction* »⁵

Dans cette section les auteurs proposent également la mise en place d'une mesure pour les sorties ou les entrées de stock. Nous verrons que cette mesure sera nécessaire pour répondre aux questions que nous avons fixées. Les mesures que nous choisissons sont donc :

1. semi-additive : quantité de stock pour un produit, dans un entrepôt, à une date. En effet on ne peut pas sommer les stocks si les dates sont différentes.
2. additive : quantité sortante d'un produit depuis la ligne précédente (depuis la semaine précédente puisque nous choisissons un grain d'une semaine) pour un entrepôt.

4. Où X correspond à une opération de type fabrication, transit, livraison, vente, selon un flag.

5. Les dimensions se déduisent de façon immédiate du grain : date, produit et entrepôt. Ceci arrive souvent avec une table de faits snapshot, dans laquelle vous ne pouvez pas exprimer un grain plus précis.

Remarque : La quantité de stock entrante est la mesure duale de la quantité sortante. On peut la recalculer à partir de celle-ci et de la variation de stock.

Table Stocks

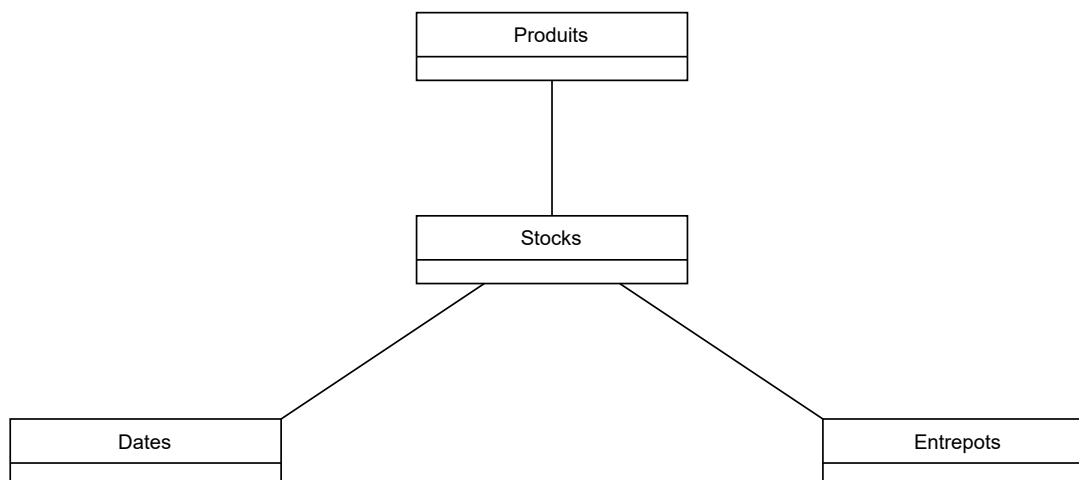


FIGURE 2 – Data-mart indépendant : Stocks

1. id date du snapshot (FK).
2. id de l'entrepôt (FK).
3. id du produit (FK).
4. mesure : stock.
5. mesure : quantité sortie de l'entrepôt durant la période.

2.3 Définissez les dimensions nécessaires aux modèles.

Pour chaque action, indiquez au moins 5 dimensions.

La dimension Produits :

Cette table ne contient pas de prix unitaire puisque les médicaments peuvent avoir un prix différent sur certains marchés. Ainsi, afin d'éviter d'avoir à décliner les produits dans tous les marchés dans lesquels ils sont vendus à un prix différent, nous nous contentons des données de chiffres d'affaires de la table des faits.

1. clé primaire.
2. id produit.
3. nom produit.
4. gamme produit
5. coût moyen de fabrication. (Note : Cette information dépend de l'historique de la table de faits)
6. GBU {MS, MG, VA, GP}
7. vendu par quantité.
8. poids total.
9. largeur de l'emballage.
10. longueur de l'emballage.
11. hauteur de l'emballage.
12. volume de l'emballage.

Les dimensions spatiales

Nous remarquons que la dimension de Lieux pour Productions et Entrepôts pour Stocks contiennent des informations très similaires. Elles partagent une majorité de leurs colonnes. On gèrera donc leur implémentation grâce à des vues qui contrôlent l'accès aux colonnes pour Lieux et aux lignes pour Entrepôt.

La dimension Lieux :

- | | |
|----------------------------|--|
| 1. clé primaire. | |
| 2. nom lieu. | 10. ville. |
| 3. désignation interne. | 11. code postal. |
| 4. nom du type de lieu | 12. numéro d'adresse. |
| 5. hémisphère {Sud, Nord } | 13. adresse. |
| 6. nom du continent. | 14. indication supplémentaire sur l'adresse. |
| 7. nom du pays. | 15. adresse complète. |
| 8. nom de la région. | 16. nombre de personnel. |
| 9. nom du département. | |

La dimension Entrepôts :

- | | |
|---|--|
| 1. clé primaire. | 10. nom de la région. |
| 2. id entrepôt. | 11. nom du département. |
| 3. nom entrepôt | 12. ville. |
| 4. aire de stockage effectif de l'entrepôt. | 13. code postal. |
| 5. nombre total de palettes stockables. | 14. numéro d'adresse. |
| 6. désignation interne. | 15. adresse. |
| 7. hémisphère {Sud, Nord } | 16. indication supplémentaire sur l'adresse. |
| 8. nom du continent. | 17. adresse complète. |
| 9. nom du pays. | 18. nombre de personnel. |

La dimension Dates :

Nous prenons ici une implémentation classique d'une dimension dates. Nous notons néanmoins qu'en raison du grain hebdomadaire choisi pour Stocks nous ferons usage d'une vue pour restreindre l'accès aux colonnes de cette dimension.

