

1. Quel est l'objectif principal de la modélisation dimensionnelle dans le contexte de l'Informatique Décisionnelle et de l'entrepôt de données (Data Warehousing), tel que décrit dans les sources ?
  - A) Minimiser la redondance des données en utilisant des formes normales élevées (par exemple, 3NF).
  - B) Optimiser les performances des systèmes transactionnels (OLTP) pour l'insertion et la mise à jour rapides des données.
  - C) Faciliter et optimiser la création de rapports, de requêtes et d'analyses pour les applications de Business Intelligence.
  - D) Créer un modèle Entité-Relation complet représentant toutes les entités et relations de l'entreprise.
2. Les sources décrivent plusieurs types de relations dans la modélisation Entité-Relation (ER). Lesquels parmi les suivants sont des types de relations mentionnés ?
  - A) Relation identifiante
  - B) Relation non-identifiante optionnelle
  - C) Relation agrégée
  - D) Relation récursive
  - E) Relation plusieurs-à-plusieurs
3. Dans un modèle dimensionnel, comment les sources distinguent-elles fondamentalement un Fait d'une Dimension ?
  - A) Un Fait est une entité métier principale (comme un Client ou un Produit), tandis qu'une Dimension est une caractéristique descriptive de cette entité.
  - B) Un Fait est une mesure numérique d'une activité commerciale (comme une quantité vendue ou un montant), tandis qu'une Dimension fournit le contexte de cette mesure (qui, quoi, où, quand, comment).
  - C) Un Fait est une table normalisée contenant des attributs détaillés, tandis qu'une Dimension est une table dénormalisée contenant des identifiants uniques.
  - D) Un Fait est utilisé dans les systèmes transactionnels (OLTP), tandis qu'une Dimension est utilisée dans les systèmes analytiques (OLAP).

4. Selon les sources, quels sont les défis spécifiques mentionnés en lien avec les analyses avancées (prédictions, visualisation des données, découverte de données) ?
- A) Un manque de professionnels qualifiés en analyse de données.
  - B) Des inquiétudes parmi les responsables informatiques.
  - C) La difficulté à obtenir des données en temps réel à partir des systèmes sources.
  - D) La complexité du modèle Entité-Relation (ER).
  - E) La confusion entre les objectifs de capture et d'analyse dans les systèmes opérationnels.
5. Quels sont les types de relation logique standard entre une table de faits et une table de dimension dans un modèle dimensionnel ?
- A) Many-to-Many entre la table de faits et la table de dimension.
  - B) One-to-One entre la table de faits et la table de dimension.
  - C) One-to-Many entre la table de dimension (le côté "un") et la table de faits (le côté "plusieurs").
  - D) Many-to-One entre la table de dimension (le côté "plusieurs") et la table de faits (le côté "un").
6. Les sources décrivent différents types d'architectures de données en concurrence ou ayant évolué historiquement. Lesquelles parmi les suivantes sont mentionnées ?
- A) L'entrepôt de données d'entreprise (EDW)
  - B) Les data marts indépendants
  - C) L'architecture en bus de données (Kimball)
  - D) Le modèle Hub-and-Spoke (Inmon / CIF)
  - E) L'architecture analytique des données (ADA)
7. Dans la modélisation dimensionnelle, une clé de substitution (Surrogate Key) est souvent utilisée comme clé primaire dans une table de dimension au lieu de la clé naturelle du système source. Quel est l'une des principales raisons ou avantages de cette approche, implicite ou explicite dans les sources ?

- A) Les clés de substitution sont généralement des valeurs numériques plus petites et plus rapides à joindre que les clés naturelles, qui peuvent être des chaînes de caractères ou des clés composites complexes.
  - B) Elles permettent de gérer facilement les changements dans les attributs de dimension (Dimensions à Évolution Lente - SCD) en créant de nouvelles lignes avec une nouvelle clé de substitution pour suivre l'historique, tout en conservant la clé naturelle constante.
  - C) Elles garantissent l'unicité de chaque enregistrement dans la table de dimension et la cohérence des relations avec la table de faits.
  - D) Les clés de substitution sont nécessaires pour modéliser les dimensions à rôle multiple.
8. Dans la modélisation dimensionnelle, la granularité d'une table de faits est un concept crucial. Quels aspects la granularité détermine-t-elle ou impacte-t-elle, selon les sources ?
- A) Le nombre de tables de dimensions liées à la table de faits.
  - B) La précision avec laquelle les analyses peuvent être menées.
  - C) La taille potentielle de la table de faits dans l'entrepôt de données.
  - D) Les performances potentielles des requêtes analytiques.
  - E) Le type de fait (additif, semi-additif, non-additif).
9. Identifier les types de faits qui ne peuvent pas être additionnés sur toutes les dimensions selon les sources.
- A) Faits additifs (ex: quantité vendue)
  - B) Faits semi-additifs (ex: solde bancaire)
  - C) Faits non-additifs (ex: prix unitaire, ratios)
  - D) Faits transactionnels
10. Qu'est-ce qui caractérise un Schéma en Constellation (Multifact Star Model) ?
- A) Il est composé de plusieurs tables de faits qui partagent une ou plusieurs tables de dimensions communes.
  - B) C'est un schéma où toutes les tables de dimensions sont interconnectées indépendamment de la table de faits.
  - C) Il s'agit d'une version entièrement normalisée du Schéma en Étoile, utilisée principalement pour l'intégration des données.

