

QCM Avances - Volume 2

Business Intelligence et Reporting

100 Questions Supplémentaires

Ecole des Sciences de l'Information (ESI)
3ACI - ICSD | 2024-2025

Contents

1 ETL Avance et Qualite des Donnees (Q1-25)	2
2 Modelisation Dimensionnelle Avancee (Q26-50)	8
3 Requetes MDX et Analyse (Q51-65)	14
4 Approches et Methodologie (Q66-80)	18
5 Tendances et Technologies Modernes (Q81-100)	22

1 ETL Avance et Qualite des Donnees (Q1-25)

Q1 - Processus ETL

Quel est l'ordre correct des phases ETL?

- A) Load, Transform, Extract
- B) Transform, Extract, Load
- C) **Extract, Transform, Load**
- D) Extract, Load, Transform

Q2 - ELT vs ETL

Dans quel contexte l'approche ELT est-elle PREFEREE a ETL?

- A) Donnees structurees de petits volumes
- B) **Big Data avec Data Lake et puissance de calcul cote destination**
- C) Systemes legacy anciens
- D) Donnees temps reel critiques

Q3 - Extraction incrementale

L'extraction incrementale consiste a:

- A) Extraire toutes les donnees a chaque fois
- B) **Extraire uniquement les donnees modifiees depuis la derniere extraction**
- C) Extraire les donnees par petits morceaux aleatoires
- D) Ne jamais extraire les memes donnees

Q4 - Detection des changements

Pour detecter les enregistrements modifies, quelle technique est la plus fiable?

- A) Comparer toutes les colonnes manuellement
- B) **Utiliser une colonne timestamp de derniere modification ou CDC (Change Data Capture)**
- C) Demander aux utilisateurs de signaler les changements
- D) Recharger tout systematiquement

Q5 - CDC

Que signifie CDC dans le contexte ETL?

- A) Central Data Center
- B) Change Data Capture (capture des données modifiées)**
- C) Controlled Data Cleaning
- D) Certified Data Compliance

Q6 - Qualité des données

Parmi ces problèmes, lequel est un problème de QUALITE des données?

- A) Le serveur est trop lent
- B) Des doublons dans la table clients avec des orthographies différentes**
- C) Le budget est insuffisant
- D) L'interface est peu ergonomique

Q7 - Deduplication

La "deduplication" ou "dedoublonnage" consiste à:

- A) Dupliquer les données pour sécurité
- B) Identifier et fusionner les enregistrements en double**
- C) Supprimer toutes les données
- D) Créer des copies de sauvegarde

Q8 - Standardisation

La standardisation des données dans l'ETL inclut:

- A) Chiffrer toutes les données
- B) Uniformiser les formats (dates, adresses, noms) selon des règles définies**
- C) Supprimer les colonnes inutiles
- D) Compresser les fichiers

Q9 - Donnees manquantes

Comment traiter une valeur manquante (NULL) pour le champ "ville" d'un client?

- A) Toujours supprimer l'enregistrement
- B) Selon le contexte: valeur par defaut "Non renseigne", imputation, ou exclusion
- C) Toujours mettre "Paris"
- D) Ignorer completement le probleme

Q10 - Validation des donnees

Une regle de validation "l'age doit etre entre 0 et 150" est une verification de:

- A) Unicite
- B) Plage de valeurs (range check)
- C) Format
- D) Integrite referentielle

Q11 - Integrite referentielle

L'integrite referentielle garantit que:

- A) Les donnees sont chiffrees
- B) Chaque cle etrangere correspond a une cle primaire existante
- C) Les dates sont au bon format
- D) Les noms sont en majuscules

Q12 - Profilage des donnees

Le "data profiling" (profilage) permet de:

- A) Creer des profils utilisateurs
- B) Analyser la structure, le contenu et la qualite des donnees sources
- C) Optimiser les performances serveur
- D) Generer des rapports automatiques

Q13 - Mapping des données

Le "data mapping" dans un projet ETL définit:

- A) La localisation géographique des serveurs
- B) La correspondance entre les champs sources et les champs cibles**
- C) Le plan du bâtiment informatique
- D) Les horaires de travail de l'équipe