- | | |
|--|--|
| 1. clé primaire. | 13. date de la fin de la semaine. |
| 2. date. | 14. numéro de la semaine dans l'année. |
| 3. date complète. | 15. nom du mois. |
| 4. jour de la semaine. | 16. numéro du mois dans l'année. |
| 5. numéro de la semaine. | 17. date au format MM-AAAA. |
| 6. numéro du mois. | 18. numéro de semestre. |
| 7. numéro du jour du mois. | 19. numéro de trimestre. |
| 8. numéro du jour dans l'année. | 20. numéro d'année. |
| 9. numéro du jour dans le mois fiscal. | 21. numéro de la semaine fiscale. |
| 10. numéro du jour dans l'année fiscale. | 22. numéro de la semaine fiscale de l'année. |
| 11. dernier jour de la semaine. | 23. numéro du mois fiscal. |
| 12. dernier jour du mois. | 24. numéro du mois fiscal de l'année. |

La dimension Prestataires :

Il faut remarquer que la fonction des prestataires n'est réalisée que par 2 éléments, la livraison et non-identifié, pour une éventuelle donnée qui ne rentre pas dans ce premier cadre. L'énumération peut être étendue si besoin. Nous avons fait ce choix avec la perspective que Sanofi pourrait faire évoluer son management et déléguer d'autres services qui ne rentrent pas dans le cadre de l'étude à ce jour.

1. clé primaire.
2. identifiant interne.
3. nom.
4. nom complet.
5. nom du responsable.
6. préfixe du téléphone.
7. numéro de téléphone.
8. fonction {livraison, non-identifié}.

2.4 Est-il possible de répondre aux traitements que vous avez indiqué avec le modèle que vous avez mis en place ?

Expliquez pourquoi et comment.

Nous utilisons un code couleur qui permet au lecteur de facilement identifier un traitement réalisable ■ d'un traitement non-réalisable ▲ avec le modèle proposé précédemment. Étant donné que notre table de faits est un modèle complet suivant toutes les étapes de vie d'un produit, elle peut donc répondre à nos questions sur l'aspect Ventes. Cependant ce n'est pas l'optique du sujet, nous ne continuerons donc pas plus avant le développement de cette partie dans la suite du sujet.

1. Ventes

- Quels sont les chiffres d'affaires liés à chacun des produits vendus dans une zone géographique ?
 - Oui, toutes les données sont présentes dans le modèle. Le chiffre d'affaires se calcule sur les ventes regroupées par produits. On peut affiner les données résultant par zone géographique de la table de faits *Productions*.
- Quels sont les volumes de vente liés à chacun des produits pour une période de l'année ?
 - Oui, le chiffre d'affaires se calcule sur les ventes regroupé par produits. On peut affiner la donnée résultant par date de début et fin de la table de faits *Productions*.
- Quels sont les pays qui génèrent les plus gros chiffres d'affaires ?
 - Oui. Le chiffre d'affaires se calcule sur les ventes regroupées par zone géographique de la table de faits *Production*.
- Quels sont les produits les plus populaires pour les pharmacies indépendantes ? Pour les répartiteurs grossistes ?
 - ▲ Non. Après réflexion, cette question n'est pas pertinente pour notre problème.
- Quelle est le chiffre d'affaires relatif aux produits stockés dans chaque entrepôt sur une même période ? Quels sont les entrepôts les plus rentables ?

- Oui. On est capable de calculer le chiffre d'affaires sur chaque produit dans un entrepôt particulier et en sortir le chiffre d'affaires de l'entrepôt en question et enfin les trier. *Production*.
- Quelle est l'évolution des chiffres d'affaires par produit par rapport à une période précédente ?
 - Oui. La requête est triviale.
- Quel est le nombre de produits vendus par maladie, par tranche temporelle et par pays ?
 - ▲ Non. Après réflexion, cette question n'est pas pertinente pour notre problème.

2. Stocks

- Quels sont les produits dont les stocks connaissent de fortes variations (des produits dont l'approvisionnement est fragile) ?
 - Oui. Pour une période demandée de n semaines on calcule la moyenne des variations relatives de stocks s . On dispose de e la quantité sortante. On regarde donc les valeurs maximales pour

$$\frac{1}{n} \sum_{j=d_0}^{d_0+n} \frac{e(j)}{s(j)}$$

- Quels produits, relativement à leur demande moyenne, sont en pénurie où proches de l'être dans une zone géographique ?
 - Oui. On peut calculer et ordonner selon le temps estimé avant épuisement des stocks en divisant le stock actuel par la demande moyenne. Cette demande moyenne peut être la demande récente sur la période, historique ou historique sur la période.
- Quel est, pour chaque produit, le temps moyen passé dans un entrepôt pour une unité ? Et quel est donc le coût de stockage relatif à ce produit ?

La question comporte deux sous-questions. Nous avons deux analyses, qui sont les suivantes :

 - Oui. Pour la première partie de la question le temps moyen sur une période se calcule de la façon suivante :

$$\frac{1}{n} \sum_{j=d_0}^{d_0+n} \frac{s(j)}{e(j)}$$

- ▲ Non. Pour la seconde partie de la question, étant donné qu'on ne mesure pas le coût dans cette table de faits *Stocks*.
- Quels sont les entrepôts ayant le plus de produits en stock à un jour donné ?
 - Oui. Trivial.

