

I. Extraction des données et Partie ETL – Power Query : (fin mi formation)

A chaque extraction :

- Vous modifier le nom de la feuille (si feuille1, etc.) afin qu'il soit plus parlant,
- Puis cliquer sur la petite flèche au niveau de **FERMER ET CHARGER** → Sélectionner **FERMET ET CHARGER DANS**
- Sélection « Connexion uniquement » afin qu'il ne crée pas une feuille **EXCEL** avec la table contenant les données de la source **PUIS** cocher « Ajouter ces données au modèle de données » afin que les données soient chargées directement dans le **CUBE (Power Pivot)**.

Pour mes tableaux de bords, j'ai besoin de modéliser un schéma qui comportera ces différentes sources de données :

1. « **Ventes et géographie** » → Document Excel contenant 5 feuilles à importer : « Géographie », « Géographie2 », « VentesAVANT2023 », « VentesDEPUIS2023 » et « DATE_Calendrier ».
2. « **Produits** » → document Excel contenant 1 feuille MAIS il faut importer le tableau intitulé Produit ».
3. « **Clients** » → fichier texte à importer,
4. « **Vendeur** » : Connexion Web → Aller chercher le tableau sur la page du site internet suivant : <http://berenguer-formation-conseil.fr/tableau-vendeur-pbi/>
Le tableau sera chargé sous le nom de table1. Il faut le renommer en « Vendeur ».

II. Nettoyage et préparation des données : création de variables :

Lors de la première importation de vos données en entreprise, il faudra impérativement vérifier que :

- Les types de données attribués par Power Query ne sont pas forcément bons car dépendent des 200 premières lignes,
- Que toutes les colonnes sont nécessaires ? sinon, les supprimer !
- Que toutes les lignes sont propres ? sinon, les nettoyer ou les supprimer si besoin de filtres !

Il faut se questionner sur comment optimiser le jeu de données pour son utilisation dans les visuels.

1. Renommer les colonnes « Nom » dans les tables « client », « vendeur » avec la bonne valeur : « nom du client », « nom du vendeur ».
2. Supprimer la colonne « StockTypeId » sur la table *Produit* mais en gardant l'assistant de suppression dans les étapes de modification (Choisir les colonnes).
3. Je souhaite remplacer les valeurs vides (dû au CSV) par des valeurs Null reconnus par Power BI dans les colonnes « genre » et « situation familiale » de la table *client* : Je sélectionne avec CTRL les deux colonnes puis je fais un clic droit « Remplacer les valeurs » :

Remplacer les valeurs

Remplacez une valeur dans les colonnes sélectionnées par une autre.

Valeur à rechercher

Remplacer par

Options avancées

Je remplace : « rien » par « null » en minuscule 😊

4. Fractionner la colonne « Pays et ville » dans la table *géographie2*,
Attention, lors du fractionnement, il faut utiliser un fractionnement personnalisé afin d'utiliser la virgule et les espaces avant et après elle comme séparateur complet. (Sinon le pays va être écrit avec un espace après et la ville avec espace avant).

Puis, renommer les colonnes proprement

5. Dans la table *client*, supprimer les lettres qui apparaissent dans la colonne idclient (fractionnement spécifique) et renommer la colonne idclient proprement (sans le 1).

6. Supprimer les doublons sur la colonne « idvendeur » de la *table vendeur* afin de n'avoir qu'une seule ligne par vendeur différente.
7. Filtrer les lignes sur la *table* « *VentesAVANT2023* » : sur la colonne « Date de ventes », il faut seulement les ventes à partir de 2019.

SUITE POWER QUERY

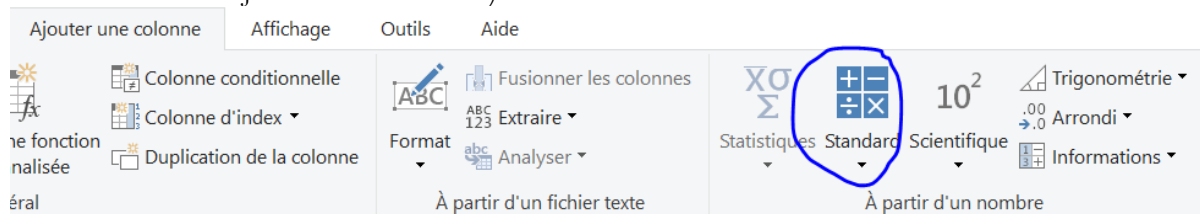
8. Concaténer les lignes de la *table VentesAVANT2023* avec celles de la *table VentesDEPUIS2023* : en utilisant « Ajouter des requêtes ». Ce qui permettra de récupérer les lignes de *VentesDEPUIS2023* dans la *table VentesAVANT2023*. → Il faut se positionner sur la table où l'on souhaite ajouter des lignes, donc bien cliquer sur la *table VentesAVANT2023* puis utiliser « Ajouter des requêtes ».

