

Business Intelligence

Devoir ETL

Rapport

ING 4

Réalisé par:

Riad EL HAJJAME

Ali BENCHRIFA

Sami HIOUAS

Hajar BALIRH

Manal CHAILOU

Manal ED-DAOUDI

2022 – 2023

Contents

-	Elaboration de l'entrepôt de données :	3
1.	Conception :	3
2.	Réalisation :	3
A.	Table de faits :	3
B.	Table de dimension :	4
3.	Visualisation :	5
-	Alimentation de l'entrepôt de données :	6

- Elaboration de l'entrepôt de données :

1. Conception :

Les entreprises doivent être en mesure de prévoir les événements futurs pour faire face aux défis économiques. Pour cela, elles peuvent utiliser des données pertinentes qui sont disponibles en interne ou auprès de fournisseurs externes. Cependant, ces données sont souvent dispersées et non organisées, il est donc important de les rassembler et de les standardiser pour permettre des analyses et faciliter les prises de décisions. Pour répondre à ces besoins, l'entreprise peut mettre en place un système de prise de décision qui lui permettra de prendre des décisions et de prévoir les événements en utilisant les informations disponibles. Les systèmes d'information ont pour objectif de conserver une trace fiable et intégrée des événements, ce qui permettra à l'entreprise de devenir plus entrepreneuriale et d'avoir une meilleure connaissance de ses clients, de sa compétitivité et de son environnement.

2. Réalisation :

Entrepôt de données est constitué de :

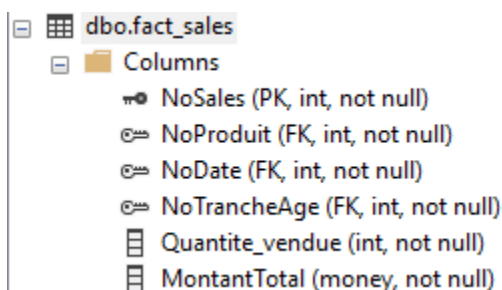
- Table de faits
- Table de dimension

A. Table de faits :

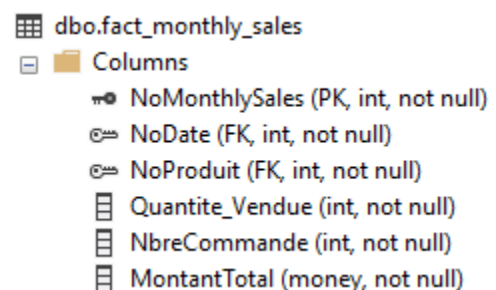
Une table de fait est un outil utilisé pour regrouper et organiser les données observées sur un sujet donné. Elle est liée à des tables dimensionnelles qui fournissent des informations supplémentaires sur les données observées. Les tables de fait contiennent des associations de données qui dépendent des données des tables dimensionnelles, et permettent une analyse plus approfondie et une meilleure compréhension des données.

Dans notre cas, nous avons utilisés deux tables de faits :

- Fact_sales
- Fact_monthly_sales



dbo.fact_sales	
Columns	
NoSales	(PK, int, not null)
NoProduit	(FK, int, not null)
NoDate	(FK, int, not null)
NoTrancheAge	(FK, int, not null)
Quantite_vendue	(int, not null)
MontantTotal	(money, not null)



dbo.fact_monthly_sales	
Columns	
NoMonthlySales	(PK, int, not null)
NoDate	(FK, int, not null)
NoProduit	(FK, int, not null)
Quantite_Vendue	(int, not null)
NbreCommande	(int, not null)
MontantTotal	(money, not null)

B. Table de dimension :

Les tables de dimensions permettent de classer les données en utilisant des attributs spécifiques tels que les catégories, les dates, les lieux, etc. Elles sont souvent utilisées pour créer des segments de données et des groupes de données, qui peuvent ensuite être utilisés pour créer des tableaux de bord, des rapports et des indicateurs clés de performance (KPI). Les tables de dimensions sont liées aux tables de faits et permettent de créer des hiérarchies de données, ce qui permet de décomposer les données à différents niveaux de détails, facilitant ainsi l'analyse et la compréhension des données.

Dans notre cas, nous avons utilisés trois tables de dimensions :