3. Productions

- Quelle est la moyenne de livraisons par semaine, mois et par ville, région, pays et continent ?
 - Oui. On pourra utiliser la fonction *ROLLUP* sur les dimensions concernées.
- Quels sont les taux d'incidence de défauts par produit ?
 - ▲ Non. Les pertes sont répertoriées, mais non détaillées. On ne peut donc pas savoir l'origine/cause de la perte (défaut, stock introuvable, péremption, destruction accidentelle, vols, etc.).

- Oui, si on généralise la question en "Quels sont les taux d'incidence de perte par produit?".
- Quel est, pour chaque produit, le coût moyen associé à la production d'une unité?
 - Oui. Il faut sommer les coûts moyens pour chaque étape du cycle de vie du produit. En effet un même produit risquerait d'être compté plusieurs fois (s'il connaît plusieurs périodes de transit + stockage).
- Quels sont les coûts totaux associés à chacune des sources de coûts, fabrication, transit, stockage, livraison, pertes?
 - Oui. On peut faire un GROUP BY sur le flag de type d'opération (sans inclure les ventes).
- Quel est le bénéfice moyen par région?
 - Oui. On fait la moyenne de toutes les lignes par région. Mais cette question n'a de sens que si les régions ne s'échangent pas de produits.
- Quels prestataires ont causées le plus de pertes en nombre d'occurrences?
 - Oui. Un GROUP BY sur le prestataire, en ordonnant le nombre de pertes.

2.5 Pour tester la pertinence de votre modèle, donner un exemple d'instance de l'entrepôt de données.

La table Date ne sera pas ici donnée puisqu'elle est déjà décrite dans le cours. On voit bien que la table Entrepôt est une simple vue sur la table Lieux.

Lieux								
id	nomLieu	nomInterne	type	ville	codePostal	noAdresse	adresse	nbDePersonnel
1	entrepotAAA	FR-MTP-1	entrepôt	montpellier	34000	349	avenue du marché de gare	30
2	usine947	FR-PARIS-23	usine	Évry-Grégy-sur-Yerre	94700	0000	chemin du breuil	160
3	entrepotBBB	RUS-SOFIA-1	entrepôt	purus lorem	nulla sed sodales	11	malesuada	42

TABLE 1 – Exemple réduit d'instance de la table Lieux

Entrepôt								
id	nomLieu	nomInterne	type	ville	codePostal	numAdresse	adresse	nbDePersonnel
1	entrepotAAA	FR-MTP-1	entrepôt	montpellier	34000	349	avenue du marché de gare	30
3	entrepotBBB	RUS-SOFIA-1	entrepôt	purus lorem	nulla sed sodales	11	malesuada	42

TABLE 2 – Exemple réduit d'instance de la table Entrepôt

Prestataires								
id	identifiantInterne	nom	nomComplet	nomResponsable	prefixeTelephone	numéroTelephone	fonction	
1	2020PSD-FOO	Foo	SANOFI	Fontaine C.	7	(922) 973-2013	non-identifié	
2	2016KKC-UPS	UPS	UPS	Rodriguez K.	33	07 38 58 27 87	livraison	
3	2007REN-FED	Fedex	Federal Express	Schimmel A.	1	521-553-6292	livraison	

TABLE 3 – Exemple d'instance de la table Prestataires

Productions								
idProduit	idLieuDepart	idLieuxArriver	idPrestataire	idDateDebut	idDateFin	operation	qteProduit	cout
1	2	2	1	20201101	20201101	fabrication	974	1984034
2	2	2	1	20201101	20201101	fabrication	3449	1987987
3	2	2	1	20201101	20201101	fabrication	4657	97636
4	1	3	2	20201101	20201102	transit	4132	6458
5	2	2	1	20201101	20201101	fabrication	126456	23178
6	2	2	1	20201101	20201101	fabrication	152646	164867
7	2	2	1	20201101	20201101	fabrication	2464	45678
8	3	1	2	20201101	20201103	transit	4567	1262
9	3	3	3	20201101	20201101	transit	16146	486978
10	1	3	3	20201101	20201102	livraison	1264	7862
3	2	2	1	20201101	20201101	perte	67	0

TABLE 4 – Exemple d'instance de la table Productions

Stocks				
idDate	idEntrepot	idProduit	stock	quantiteSortie
20201106	1	1	4300	1200
20201106	1	2	2800	600
20201106	1	3	6700	3100
20201106	1	4	4500	900
20201106	1	5	0	0
20201106	1	6	3500	800
20201106	1	7	5000	600
20201106	1	8	24000	9000
20201106	1	9	0	0
20201106	1	10	0	8500
20201106	1	11	2000	100
20201106	2	1	0	0
20201106	2	2	1400	200
20201106	2	3	3000	800
20201106	2	4	2300	400
20201106	2	5	6700	800
20201106	2	6	2800	1200
20201106	2	7	2900	700
20201106	2	8	18000	6900
20201106	2	9	6000	800
20201106	2	10	52000	25000
20201106	2	11	900	200
20201113	1	1	3900	900
20201113	1	2	2200	1000
20201113	1	3	9600	4000
20201113	1	4	4500	600
20201113	1	5	0	0
20201113	1	6	2500	500
20201113	1	7		
20201113	1	8	24000	7600
20201113	1	9	0	0
20201113	1	10	32400	6400
20201113	1	11	1900	0
20201113	2	1	0	0
20201113	2	2	1100	200
20201113	2	3	3700	1200
20201113	2	4	2600	200
20201113	2	5	6300	1500
20201113	2	6	2200	900
20201113	2	7	3200	500
20201113	2	8	21000	6200
20201113	2	9	5200	500
20201113	2	10	75000	48000
20201113	2	11	1200	300