Q14 - Transformation lookup

Une transformation "lookup" permet de:

- A) Supprimer des lignes
- B) Enrichir un flux avec des données d'une table de référence**
- C) Compresser les données
- D) Chiffrer les informations sensibles

Q15 - Aggregation ETL

L'aggregation dans l'ETL consiste à:

- A) Diviser les données en petits morceaux
- B) Calculer des sommes, moyennes ou comptages sur des groupes de lignes**
- C) Fusionner plusieurs tables
- D) Trier les données

Q16 - Chargement initial vs incremental

Le chargement initial (full load) est utilisé:

- A) À chaque exécution quotidienne
- B) Lors de la première mise en place du DW ou après une reconstruction**
- C) Uniquement pour les petites tables
- D) Jamais en production

Q17 - Parallelisation ETL

La parallelisation dans l'ETL permet de:

- A) Reduire la qualite des donnees
- B) Accelerer le traitement en executant plusieurs flux simultanement**
- C) Simplifier le code
- D) Reduire les couts de licence

Q18 - Journalisation ETL

La journalisation (logging) des jobs ETL est importante pour:

- A) Decorer les rapports
- B) Tracer les erreurs, auditer les traitements et faciliter le debugging**
- C) Augmenter la vitesse
- D) Reduire l'espace disque

Q19 - Gestion des erreurs

Que faire quand une ligne source contient une erreur de format?

- A) Arreter completement le job ETL
- B) Ignorer silencieusement l'erreur
- C) Isoler la ligne dans une table d'erreurs pour traitement ultérieur**
- D) Modifier la source automatiquement

Q20 - Orchestration ETL

L'orchestration des jobs ETL permet de:

- A) Jouer de la musique pendant le chargement
- B) Planifier, sequencer et surveiller l'execution des differents flux**
- C) Dessiner des diagrammes
- D) Compresser les fichiers

Q21 - Outil ETL

Quel outil est spécifiquement conçu pour l'ETL?

- A) Microsoft Word
- B) Talend, Informatica, ou SSIS (SQL Server Integration Services)**
- C) Adobe Photoshop
- D) Slack

Q22 - SSIS

SSIS est l'outil ETL de:

- A) Oracle
- B) SAP
- C) Microsoft (SQL Server)**
- D) IBM

Q23 - Informatica

Informatica PowerCenter est principalement utilisé pour:

- A) La visualisation de données
- B) L'intégration et la transformation de données (ETL entreprise)**
- C) La création de sites web
- D) La gestion de projet

Q24 - Data lineage

Le "data lineage" (lignage des données) permet de:

- A) Mesurer la taille des données
- B) Tracer l'origine et les transformations subies par chaque donnée**
- C) Compter le nombre de lignes
- D) Chiffrer les données

Q25 - Metadonnees ETL

Les metadonnees dans un processus ETL incluent:

- A) Les donnees clients
- B) **Les definitions des sources, transformations, mappings et statistiques d'execution**
- C) Les rapports generes
- D) Les mots de passe utilisateurs

2 Modelisation Dimensionnelle Avancee (Q26-50)

Q26 - Table de faits

Une table de faits contient principalement:

- A) Des descriptions textuelles detaillees
- B) **Des mesures numeriques et des cles etrangeres vers les dimensions**
- C) Uniquement des dates
- D) Des images et documents

Q27 - Table de dimension

Une table de dimension contient principalement:

- A) Des mesures numeriques a agreger
- B) **Des attributs descriptifs pour contextualiser les faits**
- C) Des calculs complexes
- D) Des transactions en temps reel

Q28 - Dimension conforme

Une "dimension conforme" (Kimball) est:

- A) Une dimension non utilisee
- B) **Une dimension partagee identiquement entre plusieurs datamarts**
- C) Une dimension temporaire
- D) Une dimension sans hierarchie

Q29 - Avantage dimensions conformes

L'avantage principal des dimensions conformes est:

- A) Reduire l'espace disque
- B) **Permettre des analyses croisees coherentes entre datamarts (drill-across)**
- C) Accelerer les requetes
- D) Simplifier la securite