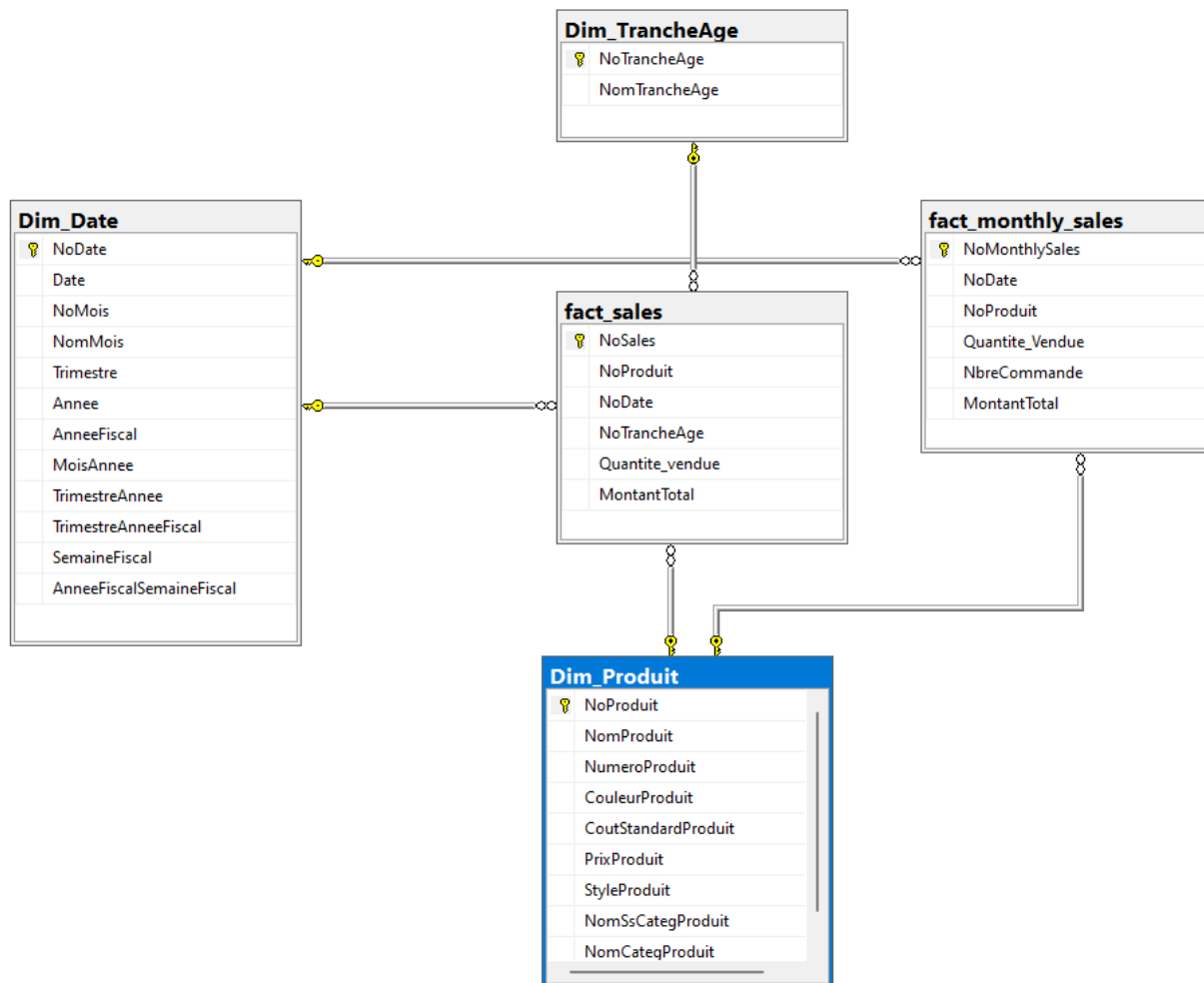
- Dim_Date
- Dim_TrancheAge
- Dim_Produit

dbo.Dim_Date
Columns
NoDate (PK, int, not null)
Date (datetime, not null)
NoMois (smallint, not null)
NomMois (varchar(10), not null)
Trimestre (varchar(10), not null)
Annee (smallint, not null)
AnneeFiscal (smallint, not null)
MoisAnnee (varchar(10), not null)
TrimestreAnnee (varchar(10), not null)
TrimestreAnneeFiscal (varchar(10), not null)
SemaineFiscal (varchar(10), not null)
AnneeFiscalSemaineFiscal (varchar(10), not null)

dbo.Dim_Produit
Columns
NoProduit (PK, int, not null)
NomProduit (nvarchar(50), not null)
NumeroProduit (nvarchar(25), not null)
CouleurProduit (nvarchar(10), not null)
CoutStandardProduit (money, not null)
PrixProduit (money, not null)
StyleProduit (nvarchar(25), not null)
NomSsCategProduit (nvarchar(50), not null)
NomCategProduit (nvarchar(50), not null)

dbo.Dim_TrancheAge
Columns
NoTrancheAge (PK, int, not null)
NomTrancheAge (nvarchar(20), not null)

3. Visualisation :



Le diagramme réalisé est un diagramme de constellation : Un diagramme de constellation est un type de diagramme utilisé pour représenter les relations entre les tables d'une base de données. Il permet de visualiser les relations de clé étrangère entre les tables, de manière à comprendre comment les données sont liées et comment elles sont utilisées dans les différentes requêtes. Il se compose généralement de formes rectangulaires pour représenter les tables, et de flèches pour représenter les relations de clé étrangère entre les tables. Les diagrammes de constellation sont généralement utilisés pour documenter les modèles de données, pour planifier des modifications à une base de données existante et pour faciliter la compréhension des relations entre les tables pour les utilisateurs et les développeurs.

- Alimentation de l'entrepôt de données :

Les données sont extraites d'une source de données, puis injectées vers SQL Server, éventuellement après une transformation des données. Les tables de fait fact_sales et fact_monthly_sales sont alimentées à partir de la base de données relationnelle d'AdventureWorks2014. Pour les tables dimension, les données sont extraites d'AdventureWorks2014, puis intégrées dans les tables des dimensions Tranche_age, Produit, et Date.

Le code envoyé en pièce jointe dans le mail s'agit de scripts SQL qui créent des tables de dimensions et de faits pour une base de données de ventes. Les tables de dimensions incluent Dim_TrancheAge, Dim_Date, et Dim_Produit, qui stockent des informations sur les tranches d'âge des clients, les dates des ventes et les produits vendus. Les tables de faits incluent fact_monthly_sales et fact_sales, qui stockent des informations sur les ventes mensuelles et les ventes totales. Les scripts incluent également des instructions pour remplir ces tables avec des données provenant d'une base de données AdventureWorks2014, en utilisant des jointures et des requêtes pour extraire les informations nécessaires.