TABLE 5 – Exemple d’instance de la table Stocks

Produit											
id	idProduit	GBU	nomProduit	gamme	coutMoyen	quantite	poids	largeur	longueur	hauteur	volume
1	MG01	MG	LANTUS 100 UI/ml	LANTUS	5.24	5	204	80	130	40	416
2	MG02	MG	LANTUS SOLOSTAR 100 UI/ml	LANTUS	7.56	5	291	120	170	60	1224
3	MG03	MG	TOUJEO SOLOSTAR 300 UI/ml	TOUJEO	7.98	3	127	120	170	60	1224
4	MG04	MG	TOUJEO DOUBLESTAR 300 UI/ml	TOUJEO	7.86	3	127	120	170	60	1224
5	MS01	MS	AUBAGIO	AUBAGIO	56.35	28	85	60	90	20	108
6	MS02	MS	DUPIXENT 300 mg	DUPIXENT	89.02	1	154	70	120	20	168
7	GP01	GP	DOLIPRANE 2,4 % SANS SUCRE	DOLIPRANE	1.02	1	102	60	80	40	192
8	GP02	GP	DOLIPRANE 1000 mg adulte	DOLIPRANE	0.87	8	85	30	70	20	42
9	GP03	GP	XYZALL	XYZALL	1.13	14		70	100	40	280
10	VA01	VA	FLUZONE HIGH-DOSE QUADRIVALENT	FLUZONE	1.67	1	102	60	80	40	192
11	VA02	VA	IMOVAX POLIO	IMOVAX	0.86	1	108	60	180	40	192

TABLE 6 – Exemple d’instance de la table Produits

2.6 Estimez la taille des tables de l'entrepôt (en terme du nombre de lignes) sur 12 mois.

Nous nous sommes efforcés, autant que possible, de donner des chiffres les plus proches de la réalité en utilisant différentes sources concernant Sanofi. Certains de chiffres ont par ailleurs été reportés dans [l'étude de cas en section 1](#) dont voici un résumé servant à l'estimation sur 12 mois. En 2019, Sanofi

- possédait 100 000+ collaborateurs,
- basé dans 100 pays,
- disponible dans 172 pays,
- possédant 73 usines,
- répartie dans 32 pays.
- a vendu pour 4.5 milliards d'unité (médicament) en 2019,
- pour un chiffre d'affaires d'un total de 36,126 milliards €
- possédait au moins une cinquantaine de médicaments différents ⁶.

De la même façon que Kimball & Ross proposent de ne pas émettre d'hypothèse sur la table de fait vente de son exemple [KR13] (p40, ch2) « *Le trafic de la distribution fluctue de manière significative d'un jour à l'autre et nous avons besoin de connaître l'activité des transactions sur une période raisonnable. Une autre manière d'estimer le nombre de lignes de la table de faits pourrait consister à diviser le chiffre d'affaires brut annuel de la chaîne de distribution par le prix de vente unitaire moyen.* »

En nous basant sur le rapport financier du S1 2020 [Sana] ainsi que sur le prix des médicaments ⁷ nous avons trouvé, pour les 13 marques pour lesquelles le chiffre d'affaires est disponible, un total d'unités (boîtes) vendues d'environ 230 millions par an. Nous choisissons de doubler ce chiffre afin de prendre en compte les autres références (environ 30% du CA) pour lesquelles les données ne sont pas disponibles. On obtient donc une estimation finale de 460 millions de produits vendus par an.

Table Produits

Vu les données du rapport financier [Sana].

La table Produits est estimée à moins d'une dizaine de milliers de ligne (assez pessimistes). Une gamme (telle que Doliprane® ou Dupixent® se décline au maximum en une dizaine de variantes (selon le dosage, présentation, etc.) sachant que Sanofi propose moins d'un millier de gammes différentes.

La taille est plus que raisonnable en toutes proportions données par rapport à la table de faits.

Table Entrepôts

Cette information n'est pas directement donnée. On sait cependant que Sanofi est présent dans 100 pays. Cependant ce serait oublier qu'elle peut avoir plus d'un entrepôt dans le même pays pour pouvoir desservir des régions cruciales plus rapidement. On fait le choix arbitraire d'estimer à 50 entrepôts en moyenne par pays. Nous obtenons donc un total de 5 000 entrepôts soit autant de lignes dans la table. Cette estimation est en réalité généreuse lorsque l'on prend en compte le nombre de total de collaborateurs qui est de 100 000.

Cette taille est tout à fait raisonnable pour une dimension.

6. Nous rappelons à ce stade de notre rapport, qu'un médicament référence tout produit distribué par Sanofi. Sous cette effigie, nous retrouvons par exemple des compléments alimentaires.

7. sources : <https://www.drugs.com/price-guide/dupixent> et <https://eurekasante.vidal.fr/>

Table Stocks

Le snapshot est réglé sur une fréquence hebdomadaire.

Sanofi possède une dizaine de milliers de références dans 5 000 entrepôts différents. On a alors 50 000 000 lignes chaque semaine.

Une année possédant 52 semaines, nous avons donc 2 600 000 000 lignes générées en une année complète.

Est ce que cette taille est raisonnable ? Justifiez votre réponse.

C'est une table de faits snapshot, on ne peut pas en espérer moins.

Table Productions

Étant donné le chiffre d'affaires de 36 126 000 000 € par an, un article coûtant en moyenne 1446 €⁸, on calcule 25 millions de lignes de transactions par an au minimum car il manque presque 25% de la représentation du chiffre d'affaires dans le calcul du prix moyen.