Renommer ensuite la *table VentesAVANT2023* : *Ventes* !

9. Fusionner la *table géographie* avec la *table géographie2* afin de pouvoir récupérer toutes les colonnes de la *table géographie2* dans la *table géographie* : en utilisant « Fusionner des requêtes » → Pas besoin de récupérer le champ idgeo de la *table géographie2* (car existe déjà dans la *table géographie*).
10. Créer groupes de requêtes : créer groupe « intermédiaire » avec les requêtes *géographie2* et *VentesDEPUIS2023*. Cliquer droit sur la requête : mettre dans autre groupe – créer.

SUITE POWER QUERY

11. Sur la *table Ventes* : Créer une colonne qui nous permette de calculer les quantités réellement vendues : il faut donc soustraire les quantités retournées aux quantités vendues (Onglet du ruban « Ajouter une colonne ») :



12. Sur la *table Ventes*. Créer une colonne personnalisée qui nous permette de calculer la part de bénéfices par ventes. On a besoin donc de soustraire le montant des ventes par celui des coûts totaux avant de diviser par le montant des ventes global.
→ Penser à renommer la colonne, renommer l'étape appliquée à droite et modifier le type de données de votre nouvelle colonne.
$$=([MontantVentes]-[CoûtTotal])/[MontantVentes]$$

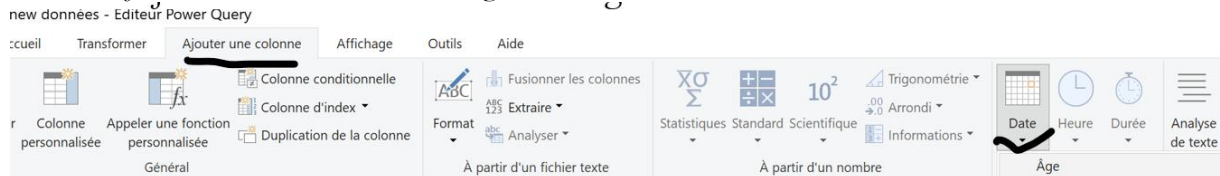
13. Sur la *table Produit* : Dans l'onglet, « ajouter une colonne », fusionner les colonnes Idproduit et nom de produit en les séparant par un point-virgule.
14. Sur la *table clients* : Créer une nouvelle colonne Genre (à partir de la colonne genre 0, avec Femme à la place de F et Homme à la place de H) en utilisant une colonne conditionnelle.
→ Penser à renommer la colonne, renommer l'étape appliquée à droite et modifier le type de données de votre nouvelle colonne.
15. Sur la *table clients* : Créer une nouvelle colonne Situation familiale (à partir de la colonne situation familiale0, avec Célibataire à la place de C, Mariée à la place de M et Pacsée à la place de P) en utilisant une colonne conditionnelle.

→ Penser à renommer la colonne, renommer l'étape appliquée à droite et modifier le type de données de votre nouvelle colonne.

16. Concaténer les deux nouvelles colonnes Genre et situation familiale ensemble dans une nouvelle colonne en fusionnant les colonnes avec un espace en guise de séparateur !

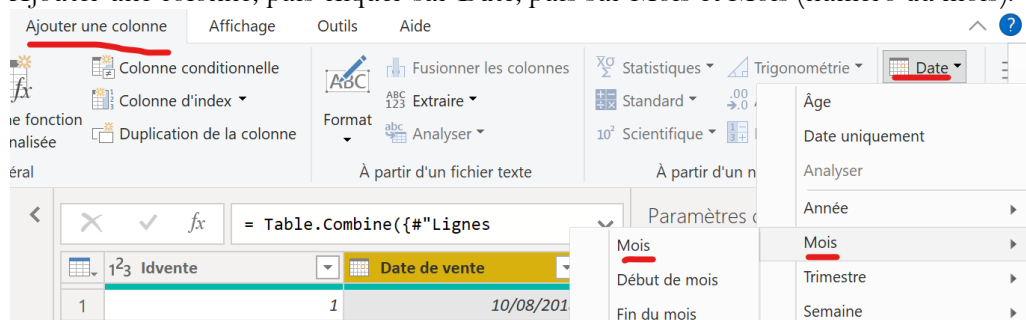
17. Création de l'âge du client en 3 étapes (une seule création de colonne puis transformations dessus) :