Q30 - Fait transactionnel

Un fait "transactionnel" correspond a:

- A) Une synthese mensuelle
- B) **Un evenement ponctuel au grain le plus fin (ex: une ligne de commande)**
- C) Un solde de compte
- D) Une prevision budgetaire

Q31 - Fait periodique

Un fait "periodique" (periodic snapshot) represente:

- A) Une transaction individuelle
- B) **L'etat d'une mesure a intervalles reguliers (ex: stock en fin de journee)**
- C) Un evenement aleatoire
- D) Une prevision future

Q32 - Fait accumulating

Un fait "cumulatif" (accumulating snapshot) est utilise pour:

- A) Stocker des transactions simples
- B) **Suivre un processus avec plusieurs etapes (ex: cycle de vie d'une commande)**
- C) Calculer des moyennes
- D) Archiver des donnees anciennes

Q33 - Junk dimension

Une "junk dimension" (dimension poubelle) regroupe:

- A) Les donnees a supprimer
- B) **Plusieurs attributs de faible cardinalite (flags, indicateurs) en une seule dimension**
- C) Les erreurs de traitement
- D) Les donnees temporaires

Q34 - Role-playing dimension

Une "role-playing dimension" est:

- A) Une dimension pour les jeux video
- B) **Une meme dimension utilisee plusieurs fois avec des roles differents (ex: Date commande, Date livraison)**
- C) Une dimension sans donnees
- D) Une dimension de test

Q35 - Outrigger

Un "outrigger" dans un schema dimensionnel est:

- A) Un type de mesure
- B) **Une table secondaire liee a une dimension (normalisation partielle)**
- C) Un outil de reporting
- D) Un type de cube

Q36 - Bridge table

Une "bridge table" (table pont) est utilisee pour gerer:

- A) Les connexions reseau
- B) **Les relations plusieurs-a-plusieurs entre faits et dimensions**
- C) Les erreurs ETL
- D) Les sauvegardes

Q37 - Hierarchie desequilibree

Une hierarchie "ragged" (desequilibree) est:

- A) Une hierarchie avec des niveaux fixes
- B) **Une hierarchie ou certaines branches ont moins de niveaux que d'autres**
- C) Une hierarchie temporelle
- D) Une hierarchie sans parent

Q38 - Parent-child hierarchy

Une hierarchie parent-enfant est adaptee pour:

- A) Les dates uniquement
- B) **Les structures organisationnelles ou le nombre de niveaux est variable**
- C) Les produits uniquement
- D) Les mesures numeriques

Q39 - Dimension mini

Une "mini-dimension" est creee pour:

- A) Reduire le nombre de dimensions
- B) **Isoler les attributs qui changent frequemment d'une grande dimension**
- C) Stocker les petites valeurs
- D) Compresser les donnees

Q40 - SCD Type 3

Une SCD Type 3 permet de:

- A) Conserver tout l'historique
- B) Ecraser l'ancienne valeur
- C) **Conserver uniquement la valeur precedente (ajout d'une colonne "ancien_xxx")**
- D) Creer une nouvelle table

Q41 - SCD Type 4

Une SCD Type 4 utilise:

- A) Une seule table avec ecrasement
- B) **Une table principale courante + une table d'historique separatee**
- C) Trois colonnes par attribut
- D) Aucune historisation

Q42 - SCD Type 6

La SCD Type 6 (hybride) combine:

- A) Type 1 uniquement
- B) Type 2 uniquement
- C) **Types 1, 2 et 3 (valeur courante, historique complet, valeur precedente)**
- D) Aucune des reponses

Q43 - Date de validite

Dans une SCD Type 2, les colonnes date_debut et date_fin servent a:

- A) Calculer la duree du projet
- B) **Delimiter la periode de validite de chaque version de l'enregistrement**
- C) Enregistrer les dates de creation ETL
- D) Stocker des dates de naissance

Q44 - Indicateur courant

Dans une SCD Type 2, la colonne "est_courant" (is_current) indique:

- A) Si la donnee est electrique
- B) **Si l'enregistrement est la version active/la plus recente**
- C) Le courant electrique du serveur
- D) La devise monetaire