Est ce que cette taille est raisonnable ? Justifiez votre réponse.

Mais cette estimation paraît plausible malgré que Sanofi ne fasse pas de la vente au détail de chacun de ces médicaments, on imagine bien une pharmacie ou bien un hôpital acheté certains médicaments les plus coûteux et par la même occasion un médicament traitant une maladie un peu plus rare, que lorsqu'un patient en a besoin.

En comparaison, Kimball et Ross avaient calculé une estimation de 2 milliards par an de transaction pour une chaîne de magasins.

Table Dates

Pour l'estimation de la table en question, nous avons pris en compte la remarque de M. Ulliana pour avoir un historique sur 20 ans dans le passé. Il se trouve par ailleurs que ces 20 ans correspondent à la durée maximale des brevets pharmaceutiques. Aussi nous voudrions analyser les données des jours à venir sur 20 ans au minimum. on peut donc estimer assez grossièrement la taille de la table de la façon suivante :

— $(366 * 40) = 14\,640$ lignes.

Table Lieux

Pour cette table nous devons déjà reprendre les données de la table Entrepôt qui contient donc 5000 lignes. On a également les usines qui représentent 73 sites industriels. On obtient donc un total de 5100 lignes environ.

Cette taille est raisonnable pour une dimension.

Table Prestataires

On peut estimer que le nombre de prestataires par pays ne dépasse pas 10. De plus on peut considérer une centaine de transporteurs transnationaux et globaux. On a donc un nombre de lignes total de 2000 lignes.

Encore une fois cette taille est raisonnable.

8. change rate 1 USD = 0.8596 EUR, 02 novembre 2020 16h20

3 L'implémentation et le requêtage

3.1 Implantez en ORACLE les tables de faits et des dimensions prévues par votre modèle.

```
1 DROP TABLE PRODUCTIONS;
2 DROP TABLE STOCKS;
3 DROP TABLE PRODUITS;
4 DROP TABLE PRESTATAIRES;
5 DROP TABLE LIEUX;
6 DROP TABLE DATES_TABLE;
7
8
9
10 CREATE TABLE PRODUITS (
11     id int PRIMARY KEY,
12     idProduit varchar(38),
13     nomProduit varchar(38),
14     gammeProduit varchar(38),
15     coutMoyen integer,
16     gbuType varchar(5) ,
17     venduParQuantite integer,
18     poidsTotal integer,
19     largeurEmballage integer,
20     longueurEmballage integer,
21     hauteurEmballage integer,
22     volumeEmballage integer,
23     constraint CHK_PRODUITS CHECK ( gbuType in ('MS', 'MG', 'VA', 'GP'))
24 );
25
26 CREATE TABLE PRESTATAIRES (
27     id int PRIMARY KEY,
28     idInterne varchar(38),
29     nom varchar(38),
30     nomComplet varchar(38),
31     nomResponsable varchar(38),
32     indicatif varchar(38),
33     telephone varchar(38),
34     fonction varchar(38),
35     constraint CHK_PRESTATAIRES CHECK (fonction in ('livraison', 'non-defini'))
36 );
37
38 CREATE TABLE LIEUX (
39     id int PRIMARY KEY,
40     idLieu int,
41     nom varchar(38),
42     designationInterne varchar(38),
43     aireDeStockageEffectifEntrepot int,
44     nbTotalPaletteStockable int,
45     typeLieu varchar(38),
46     hemisphere varchar(38),
47     continent varchar(38),
48     pays varchar(38),
49     region varchar(38),
50     departement varchar(38),
51     ville varchar(38),
52     codePostal varchar(38),
53     numAdresse int,
54     adresse varchar(38),
55     indicationSupplementaire varchar(38),
56     adresseComplete varchar(38),
57     nbPersonnel int,
58     constraint CHK_LIEUX_typeLieu CHECK (typeLieu in ('entrepot', 'usine', 'non-defini')),
59     constraint CHK_LIEUX_hemisphere CHECK (hemisphere in ('nord', 'sud'))
60 );
61
62 CREATE OR REPLACE VIEW ENTREPOTS AS
```

```

63 SELECT id, idLieu as idEntrepot, nom as nomEntrepot, designationInterne,
    aireDeStockageEffectifEntrepot, nbTotalPaletteStockable, hemisphere, continent, pays,
    region, departement, ville, codePostal, numAdresse, adresse, indicationSupplementaire,
    adresseComplete, nbPersonnel
64 FROM LIEUX WHERE typeLieu = 'entrepot';
65
66 CREATE TABLE DATES_TABLE (
67     idDate int PRIMARY KEY,
68     dateSQL DATE,
69     dateComplete varchar(38),
70     jourSemaine varchar(38),
71     numSemaine int,
72     numMois int,
73     numJourMois int,
74     numJourAnnee int,
75     numJourMoisFiscal int,
76     numJourAnneeFiscal int,
77     dernierJourSemaine varchar(38),
78     dernierJourMois varchar(38),
79     dateFinSemaine varchar(38),
80     numSemaineAnnee int,
81     nomMois varchar(38),
82     numMoisAnnee int,
83     dateFormatMM_AAAA varchar(38),
84     numSemestre int,
85     numTrimestre int,
86     numAnnee int,
87     numSemaineFiscal int,
88     numSemaineFiscalAnnee int,
89     numMoisFiscal int,
90     numMoisFiscalAnnee int
91 );
92
93 CREATE TABLE PRODUCTIONS (
94     idProduit number(8),
95     idLieuDepart number(8),
96     idLieuArrivee number(8),
97     idPrestataire number(8),
98     idDateDebut number(8),
99     idDateFin number(8),
100    typeOperation varchar(15),
101    quantiteDeProduit number(8,2),
102    coutOperation number(12,2),
103    constraint FK_PRODUCTIONS_produit foreign key (idProduit) references PRODUITS (id),
104    constraint FK_PRODUCTIONS_lieuDepart foreign key (idLieuDepart) references LIEUX (id),
105    constraint FK_PRODUCTIONS_lieuArrivee foreign key (idLieuArrivee) references LIEUX (id),
106    constraint FK_PRODUCTIONS_prestataire foreign key (idPrestataire) references PRESTATAIRES (
        id),
107    constraint FK_PRODUCTIONS_dateDebut foreign key (idDateDebut) references DATES_TABLE (idDate
        ),
108    constraint FK_PRODUCTIONS_dateFin foreign key (idDateFin) references DATES_TABLE (idDate),
109    constraint PK_PRODUCTIONS primary key (idProduit, idLieuDepart, idLieuArrivee, idPrestataire
        , idDateDebut, idDateFin, typeOperation),
110    constraint CHK_PRODUCTIONS_typeOperation CHECK (typeOperation in ('fabrication', 'transit',
        'entrepot', 'vente', 'livraison', 'perte', 'non-defini'))
111 );
112
113 CREATE TABLE STOCKS (
114     idDateSnapshot number(8),
115     idEntrepot number(8),
116     idProduit number(8),
117     stock number(8,2),
118     quantiteSortieDurantLaPeriode number(8,2),
119     constraint FK_STOCKS_dateSnapshot foreign key (idDateSnapshot) references DATES_TABLE (
        idDate),
120     constraint FK_STOCKS_entrepot foreign key (idEntrepot) references LIEUX (id),
121     constraint FK_STOCKS_produit foreign key (idProduit) references PRODUITS (id)
122 );