- D) C'est un modèle multidimensionnel optimisé pour les calculs OLAP sur un seul cube de données.

11. Identifier les pièges courants à éviter lors de la justification d'un projet BI, selon les sources.

- A) Avoir des sponsors métier trop enthousiastes.
- B) Avoir le DSI (CIO) comme unique sponsor.
- C) Définir uniquement des bénéfices tangibles et quantifiables.
- D) Mal définir les avantages intangibles.
- E) Confondre la technologie BI avec la valeur métier.

12. L'ADA (Architecture Analytique des Données) est décrite comme un modèle hybride évolué. Quels sont les principes de modélisation qu'elle combine et certaines technologies qu'elle peut intégrer ?

- A) Principes de normalisation (par exemple, 3NF).
- B) Modèles dimensionnels.
- C) Bases de données transactionnelles classiques uniquement.
- D) Technologies de Big Data (Data Lakes, Lakehouses) et NoSQL.
- E) Bases de données HTAP (Hybrid Transactional/Analytical Processing).

13. Les sources décrivent plusieurs styles analytiques de BI. Lesquels parmi les suivants sont mentionnés ?

- A) Rapports et alertes prédéfinis.
- B) Analyse ad hoc et tableaux de bord.
- C) OLAP (Traitement Analytique en Ligne).
- D) Extraction de données pour intégration dans des feuilles de calcul.
- E) Analyse prédictive et découverte de données.
- F) Recherche de BI et BI mobile.

14. Quelles sont les caractéristiques d'une table de faits sans mesures directes (Factless Fact Table) selon les sources ?

- A) Elle enregistre la survenue d'un événement ou une relation entre dimensions.
- B) Elle ne contient pas de mesures numériques quantifiables directement (même si un compteur factice peut être ajouté).

- C) Elle est utilisée principalement pour stocker des données de référence (master data).
- D) Elle est utilisée pour modéliser des processus accumulatifs (Accumulating Snapshot).
- E) Un exemple donné est la présence d'un étudiant à un cours.

15. Quels sont les éléments clés à inclure dans une feuille de route BI selon les sources ?

- A) Les principales initiatives commerciales soutenues et les processus métier impliqués.
- B) Les livrables commerciaux de haut niveau attendus.
- C) La liste détaillée de toutes les tables et colonnes des sources de données.
- D) Les technologies qui seront introduites dans le cadre du programme BI.
- E) Les groupes métiers qui seront impliqués dans l'utilisation de la BI.

16. L'architecture Hub-and-Spoke est présentée comme une bonne pratique, mais son implémentation peut échouer en raison d'erreurs courantes. Parmi les options suivantes, lesquelles sont citées comme des erreurs typiques dans l'implémentation du modèle Hub-and-Spoke ?

- A) Concevoir l'entrepôt de données (EDW) de manière trop rigide, par exemple en utilisant un modèle strictement normalisé (3NF).
- B) Permettre aux utilisateurs métier d'accéder directement à l'EDW pour leurs analyses en libre-service.
- C) Construire une structure BI distincte (data mart, base de données) pour chaque rapport spécifique.
- D) Ne pas utiliser de clés de substitution (surrogate keys) dans les tables de dimensions.
- E) Concevoir des data marts soit trop vastes, soit trop spécialisés.

17. La gestion des données de référence (MDM) est abordée dans les sources. Pourquoi la MDM est-elle nécessaire et dans quelles zones problématiques potentielles peut-elle intervenir ?

- A) Elle résout les incohérences de données liées à la multiplicité des sources.
- B) Elle permet de gérer des dimensions conformes dans les EDW.