- Ajouter une colonne : Date = Age



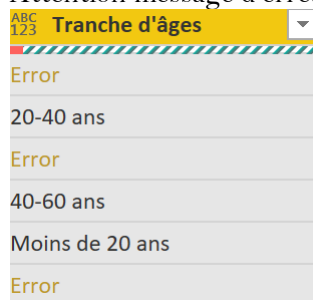
- Transformer : Durée – Nombre d'années
- Transformer : Arrondir à l'entier inférieur

18. Création du numéro du mois dans la table *Ventes* en utilisant le champ *Date de vente*. Ajouter une colonne, puis cliquer sur Date, puis sur Mois et Mois (numéro du mois).



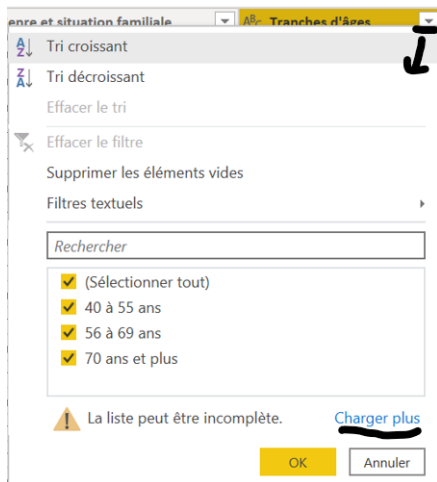
19. Création des tranches d'âges des **CLIENTS** = Moins de 20 ans, 20-39 ans (moins de 40), 40 à 59 ans, 60 ans et plus.

Attention message d'erreur possible ou erreurs qui s'affichent d'office dans les données.



Depuis version août 2022 : cliquer sur la cellule à droite du mot error pour afficher le message d'erreur.

Sur les anciennes versions, utiliser la petite flèche blanche à droite de l'entête de colonne créée et faire « Charger plus ». Les erreurs et l'explication s'affichera à gauche au-dessus du bouton OK.



! Expression.Error : Désolé... Nous n'avons pas pu convertir la valeur null en type Logical.
 Détails :
 Value=
 Type=[Type]

Il est nécessaire de modifier l'erreur.

2 possibilités :

- Remplacer erreurs (en cliquant droit sur la colonne en question puis dire par quelle valeur remplacer toutes vos erreurs),
- Retourner sur l'étape appliquée précédente (colonne conditionnelle) et rajouter la notion des valeurs Null dans la colonne conditionnelle en tout premier comme ce qui suit :

1 Cliquer sur ajouter une clause

2 Cliquer tout à droite, sur les 3 petits points de la ligne « Plus » et « Déplacer vers le haut » afin de faire remonter cette nouvelle clause tout en haut

3 Remplir condition : Age est égal à null et là il les gère bien

Nouveau nom de colonne

Tranche d'âges

	Nom de la colonne	Opérateur	Valeur ①		Sortie ①
Si	Âge	égal à	ABC 123 null	Alors	ABC 123 NC
Autre...	Âge	est inférieur à	ABC 123 20	Alors	ABC 123 Moins de 20 ans

20. Fermer et charger

Fermer Power Query et de retour sur Excel, aller modifier le type d'importation des données pour les tables « Ventesdepuis2023 » et « geography2 ». Décocher « ajouter ces données au modèle de données » car on a déjà les lignes et les colonnes de ces tables dans d'autres tables 😊

21. Seulement pour les plus avancés et ayant déjà vu du code SQL ou autre :

Tenter de créer directement une nouvelle colonne (via colonne personnalisée) qui aura les valeurs « Homme marié », 'Femme Mariée' etc. mais sans créer/utiliser les colonnes conditionnelles en amont...

Juste du code → si H pour la colonne Genre et si C pour la colonne Situation familiale alors Homme célibataire....