```

3.2 Proposez 10 requêtes analytiques correspondant aux traitements que vous avez indiqués.

```

1  -- Stocks --
2  -- 1
3  SELECT AVG(quantiteSortieDurantLaPeriode/stock) variation, nomProduit
4  FROM PRODUITS JOIN STOCKS ON STOCKS.idProduit = PRODUITS.id
5  JOIN DATES_TABLE ON STOCKS.idDateSnapshot = DATES_TABLE.idDate
6  WHERE (stock > 0)
7  GROUP BY nomProduit
8  ORDER BY variation DESC;
9  -- 2
10 SELECT nomProduit, stock/moyenne tempsAvantEpuisement
11 FROM Stocks JOIN
12 (SELECT AVG(quantiteSortieDurantLaPeriode) moyenne, idProduit FROM STOCKS GROUP BY idProduit)
13 M
14 ON M.idProduit = STOCKS.idProduit
15 JOIN Produits on Produits.id = Stocks.idProduit
16 JOIN DATES_TABLE ON Stocks.idDateSnapshot = DATES_TABLE.idDate
17 JOIN ENTREPOTS ON Stocks.idEntrepot = ENTREPOTS.id
18 WHERE numSemaine = 46 AND pays = 'france';
19 -- 3
20 SELECT Stocks.idProduit, nomProduit, SUM(stock/(quantiteSortieDurantLaPeriode*joursAvecSortie)
21 )
22 FROM Stocks
23 JOIN Produits on Stocks.idProduit = Produits.id
24 JOIN (SELECT idProduit, count(*) joursAvecSortie FROM Stocks WHERE
25 quantiteSortieDurantLaPeriode>0 GROUP BY idProduit) M
26 ON M.idProduit = Stocks.idProduit
27 WHERE quantiteSortieDurantLaPeriode >0
28 GROUP BY (Stocks.idProduit, nomProduit);
29 -- 4
30 SELECT idEntrepot, nomEntrepot, designationInterne, SUM(stock)
31 FROM STOCKS, DATES_TABLE, ENTREPOTS
32 WHERE idDateSnapshot = idDate
33 AND idDateSnapshot = '20201106'
34 AND idEntrepot = id
35 GROUP BY (idEntrepot)
36 ORDER BY stock DESC;
37 -- Productions --
38 -- 5
39 SELECT AVG(count(*))
40 FROM PRODUCTIONS, DATES_TABLE, LIEUX
41 WHERE typeOperation = 'livraison'
42 AND idDateDebut = idDate
43 AND idLieuDepart = LIEUX.id
44 GROUP BY ROLLUP(numSemaine, numMois, ville, region, pays, continent);
45 -- 6
46 SELECT perte, SUM(quantiteDeProduit), Productions.idProduit, perte/SUM(quantiteDeProduit)
47 FROM (SELECT idProduit, sum(quantiteDeProduit) perte FROM Productions
48 WHERE typeOperation = 'perte' GROUP BY idProduit) M
49 JOIN Productions ON M.idProduit = Productions.idProduit
50 GROUP BY Productions.idProduit, perte;
51 -- 7
52 SELECT idProduit, SUM(coutOperation) / quantiteDeProduit
53 FROM PRODUCTIONS
54 WHERE typeOperation != 'perte'
55 GROUP BY ( idProduit, quantiteDeProduit);
56 -- 8
57 SELECT typeOperation, SUM(coutOperation)
58 FROM PRODUCTIONS
59 GROUP BY (typeOperation);
60 -- 9
61 SELECT region, SUM(coutOperation)
62 FROM PRODUCTIONS JOIN Lieux ON idLieuArrivee = Lieux.id
63 GROUP BY region;
64 -- 10
65 SELECT idPrestataire, COUNT(*) as OccurencePerte

```

```
63 FROM PRODUCTIONS
64 WHERE typeOperation = 'perte'
65 GROUP BY (idPrestataire)
66 ORDER BY (OccurencePerte);
```

3.3 Donnez l'ensemble des vues matérialisées permettant de répondre à l'ensemble de vos requêtes.

Partie Stocks

Nous voyons que nos requêtes n'ont pas nécessité de GROUP BY sur Dates. L'ensemble des requêtes peut donc être effectué sur une vue matérialisée qui effectue un GROUP BY sur les colonnes idProduit et idEntrepot. Ce niveau de détail est nécessaire suffisant dans notre cas.

