

# UNIVERSITÉ DE THIES



UFR DES SCIENCES ECONOMIQUES ET  
SOCIALES

UFR DES SCIENCES ET  
TECHNOLOGIQUES

## **Master Science Des Données et Applications**

*Par*

*COUMBA SY*

*ALMAMY YOUSOUF LY*

*Sur le sujet*

---

## **Projet Business Intelligent**

---

Professeur : M. Boly

**Année universitaire 2019-2020**

### **Exercice 1 :**

3. Créer un rapport qui affiche pour chaque **Citystate** (concaténation des colonnes **CityName** et **StateProvinceName** séparées par une **virgule**), le montant total de ventes **Sum of SalesAmount**.

- **CityState** est la concaténation des colonnes **CityName** et **StateProvinceName** séparées par une virgule
- Le montant **Sum of SalesAmount** est une nouvelle mesure.

- $\text{Sum of SalesAmount} = \text{SUM}(\text{FactInternetSales}[\text{SalesAmount}])$

4. Créer une mesure **Number of Sales** qui donne le nombre total de ventes à partir de la table de faits **FactInternetSales**.

- $\text{Number of Sales} = \text{COUNTROWS}(\text{FactInternetSales})$

5. Créer une deuxième mesure **Number of Unique Products** qui donne le nombre distinct de produits à partir de la table de faits **FactInternetSales**

- $\text{Number of Unique Products} = \text{DISTINCTCOUNT}(\text{FactInternetSales}[\text{ProductKey}])$

6. Créer une mesure qui donne le nombre total de produits à date (depuis le début), **Number total To date**

- $\text{Number total To date} = \text{CALCULATE}(\text{COUNT}(\text{DimProduct}[\text{ProductKey}]), \text{ALL}(\text{DimProduct}[\text{StartDate}]))$

7. Créer une troisième mesure **Average Sales Count in Store** qui donne la moyenne du nombre de ventes (**Number of Sales**) à partir de la table **DimStore**

8. Créer un rapport qui affiche pour chaque continent les valeurs des mesures **Number of Sales**, **Number of Unique Products**, **Number total To date** et **Average Sales Count in Store**.

- Voir rapport

9. Créer une mesure **Count of All Stores** qui compte le nombre total de magasins

- $\text{Count of All Stores} = \text{COUNTROWS}(\text{DimStore})$

10. Créer une mesure **Count of Active Stores** qui compte le nombre total de magasins actifs.

- `Count of Active Stores = COUNTROWS(Filter(All(DimStore), DimStore[Status] = "On"))`

11. Créer un rapport qui affiche pour chaque continent les valeurs des mesures **Count of All Stores** et **Count of Active Stores**.

- Voir rapport

12. Créer une mesure **Sum of SalesAmount Prev Year** qui donne le montant total des ventes de l'année précédente

- `Sum of SalesAmount Prev Year = VAR annee = YEAR(LASTDATE(FactInternetSales[ShipDate])) - 1  
return CALCULATE( [Sum of SalesAmount], FactInternetSales[ShipDate].[Année] = annee )`

13. Créer une mesure **Sum of SalesAmount To Date** qui donne le montant total des ventes à date (depuis le début).

- `Sum of SalesAmount To Date = CALCULATE([Sum of SalesAmount], ALL(FactInternetSales[OrderDate]))`

14. Créer un rapport qui affiche pour chaque année les mesures **Sum of SalesAmount**, **Sum of SalesAmount PrevYear** et **Sum of SalesAmount To Date**

- Voir rapport

15. Créer une mesure qui correspond au cumul annuel des ventes sur les mois de la dimension Date

- `Cumul Annuel des Ventes = CALCULATE([Sum of SalesAmount], DATESYTD(DimDate[FullDateAlternateKey].[Date]))`

16. Créer une mesure qui correspond au cumul trimestriel des ventes sur les mois de la dimension Date

- Cumul Trimestriel des Ventes = CALCULATE ([Sum of SalesAmount], DATESQTD(DimDate[FullDateAlternateKey].[Date]))

17. Créer une mesure qui correspond au cumul semestriel des ventes sur les mois de la dimension Date

- Cumul Semestriel des Ventes =  
IF (  
    ISFILTERED(DimDate[EnglishMonthName]),  
    CALCULATE ([Sum of SalesAmount],  
        FILTER(ALL(DimDate),  
            DimDate[FullDateAlternateKey].[Date] <=  
            MAX(DimDate[FullDateAlternateKey].[Date])),  
    VALUES(DimDate[CalendarSemester])  
)  
)

## Exercice 2 :

1. Donner 2 raisons pour lesquelles les systèmes transactionnels ne sont pas adaptés à la gestion d'informations pour le décisionnel concernant le temps.
  - Primo, le traitement des données s'effectue au jour le jour dans le transactionnel alors que le décisionnel adopte une stratégie à long terme.
  - Secundo, nous avons les données temporelles réparties qui rendent la vue historique des données difficiles (volatile : pas d'historisation systématique) dans le système d'information transactionnel. Alors que dans le décisionnel les données sont non volatiles et historisées.
2. Considérant la caractéristique « Hétérogénéité des données » dans un système transactionnel, expliquer pourquoi cette caractéristique est un facteur bloquant en décisionnel ?
  - Dans le transactionnel, les données sont détaillées et mise à jour alors que dans le décisionnel les données sont résumées, agrégées, historisées, non sujette à des mises à jour.
3. Dans le chargement d'un Datawarehouse, dans quel type de schéma on gagne en termes de temps de chargement des données ? Justifier votre réponse
  - Le temps nécessaire pour le chargement dans un schéma en étoile est inférieur car il utilise moins de jointure. A l'inverse, le schéma en flocon consomme plus de temps en raison de la complexité du modèle due à une grande utilisation des jointures.
4. Pourquoi fait-on le choix de « normaliser » les données dans le transactionnel alors qu'en décisionnel on peut faire le choix de « dénormaliser » ?
  - Les systèmes d'informations décisionnels utilisent souvent des schémas dénormalisés ou partiellement dénormalisés (tels que le schéma en étoile) pour optimiser les performances des interrogations. A l'inverse, les systèmes transactionnels ont souvent recours à des schémas totalement normalisés pour optimiser les performances des opérations de mise à jour, d'insertion et de suppression, et pour garantir la cohérence des données.
5. Pourquoi dit-on que dans la zone de préparation, les données sont souvent détruites après chargement dans le Datawarehouse ?
  - C'est parce qu'il s'agit d'une zone temporaire de la base de données dans laquelle les données à charger dans le Datawarehouse sont prétraitées.

6. Pourquoi dit-on qu'une table de dimension contient en général moins d'enregistrements qu'une table de faits ?

- C'est parce que la table de faits contient des mesures sur les attributs des tables de dimensions.

Par exemple le fait « vente » peut être constitué d'activités : quantités vendues, montant des ventes....

7. Pourquoi dit-on que les tables de dimensions contiennent l'ensemble des informations descriptives des faits ?

- Par définition la table des faits modélise le sujet de l'analyse (ce que l'on souhaite mesurer) alors que dans la table de dimension, on parle d'axe d'analyse selon lequel vont être étudié les données observables c'est-à-dire les faits.

Par exemple le fait « vente » peut être analysé suivant différentes perspectives correspondant à trois dimensions : la dimension Temps, la dimension Géographie et la dimension Catégorie.

Il s'agit donc d'étudier les données observables c'est-à-dire le fait « vente » à travers les différents axes d'analyse (les dimensions).

Donc les tables de dimensions contiennent l'ensemble des informations descriptives des faits.