Q45 - Bus matrix

La "bus matrix" de Kimball est un outil pour:

- A) Planifier les trajets en bus
- B) **Documenter les dimensions conformes partagees entre processus metier**
- C) Calculer les performances
- D) Gerer les utilisateurs

Q46 - Grain declaration

La "declaration du grain" doit etre faite:

- A) Apres le chargement des donnees
- B) A la fin du projet
- C) **Au tout debut de la conception, avant de choisir les dimensions et mesures**
- D) Jamais explicitement

Q47 - Conformed fact

Un "fait conforme" est:

- A) Un fait sans erreur
- B) **Une mesure definie et calculee de maniere identique dans tous les datamarts**
- C) Un fait obligatoire
- D) Un fait temporaire

Q48 - Late arriving dimension

Une "late arriving dimension" (dimension tardive) est:

- A) Une dimension creee en retard dans le projet
- B) **Un cas ou le fait arrive avant les informations completes de la dimension**
- C) Une dimension lente
- D) Une dimension archivee

Q49 - Late arriving fact

Un "late arriving fact" (fait tardif) est:

- A) Un fait calcule apres coup
- B) **Une transaction arrivee apres la periode de chargement normale**
- C) Un fait avec valeur NULL
- D) Un fait supprime

Q50 - Dimension type date

Pourquoi creer une dimension Date plutot qu'utiliser directement le type DATE SQL?

- A) Par convention uniquement
- B) **Pour avoir des attributs pre-calculés (jour semaine, ferie, trimestre) et optimiser les jointures**
- C) Pour occuper plus d'espace
- D) Pour complexifier le modèle

3 Requetes MDX et Analyse (Q51-65)

Q51 - MDX definition

MDX signifie:

- A) Microsoft Data eXchange
- B) **MultiDimensional eXpressions**
- C) Multiple Database XML
- D) Managed Data eXtraction

Q52 - MDX vs SQL

La principale difference entre MDX et SQL est:

- A) MDX est plus ancien
- B) **MDX est concu pour interroger des cubes OLAP multidimensionnels**
- C) SQL est uniquement pour Oracle
- D) Il n'y a pas de difference

Q53 - Clause FROM MDX

Dans une requete MDX, la clause **FROM** specifie:

- A) La table relationnelle source
- B) **Le cube OLAP a interroger**
- C) Le fichier de sortie
- D) L'utilisateur connecte

Q54 - Clause SELECT MDX

Dans MDX, la clause **SELECT** definit:

- A) Les tables a joindre
- B) **Les axes de resultat (colonnes et lignes)**
- C) Les filtres de securite
- D) Les procedures stockees

Q55 - Clause WHERE MDX

Dans MDX, la clause **WHERE** (slicer) sert a:

- A) Joindre des tables
- B) **Filtrer le cube sur des dimensions non affichees (contexte de selection)**
- C) Trier les resultats
- D) Grouper les donnees

Q56 - Members MDX

La fonction **.Members** en MDX retourne:

- A) Les utilisateurs du systeme
- B) **Tous les membres d'une hierarchie ou d'un niveau**
- C) Les mesures disponibles
- D) Les cubes existants

Q57 - Children MDX

La fonction .Children en MDX retourne:

- A) Tous les membres de la hierarchie
- B) Les membres enfants directs d'un membre specifique**
- C) Le membre parent
- D) Les mesures

Q58 - CurrentMember MDX

La fonction .CurrentMember en MDX est utilisee pour:

- A) Supprimer un membre
- B) Faire reference au membre actuel dans un contexte d'iteration ou de calcul**
- C) Creer un nouveau membre
- D) Renommer un membre

Q59 - Calculated member

Un "calculated member" (membre calcule) en MDX est:

- A) Un membre stocke physiquement
- B) Une mesure ou membre cree par formule, calcule a la volee**
- C) Un membre supprime
- D) Un membre duplique

Q60 - WITH clause MDX

La clause WITH en MDX permet de:

- A) Specifier le serveur
- B) Definir des membres calcules ou des ensembles nommes pour la requete**
- C) Joindre des cubes
- D) Filtrer les erreurs