- C) Elle aide à synchroniser les données de référence entre les systèmes opérationnels.
- D) Elle assure la cohérence des données brutes directement dans les systèmes sources transactionnels.
- E) Elle peut nécessiter une application MDM autonome pour gérer des données de référence d'entreprise.

18. Selon les sources, quelles sont les raisons principales justifiant la séparation de l'entrepôt de données (EDW) et des systèmes sources opérationnels ?

- A) L'incompatibilité fondamentale entre les objectifs des systèmes (capture de données vs. analyse d'informations).
- B) Les incohérences fréquentes des données (formats, définitions, etc.) provenant de différentes sources opérationnelles.
- C) Les problèmes de qualité des données qui sont révélés lors du processus d'intégration.
- D) La difficulté technique et les contraintes (performance, sécurité) liées à l'accès direct et en temps réel à de nombreux systèmes sources.
- E) Le besoin de gérer l'évolution lente des dimensions (SCD - Slowly Changing Dimensions) indépendamment des systèmes opérationnels.

19. Les métadonnées sont décrites comme essentielles dans l'architecture de l'information. Quels sont les types de métadonnées et certaines de leurs sources ou utilisations mentionnées ?

- A) Les métadonnées techniques décrivent les données pour les outils logiciels (format, taille, mappages ETL, champs BI).
- B) Les métadonnées commerciales décrivent les données du point de vue de l'entreprise (contexte, couches sémantiques).
- C) Les métadonnées sont uniquement générées automatiquement par les outils ETL et BI.
- D) Les définitions des données (Data Definitions) sont une catégorie de métadonnées qui peuvent être documentées dans des dictionnaires ou référentiels.
- E) Les utilisateurs commerciaux doivent savoir ce que représentent les données, leur origine et leur signification (couvert par les métadonnées commerciales).

20. Quelles sont les caractéristiques ou utilisations des "Dimensions à Rôle Multiple" (Role-Playing Dimensions) et les approches recommandées pour les gérer ?

- A) Elles surviennent lorsqu'une dimension est utilisée plusieurs fois dans une même table de faits avec des significations différentes.
- B) L'une des approches consiste à créer une copie physique distincte de la table de dimension pour chaque rôle.
- C) L'approche de "meilleure pratique" consiste à utiliser une seule table de dimension et créer des vues logiques pour chaque rôle.
- D) Elles sont nécessaires pour gérer l'évolution lente des dimensions (SCD).
- E) Un exemple typique est l'utilisation de la dimension "Date" pour représenter la date de commande, la date d'expédition, et la date de livraison dans une table de faits "Commandes".

21. Concernant la Granularité dans une table de faits, quelle affirmation est la plus précise selon les sources ?

- A) La granularité fait référence au nombre total de tables de dimensions liées à une table de faits.
- B) Une granularité faible (par exemple, résumé quotidien) implique une table de faits plus grande mais permet des analyses plus précises.
- C) La granularité est le niveau de détail des données dans la table de faits et est importante car elle détermine la précision des analyses et impacte la taille/performance de la table.
- D) La granularité est un attribut des dimensions qui définit leur niveau de hiérarchie.

22. Le Schéma en Étoile est décrit comme ayant une table de faits centrale entourée de tables de dimensions dénormalisées. Le Schéma en Flocon, en revanche, normalise les hiérarchies dimensionnelles en tables séparées. Quel est un avantage clé du Schéma en Étoile par rapport au Schéma en Flocon pour les performances des requêtes analytiques ?

- A) Le Schéma en Étoile réduit la redondance des données dans les tables de dimensions par rapport au Schéma en Flocon.
- B) Le Schéma en Étoile gère de manière plus native les hiérarchies complexes sans nécessiter de jointures supplémentaires.

- C) Le Schéma en Étoile est généralement plus performant pour les requêtes analytiques car il nécessite moins de jointures pour relier la table de faits aux attributs de dimension.
- D) Le Schéma en Étoile est la structure recommandée pour implémenter des dimensions à rôle multiple.

23. Une Dimension à Rôle Multiple (Role-Playing Dimension) se produit lorsqu'une même dimension est utilisée plusieurs fois dans une seule table de faits, chaque fois avec un rôle différent (ex: Date de Commande, Date d'Expédition utilisant la dimension Date). Quelle est la bonne pratique recommandée dans les sources pour gérer ce scénario ?

- A) Créer une copie physique distincte de la table de dimension pour chaque rôle afin d'éviter toute ambiguïté.
- B) Utiliser une seule table de dimension et lier la table de faits à celle-ci via plusieurs clés étrangères, en créant des vues pour chaque rôle afin de simplifier l'accès métier.
- C) Stocker les dates directement comme attributs dans la table de faits sans utiliser de dimension Date séparée.
- D) Normaliser la dimension à rôle multiple en plusieurs sous-tables pour chaque rôle.