Partie Production

4 Annexe

```
1 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('1', 'MG01', 'MG', 'LANTUS 100 UIml', 'LANTUS', '5,24', '5', '204',
    '80', '130', '40', '416');
2 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('2', 'MG02', 'MG', 'LANTUS SOLOSTAR 100 UIml', 'LANTUS', '7,56',
    '5', '291', '120', '170', '60', '1224');
3 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('3', 'MG03', 'MG', 'TOUJEO SOLOSTAR 300 UIml', 'TOUJEO', '7,98',
    '3', '127', '120', '170', '60', '1224');
4 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('4', 'MG04', 'MG', 'TOUJEO DOUBLESTAR 300 UIml', 'TOUJEO', '7,86',
    '3', '127', '120', '170', '60', '1224');
5 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('5', 'MS01', 'MS', 'AUBAGIO', 'AUBAGIO', '56,35', '28', '85', '60',
    '90', '20', '108');
6 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('6', 'MS02', 'MS', 'DUIXENT 300 mg', 'DUIXENT', '89,02', '1',
    '154', '70', '120', '20', '168');
7 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('7', 'GP01', 'GP', 'DOLIPRANE 2,4 % SANS SUCRE', 'DOLIPRANE',
    '1,02', '1', '102', '60', '80', '40', '192');
8 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('8', 'GP02', 'GP', 'DOLIPRANE 1000 mg adulte', 'DOLIPRANE',
    '0,87', '8', '85', '30', '70', '20', '42');
9 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('9', 'GP03', 'GP', 'XYZALL', 'XYZALL', '1,13', '14', '1', '70',
    '100', '40', '280');
10 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('10', 'VA01', 'VA', 'FLUZONE HIGH-DOSE QUADRIVALENT', 'FLUZONE',
    '1,67', '1', '102', '60', '80', '40', '192');
11 INSERT INTO PRODUITS (id, idProduit, gbuType, nomProduit, gammeProduit, coutMoyen,
    venduParQuantite, poidsTotal, largeurEmballage, longueurEmballage, hauteurEmballage,
    volumeEmballage) VALUES('11', 'VA02', 'VA', 'IMOVAX POLIO', 'IMOVAX', '0,86', '1', '108',
    '60', '180', '40', '192');
12
13
14 INSERT INTO PRESTATAIRES (id, idInterne, nom, nomComple, nomResponsable, indicatif, telephone
    , fonction) VALUES('1', '2020PSD-F00', 'Foo', 'SANOFI', 'Fontaine C.', '7', '(922)
    973-2013', 'non-defini');
15 INSERT INTO PRESTATAIRES (id, idInterne, nom, nomComple, nomResponsable, indicatif, telephone
    , fonction) VALUES('2', '2016KKC-UPS', 'UPS', 'UPS', 'Rodriguez K.', '33', '07 38 58 27 87',
    'livraison');
16 INSERT INTO PRESTATAIRES (id, idInterne, nom, nomComple, nomResponsable, indicatif, telephone
    , fonction) VALUES('3', '2007REN-FED', 'Fedex', 'Federal Express', 'Schimmel A.', '1',
    '521-553-6292', 'livraison');
17
18 INSERT INTO LIEUX (id, nom, designationInterne, typeLieu, ville, codePostal, numAdresse,
    adresse, nbPersonnel) VALUES('1', 'entrepotAAA', 'FR-MTP-1', 'entrepot', 'montpellier',
    '34000', '349', 'avenue du marche de gare', '30');
19 INSERT INTO LIEUX (id, nom, designationInterne, typeLieu, ville, codePostal, numAdresse,
    adresse, nbPersonnel) VALUES('2', 'usine947', 'FR-PARIS-23', 'usine', 'Evry-GrEgy-sur-
    Yerre', '94700', '0000', 'chemin du breuil', '160');
20 INSERT INTO LIEUX (id, nom, designationInterne, typeLieu, ville, codePostal, numAdresse,
    adresse, nbPersonnel) VALUES('3', 'entrepotBBB', 'RUS-SOFIA-1', 'entrepot', 'purus lorem',
    'nulla sed sodales', '11', 'malesuada', '42');
21
```



```

22 INSERT INTO DATES_TABLE (idDate, dateSQL, dateComplete, jourSemaine, numSemaine, numMois,
    numJourAnnee) VALUES ('20201101', TO_DATE('20201101', 'YYYYMMDD'), '01 novembre 2020', '
    dimanche', '44', '11', '306');
23 INSERT INTO DATES_TABLE (idDate, dateSQL, dateComplete, jourSemaine, numSemaine, numMois,
    numJourAnnee) VALUES ('20201102', TO_DATE('20201102', 'YYYYMMDD'), '02 novembre 2020', '
    lundi', '45', '11', '307');
24 INSERT INTO DATES_TABLE (idDate, dateSQL, dateComplete, jourSemaine, numSemaine, numMois,
    numJourAnnee) VALUES ('20201103', TO_DATE('20201103', 'YYYYMMDD'), '03 novembre 2020', '
    mardi', '45', '11', '308');
25 INSERT INTO DATES_TABLE (idDate, dateSQL, dateComplete, jourSemaine, numSemaine, numMois,
    numJourAnnee) VALUES ('20201106', TO_DATE('20201106', 'YYYYMMDD'), '06 novembre 2020', '
    vendredi', '45', '11', '311');
26 INSERT INTO DATES_TABLE (idDate, dateSQL, dateComplete, jourSemaine, numSemaine, numMois,
    numJourAnnee) VALUES ('20201113', TO_DATE('20201113', 'YYYYMMDD'), '13 novembre 2020', '
    vendredi', '46', '11', '318');
27
28 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('1', '2', '2', '1', '
    20201101', '20201101', 'fabrication', '974', '4034,00');
29 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('2', '2', '2', '1', '
    20201101', '20201101', 'fabrication', '3449', '1987987');
30 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('3', '2', '2', '1', '
    20201101', '20201101', 'fabrication', '4657', '97636');
31 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('4', '1', '3', '2', '
    20201101', '20201102', 'transit', '4132', '6458');
32 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('5', '2', '2', '1', '
    20201101', '20201101', 'fabrication', '126456', '23178');
33
34 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('6', '2', '2', '1', '
    20201101', '20201101', 'fabrication', '152646', '164867');
35 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('7', '2', '2', '1', '
    20201101', '20201101', 'fabrication', '2464', '45678');
36 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('8', '3', '1', '2', '
    20201101', '20201103', 'transit', '4567', '1262');
37 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('9', '3', '3', '3', '
    20201101', '20201101', 'transit', '16146', '486978');
38 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('10', '1', '3', '3', '
    20201101', '20201102', 'livraison', '1264', '7862');
39 INSERT INTO PRODUCTIONS (idProduit, idLieuDepart, idLieuArrivee, idPrestataire, idDateDebut,
    idDateFin, typeOperation, quantiteDeProduit, coutOperation) VALUES('3', '2', '2', '1', '
    20201101', '20201101', 'perte', '67', '0');
40
41
42 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '1', '4300', '1200');
43 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '2', '2800', '600');
44 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '3', '6700', '3100');
45 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '4', '4500', '900');
46 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '5', '0', '0');
47 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '6', '3500', '800');
48 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '7', '5000', '600');
49 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '8', '24000', '9000');