Q61 - CrossJoin MDX

La fonction CROSSJOIN en MDX produit:

- A) L'union de deux ensembles
- B) **Le produit cartesien de deux ensembles (toutes les combinaisons)**
- C) L'intersection de deux ensembles
- D) La difference de deux ensembles

Q62 - Filter MDX

La fonction FILTER en MDX permet de:

- A) Trier un ensemble
- B) **Selectionner les membres d'un ensemble qui satisfont une condition**
- C) Compter les membres
- D) Supprimer les doublons

Q63 - TopCount MDX

La fonction TOPCOUNT(Set, n, Measure) retourne:

- A) Les n derniers éléments
- B) **Les n premiers éléments triés par la mesure (ordre décroissant)**
- C) Un comptage total
- D) Les éléments au hasard

Q64 - YTD MDX

La fonction YTD (Year-To-Date) en MDX calcule:

- A) La date du jour
- B) **Le cumul depuis le début de l'année jusqu'au membre courant**
- C) L'année suivante
- D) La différence entre deux années

Q65 - ParallelPeriod MDX

La fonction ParallelPeriod en MDX permet de:

- A) Calculer en parallele
- B) Comparer avec la même période de l'année précédente
- C) Fusionner des périodes
- D) Supprimer des périodes

4 Approches et Methodologie (Q66-80)

Q66 - Inmon vs Kimball

La principale différence entre les approches Inmon et Kimball est:

- A) Inmon est plus récent
- B) Inmon préconise un DW normalisé top-down, Kimball des datamarts dimensionnels bottom-up
- C) Kimball n'utilise pas de dimensions
- D) Inmon n'utilise pas d'ETL

Q67 - EDW

Un EDW (Enterprise Data Warehouse) dans l'approche Inmon est:

- A) Un datamart unique
- B) Un entrepot de données centralisé, intégré et normalisé (3NF)
- C) Un cube OLAP
- D) Un outil ETL

Q68 - Datamarts Kimball

Dans l'approche Kimball, les datamarts sont:

- A) Indépendants sans lien entre eux
- B) Construits autour de processus métier et reliés par des dimensions conformes
- C) Uniquement pour les petites entreprises
- D) Temporaires et supprimés après usage

Q69 - Avantage Top-Down

L'avantage principal de l'approche Top-Down (Inmon) est:

- A) Rapide de mise en oeuvre
- B) **Coherence et integration des donnees a l'echelle de l'entreprise**
- C) Faible cout
- D) Simplicite technique

Q70 - Avantage Bottom-Up

L'avantage principal de l'approche Bottom-Up (Kimball) est:

- A) Meilleure normalisation
- B) **Livraison rapide de valeur metier avec des datamarts incrementaux**
- C) Moins de dimensions
- D) Pas besoin d'ETL

Q71 - Cycle de vie Kimball

Le cycle de vie d'un projet BI selon Kimball commence par:

- A) Le developpement technique
- B) **La planification du projet et la definition des besoins metier**
- C) L'installation des outils
- D) Le deploiement

Q72 - Trois trajectoires

Les trois trajectoires paralleles du cycle Kimball sont:

- A) Budget, Planning, Ressources
- B) **Technologie, Donnees, Applications BI**
- C) Analyse, Design, Test
- D) Extraction, Transformation, Chargement

Q73 - Besoins metier

La collecte des besoins metier est critique car:

- A) Elle determine le budget
- B) Elle guide toutes les decisions de conception (dimensions, mesures, priorites)**
- C) Elle est obligatoire legalement
- D) Elle est facile et rapide

Q74 - Sponsor projet

Le "sponsor" d'un projet BI est:

- A) Le developpeur principal
- B) Le dirigeant metier qui porte et finance le projet**
- C) L'administrateur base de donnees
- D) L'utilisateur final

Q75 - MVP BI

Un MVP (Minimum Viable Product) en BI consiste a:

- A) Livrer toutes les fonctionnalites d'un coup
- B) Livrer rapidement un premier datamart fonctionnel pour valider l'approche**
- C) Ne rien livrer avant 2 ans
- D) Livrer uniquement la documentation