24. Selon les sources, quel est l'une des erreurs les plus courantes lors de l'implémentation du modèle Hub-and-Spoke qui mène à des problèmes de performance et de confusion pour les utilisateurs métier ?

- A) Utiliser une conception de l'EDW trop dénormalisée.
- B) Une conception de l'EDW trop rigide ou trop normalisée (3NF) qui rend les requêtes complexes et lentes.
- C) Créer trop peu de data marts, forçant les utilisateurs à effectuer des transformations complexes eux-mêmes.
- D) Ne pas utiliser de dimensions conformes.

25. L'ADA (Architecture Analytique des Données) propose un modèle de données hybride combinant deux principes de modélisation. Quels sont ces deux principes ?

- A) Modélisation Entité-Relation et Modélisation Objet-Relationnel.
- B) Principes de normalisation (par exemple, 3NF) et modèles dimensionnels.



- C) Modèles en étoile et Modèles en flocon.
- D) Modèles transactionnels et Modèles opérationnels.

26. L'un des pièges à éviter lors de la justification d'un projet BI mentionné dans les sources est la "Confusion entre la technologie BI et la valeur métier". En relation avec la modélisation dimensionnelle et les diagrammes, quelle erreur typique pourrait découler de ce piège ?

- A) Se focaliser excessivement sur la structure normalisée de l'EDW (modèle d'Inmon) au détriment des besoins analytiques des utilisateurs (modèle de Kimball).
- B) Croire que le choix d'un outil BI sophistiqué seul suffira à structurer les données et à fournir des analyses pertinentes, sans une architecture de données sous-jacente bien conçue (incluant la modélisation dimensionnelle).
- C) Utiliser des data marts indépendants sans dimensions conformes.
- D) Négliger la gestion des métadonnées dans le cadre d'intégration des données (DIF).

27. Selon la typologie des faits décrite dans les sources, identifiez l'exemple de fait Semi-additif parmi les options suivantes :

- A) La quantité d'unités vendues lors d'une transaction (sommable sur toutes les dimensions).
- B) Le prix unitaire d'un produit (non sommable).
- C) Le solde d'un compte bancaire à la fin du jour (sommable sur certaines dimensions, mais pas sur le temps).
- D) Le montant total des ventes pour une période donnée (sommable sur toutes les dimensions).

28. Quel est l'avantage principal de l'utilisation de Dimensions Conformées (Conformed Dimensions) dans une architecture BI comprenant plusieurs tables de faits ou data marts ?

- A) Elles permettent de normaliser les dimensions pour réduire l'espace de stockage.
- B) Elles garantissent que les analyses et les rapports provenant de différentes tables de faits ou data marts peuvent être comparés ou combinés de manière cohérente, car les dimensions partagées ont les mêmes définitions et attributs.

- C) Elles sont nécessaires pour gérer l'évolution lente des dimensions (SCD - Slowly Changing Dimensions).
- D) Elles facilitent la création de tables de faits accumulatives.

29. Décrivez ce qu'est une table de faits sans mesures directes (Factless Fact Table) selon les sources :

- A) Une table qui contient uniquement les clés primaires de toutes les dimensions liées.
- B) Une table de faits qui enregistre la survenue d'un événement ou une relation entre dimensions (comme une présence à un cours), sans avoir de mesures numériques quantifiables directes, bien qu'un compteur factice puisse être ajouté.
- C) Une table de dimensions qui n'est pas liée à une table de faits, utilisée pour stocker des données de référence indépendantes.
- D) Une table de faits qui contient uniquement des mesures non-additives.

30. Dans le contexte de l'Architecture Analytique des Données (ADA) ou du modèle Hub-and-Spoke, quel est le rôle typique des Data Marts par rapport à l'Entrepôt de Données d'Entreprise (EDW) central ?

- A) Les Data Marts servent de zone de staging pour les données brutes avant leur chargement dans l'EDW.
- B) L'EDW est une base de données transactionnelle (OLTP) à partir de laquelle les Data Marts extraient des données pour les analyses.
- C) L'EDW centralise et intègre les données de l'entreprise, et les Data Marts sont créés à partir de l'EDW pour fournir des vues des données structurées et optimisées pour les besoins analytiques spécifiques des différents métiers ou fonctions (les "spokes").
- D) Les Data Marts remplacent l'EDW et contiennent toutes les données de l'entreprise pour des analyses ad hoc.