```

```

50 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '9', '0', '0');
51 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '10', '0', '8500');
52 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '1', '11', '2000', '100');
53 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '1', '0', '0');
54 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '2', '1400', '200');
55 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '3', '3000', '800');
56 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '4', '2300', '400');
57 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '5', '6700', '800');
58 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '6', '2800', '1200');
59 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '7', '2900', '700');
60 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '8', '18000', '6900');
61 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '9', '6000', '800');
62 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '10', '52000', '25000');
63 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201106', '2', '11', '900', '200');
64 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '1', '3900', '900');
65 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '2', '2200', '1000');
66 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '3', '9600', '4000');
67 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '4', '4500', '600');
68 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '5', '0', '0');
69 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '6', '2500', '500');
70 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '7', '', '');
71 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '8', '24000', '7600');
72 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '9', '0', '0');
73 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '10', '32400', '6400');
74 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '1', '11', '1900', '0');
75 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '1', '0', '0');
76 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '2', '1100', '200');
77 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '3', '3700', '1200');
78 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '4', '2600', '200');
79 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '5', '6300', '1500');
80 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '6', '2200', '900');
81 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '7', '3200', '500');
82 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '8', '21000', '6200');
83 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '9', '5200', '500');

```

```
84 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,  
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '10', '75000', '48000');  
85 INSERT INTO STOCKS (idDateSnapshot, idEntrepot, idProduit, stock,  
    quantiteSortieDurantLaPeriode) VALUES('20201113', '2', '11', '1200', '300');
```

Références

- [JPT10] JENSEN, C. S., PEDERSEN, T. B. et THOMSEN, C. *Multidimensional Databases and Data Warehousing (Synthesis Lectures on Data Management)*. Morgan et Claypool Publishers, 2010.
- [KR13] KIMBALL, R. et ROSS, M. *The Datawarehouse Toolkit*. Wiley 3rd ed, 2013.
- [Sana] SANOFI. *Rapport financier Semestriel Edition 2020*. URL : https://www.sanofi.com/-/media/Project/One-Sanofi-Web/Websites/Global/Sanofi-COM/Home/fr/investisseurs/docs/2020_07_29_HY_financial_report_FR.pdf?la=fr&hash=013EE70D258AF65220EF0253587D7AF0. (en date du : 22.10.2020).
- [Sanb] SANOFI. *Sanofi : un leader mondial de la santé*. URL : <https://www.sanofi.com/fr/nous-connaître/sanofi-en-bref>. (en date du : 22.10.2020).
- [Sanc] SANOFI. *Stratégie : Sanofi*. URL : <https://www.sanofi.com/fr/investisseurs/connaître-sanofi/strategie>. (en date du : 22.10.2020).
- [Ull20a] ULLIANA, F. *Datawarehouse : Evolving Data*. 2020.
- [Ull20b] ULLIANA, F. *EDBD 1 : Intro*. 2020.
- [Ull20c] ULLIANA, F. *EDBD 2 Case Study Part : 1*. 2020.
- [Ull20d] ULLIANA, F. *EDBD 2 Case Study Part : 2*. 2020.
- [Ull20e] ULLIANA, F. *EDBD 5 Optimization*. 2020.