Q76 - Agile BI

L'approche Agile en BI favorise:

- A) Un cahier des charges fige de 500 pages
- B) Des iterations courtes avec feedback continu des utilisateurs**
- C) Aucune planification
- D) Des projets de 5 ans minimum

Q77 - Sprint BI

Dans un projet BI Agile, un sprint dure typiquement:

- A) 1 jour
- B) **2 à 4 semaines**
- C) 6 mois
- D) 1 an

Q78 - User story BI

Une "user story" en BI pourrait être:

- A) "Installer Oracle 19c"
- B) **"En tant que directeur commercial, je veux voir le CA par région pour identifier les zones à développer"**
- C) "Créer 50 tables"
- D) "Acheter des serveurs"

Q79 - Prototype BI

Le prototypage rapide en BI permet de:

- A) Retarder le projet
- B) **Valider les besoins avec les utilisateurs avant le développement complet**
- C) Éviter toute documentation
- D) Remplacer les tests

Q80 - Conduite du changement

La conduite du changement dans un projet BI est importante car:

- A) Elle réduit le budget
- B) **Elle assure l'adoption de la solution par les utilisateurs**
- C) Elle accélère l'ETL
- D) Elle simplifie le code

5 Tendances et Technologies Modernes (Q81-100)

Q81 - Cloud BI

Un avantage majeur de la BI cloud est:

- A) Complexité accrue
- B) Scalabilité élastique et réduction des coûts d'infrastructure
- C) Dépendance à Internet uniquement
- D) Moins de fonctionnalités

Q82 - Data Lake vs DW

La différence principale entre Data Lake et Data Warehouse est:

- A) Le Data Lake est plus petit
- B) Le Data Lake stocke des données brutes (schema-on-read), le DW des données structurées (schema-on-write)
- C) Le DW est gratuit
- D) Il n'y a pas de différence

Q83 - Data Lakehouse

Un "Data Lakehouse" combine:

- A) Deux Data Lakes
- B) La flexibilité du Data Lake avec les capacités analytiques du Data Warehouse
- C) Deux bases relationnelles
- D) Un lac et une maison physiques

Q84 - Delta Lake

Delta Lake est une technologie pour:

- A) La visualisation de données
- B) Ajouter des transactions ACID et du versioning aux Data Lakes
- C) L'envoi d'e-mails
- D) La gestion de projet

Q85 - Real-time BI

La BI temps reel necessite:

- A) Un ETL batch quotidien
- B) Des pipelines de streaming (Kafka, Spark Streaming) et des tableaux de bord reactifs
- C) Uniquement des fichiers Excel
- D) Des rapports papier

Q86 - Apache Kafka

Apache Kafka est utilise pour:

- A) Creer des graphiques
- B) Le streaming de donnees en temps reel (messaging distribue)
- C) Editer des documents Word
- D) Gerer les ressources humaines

Q87 - Augmented Analytics

L'"Augmented Analytics" (analytique augmentee) utilise:

- A) Uniquement des requetes manuelles
- B) L'IA et le machine learning pour automatiser l'analyse et la decouverte d'insights
- C) Des lunettes de realite augmentee
- D) Des calculs papier

Q88 - NLQ

NLQ (Natural Language Query) permet de:

- A) Programmer en assembleur
- B) Interroger les donnees en langage naturel ("Quel est le CA de janvier?")
- C) Traduire des documents
- D) Compresser des fichiers

Q89 - Data Catalog

Un "Data Catalog" (catalogue de données) permet de:

- A) Stocker les données
- B) Inventorier, décrire et rendre découvrable le patrimoine de données**
- C) Supprimer les données anciennes
- D) Chiffrer toutes les tables

Q90 - Data Mesh

L'architecture "Data Mesh" préconise:

- A) Un Data Warehouse centralisé unique
- B) Une décentralisation avec des domaines métier propriétaires de leurs données**
- C) L'abandon de toute gouvernance
- D) Uniquement des fichiers CSV

Q91 - Data Fabric

"Data Fabric" est une approche qui:

- A) Utilise des tissus physiques
- B) Cree une couche d'intégration intelligente au-dessus de sources distribuées**
- C) Supprime toutes les données
- D) Ne fonctionne qu'avec Oracle

