

Sujet examen entrepôts de données,

Le lundi 10 mai 2021 14h00 – 15h30 (durée 1h30 distanciel)
Documents autorisés.

Module : Entrepôts de Données - Master2 MIAGE
Note /30



Question 1 (1 point – 5 min) :

Qu'est ce qui caractérise les entrepôts de données quelque-soit le type :

- A. Volumétrie
- B. Volatile
- C. Historisation
- D. Typage des données
- E. Ordonnées les données
- F. Intégré
- G. Unifier
- H. Restituer
- I. Non Volatile.
- J. Archiver.

⇒ Répondez simplement en donnant la séquence de lettres qui vous semble le plus pertinent dans l'ordre alphabétique.

Question 2 (1 point – 5 min) :

Quelles sont les étapes pour construire un DataLake d'entreprise pour mon business ?
Ou autrement dit : Quelles questions se poser pour mettre en place un DataLake ?
Citez 7 éléments clés (juste les noms).

Question 3 (2 points – 5 min) :

Définir ce qu'est une donnée Structurée, une donnée non structurée et semi structurée.
Donnez un exemple pour chaque type.

Question 4 (2 points – 5 min) :

Donnez les éléments d'architectures d'un entrepôt de type DataWareHouse.
Citez-les et définir rapidement.

Question 5 (2 points – 5 min) :

Qu'elles sont les fonctions d'un ETL (définir rapidement celles-ci) dans un DataWareHouse ou un Datamart.

Question 6 (2 points – 5 min) :

Définir ce qu'est un Datalake, ses principales caractéristiques techniques.

Question 7 (2 points – 5 min) :

Quelles sont les différences majeures entre un DataWareHouse, un DataMart et un DataLake ?
Faire un petit tableau comparatif des différentes caractéristiques.

Question 8 (2 points – 5 min) :

Donnez le cycle de vie d'un projet en BI sous forme d'un Diagramme.

Question 9 (3 points – 5 min) :

Dans un DataWareHouse ou un DataMart quel modèles puis-je utiliser pour obtenir de bonnes de performances d'analyse avec les outils BI du marché.
Citez ces modèles, avec leur schéma de base.
Comment nomme t'on les différentes tables.

Question 10 (3 points – 5 min) :

Pour construire un DataLake quels sont les outils les plus communément utilisés pour mettre en place les connecteurs d'ingestions des données.
Quels sont les principes de ces connecteurs, schématiser les.

Question 11 (4 points – 15 min) :

On veut construire un entrepôt de données afin de stocker les informations sur les consultations d'un pays.
On veut notamment connaître le nombre de consultations, par rapport à différents critères (personnes, médecins, spécialités, etc.).

Ces informations sont stockées dans les relations suivantes :

PERSONNE (id, nom, tel, adresse, sexe)
MEDECIN (id, tel, adresse, spécialité)
CONSULTATION (id_med, id_pers, date, prix)

1. ***Proposer un schéma relationnel qui tienne compte de la date, du jour de la semaine, du mois, du trimestre et de l'année.***

2. *Quelle est la table des faits ? Quels sont les faits ?*
3. *Combien de dimensions ont été retenues ? Quelles sont-elles ?*
4. *Quelles sont les hiérarchies des dimensions ? Dessinez-les.*

Question 12 (6 points – 25 min) :

Le Centre Hospitalier Utopia, fort de 4000 lits et de plus de 8000 employés, produit beaucoup d'informations provenant de différents plateaux techniques et de soins.

Après un audit sur la production des données de cet établissement, il en ressort les éléments suivants :

- Radiologie, 183000 passages de patients par an et autant de compte rendus et plus de 482000 actes.
- Laboratoire, 1845230 analyses par an.
- Anapath, 452698 Comptes rendu écrits.
- Consultations 874956 comptes rendus.

Nb : Les comptes rendus sont des documents au format PDF.

Les données de laboratoire, radiologie, anapath et consultations sont également stockées dans des base de données séparées pour chaque métier.

De plus, un outil que l'on nomme la GAM gère l'ensemble des identités et des mouvements des patients dans une base de données séparées également.

1. *Proposez un entrepôt de données qui permettra de conserver ces données afin de réaliser des études prospectives avec des algos « d'IA » et des études rétrospectives avec des outils de fouilles de données et statistiques.*
2. *Faire un schéma d'architecture et donnez les noms des outils proposé sur l'ensemble de la chaine de ce Data Lake, de l'ingestion à la restitution des données.*

FIN DES QUESTIONS

Modalité d'examen à distance pour le Master2 MIAGE Entrepôts de données

Etape 1 :

Vous vous connecterez sur le service meet JitSi habituel (<https://meet.jit.si/d1d5a3fe-74c9-480d-96d1-728cd8927cd7>) selon l'horaire prévu le **lundi 10 mai 14h00 au plus tard**, [prenez un peu d'avance et connectez-vous 5 minutes avant 14H00].

⇒ Vous recevrez dans le chat jitsi le lien pour récupérer le document PDF comportant les questions de l'examen à 14h00 exactement, je resterai disponible sur ce chat pour toutes questions.

Etape 2 :

Au plus tard à 15h30 vous me retournerez un document contenant vos réponses aux formats acceptés

suivants :

- Microsoft Word (doc ou docx)
- LibreOffice Writer (odt)
- OpenOffice Writer (odt)
- Le format RTF est accepté également
- Le format PDF est accepté également
- Pages pour Mac (pages)
- Microsoft PowerPoint (ppt ou pptx)
- LibreOffice Impress (odp)
- OpenOffice Impress (odp)
- Keynote pour Mac (key)

⇒ Vous ferez le choix d'un de ces formats ci-dessus pour l'ensemble du partiel et mettrez toutes vos réponses dans un même document en indiquant à chaque fois le numéro de la question.

⇒ Vous indiquerez également en début de document votre nom complet.

⇒ Dans le cas où il y a des schéma à réaliser, mettre une capture d'écran dans le document de ce schéma correspondant à la question à laquelle vous répondez où illustrer.

Dans le cas où votre PC est HS ou non disponible, vous pouvez mettre vos réponses sur papier libre et me renvoyer par mail des photos de ces pages via votre Smartphone (bien cadré avec une résolution suffisante pour permettre la relecture sur machine, vérifiez donc que vos photos ne sont pas floues ni coupées et que le texte soit lisible.).

Ceci reste une solution de secours et n'est pas sollicité si vous pouvez réaliser l'examen sur votre PC.

Votre fichier ou vos photos sont à envoyer obligatoirement aux deux adresses suivantes :

tondeur.herve@yahoo.fr

herve.tondeur@hopsiia.fr (adresse de sécurité)

Vous mettrez dans le sujet du mail (objet) : [M2MIAGE_examen] : votre nom complet.

Votre ou vos pièces jointes à importer obligatoirement dans ce mail.

Dans le corps du mail "EXAMEN ENT".

⇒ ***Merci de contrôler que vous avez bien mis toutes vos pièces jointes.***

Il y aura une tolérance sur l'heure de réception de votre mail de 10 minutes maximum (temps de transport long et préparation de votre mail)

La date et heure de réception sur les boîtes mails faisant fois, tout éléments reçu après le lundi 10 mai 15h40, ne sera pas pris en compte.

⇒ Vous recevrez de ma part dès réception de votre mail un accusé de réception confirmant la réception et la complétude de celui-ci (veuillez surveiller la réception de ce mail en retour comme preuve).