Q92 - ML in BI

L'intégration du Machine Learning dans la BI permet:

- A) De supprimer les tableaux de bord
- B) Des analyses predictives, de la détection d'anomalies et des recommandations automatiques**
- C) De remplacer complètement les analystes
- D) Uniquement du tri alphabétique

Q93 - Predictive Analytics

L'analytique predictive vise a:

- A) Decrire le passe uniquement
- B) **Anticiper les tendances futures a partir des donnees historiques**
- C) Supprimer les donnees anciennes
- D) Bloquer les utilisateurs

Q94 - Prescriptive Analytics

L'analytique prescriptive va au-delà de la predictive en:

- A) Decrivant mieux le passe
- B) **Recommandant des actions optimales a entreprendre**
- C) Supprimant les predictions erronees
- D) Affichant plus de graphiques

Q95 - DataOps

DataOps est une approche qui:

- A) Supprime les operations sur les donnees
- B) **Applique les principes DevOps (CI/CD, automatisation) aux pipelines de donnees**
- C) Ne concerne que les DBA
- D) Est uniquement theorique

Q96 - dbt

dbt (data build tool) est utilise pour:

- A) La visualisation de donnees
- B) **La transformation de donnees dans le warehouse (T de ELT) avec versioning et tests**
- C) L'extraction de donnees
- D) La gestion des utilisateurs

Q97 - Snowflake platform

Snowflake est principalement:

- A) Un outil de visualisation
- B) Un ETL on-premise
- C) Un Data Warehouse cloud avec separation calcul/stockage**
- D) Un réseau social

Q98 - Databricks

Databricks est basé sur:

- A) Microsoft Excel
- B) Apache Spark, avec un environnement unifié pour data engineering et data science**
- C) Oracle Database
- D) WordPress

Q99 - Embedded BI

L'"Embedded BI" (BI embarquée) consiste à:

- A) Installer la BI sur des téléphones
- B) Intégrer des analyses et tableaux de bord directement dans des applications métier**
- C) Enterrer les serveurs
- D) Cacher les rapports

Q100 - Reverse ETL

Le "Reverse ETL" consiste à:

- A) Annuler un ETL
- B) Pousser les données du Data Warehouse vers des outils opérationnels (CRM, marketing)**
- C) Inverser l'ordre des colonnes
- D) Supprimer l'ETL

Corrige - Grille de Reponses

Reponses (100 Questions)

Q1-20	Q21-40	Q46-65	Q66-85	Q86-100
1. C	21. B	46. B	66. B	86. B
2. B	22. C	47. B	67. B	87. B
3. B	23. B	48. B	68. B	88. B
4. B	24. B	49. B	69. B	89. B
5. B	25. B			
6. B	Q26-45	50. B	70. B	90. B
7. B	26. B	51. B	71. B	91. B
8. B	27. B	52. B	72. B	92. B
9. B	28. B	53. B	73. B	93. B
10. B	29. B	54. B	74. B	94. B
11. B	30. B	55. B	75. B	95. B
12. B	31. B	56. B	76. B	96. B
13. B	32. B	57. B	77. B	97. B
14. B	33. B	58. B	78. B	98. B
15. B	34. B	59. B	79. B	99. B
16. B	35. B	60. B	80. B	100. B
17. B	36. B			
18. B	37. B			
19. B	38. B			
20. B	39. B			
	40. B			

Recapitulatif par Theme

Section	Themes couverts	Questions
1. ETL Avance	CDC, extraction incrementale, qualite, deduplication, logging, orchestration	Q1-25
2. Modelisation	Dimensions conformes, SCD Types 1-6, faits transactionnels/periodiques, bridge tables	Q26-50
3. MDX	Syntaxe, fonctions (Members, Children, TopCount, YTD), membres calcules	Q51-65
4. Methodologie	Inmon vs Kimball, cycle de vie, Agile BI, user stories	Q66-80
5. Tendances	Cloud, Data Lake/Lakehouse, ML, DataOps, Snowflake, dbt, Reverse ETL	Q81-100