La syntaxe : Si dans la colonne champ1, la valeur est A1 ET dans la colonne champ2 la valeur est B ALORS je dois créer la valeur « valeur AB » dans ma nouvelle colonne

if [champ1]="valA" and [champ2]="B" then "valeur AB"

Else if [champ1]="valB" and [champ2]="B" then "valeur BB"

Else if etc.

Else "blabla"

```
=if [Genre0] = "H" and [Situation familiale0] = "M" then "Homme Marié"
else if [Genre0] = "H" and [Situation familiale0] = "C" then "Homme célibataire"
else if [Genre0] = "H" and [Situation familiale0] = "P" then "Homme pacsé"
else if [Genre0] = "F" and [Situation familiale0] = "M" then "Femme mariée"
else if [Genre0] = "F" and [Situation familiale0] = "C" then "Femme célibataire"
else if [Genre0] = "F" and [Situation familiale0] = "P" then "Femme pacsée"
else "NC"
```

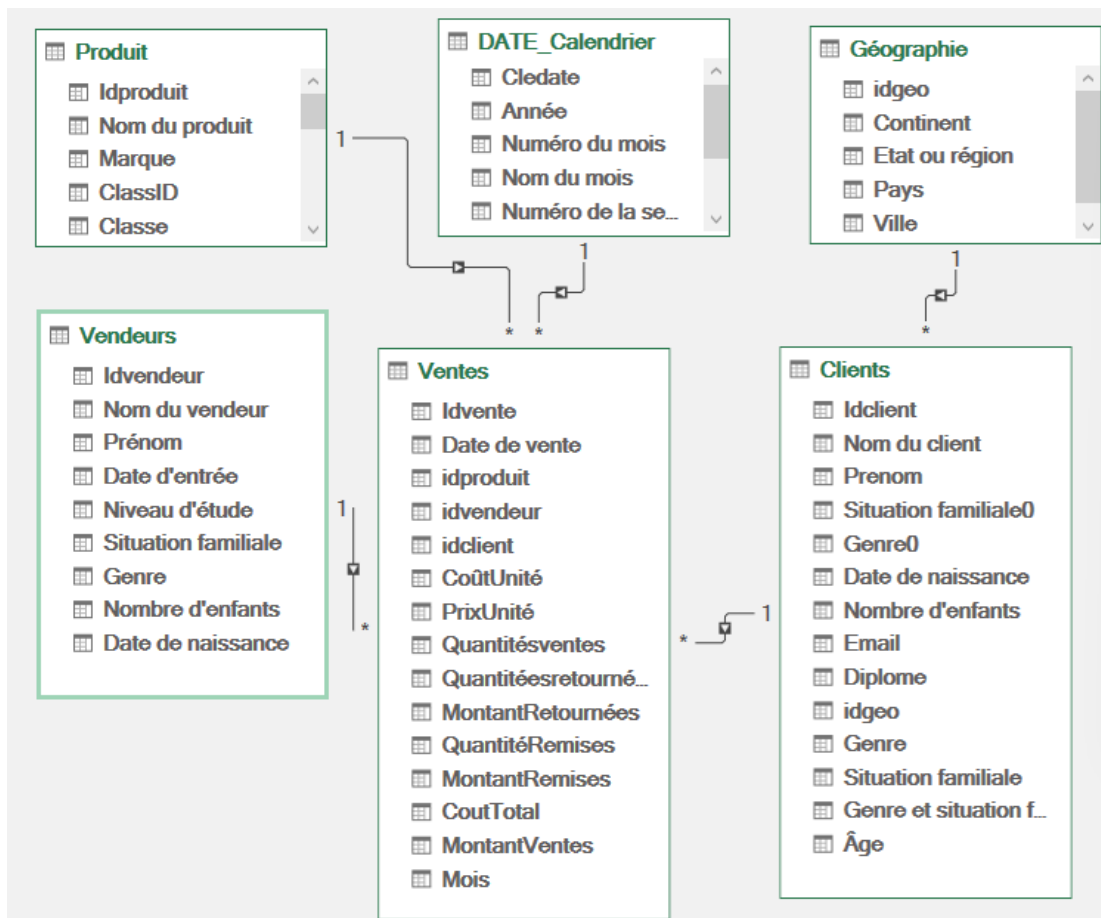
III. Power Pivot – Création du modèle : fin seconde AM

a. Vue diagramme :

Créer les relations suivantes : (onglet en haut «conception» puis « gérer les relations »)

Gérer les relations				
<div>Créer Modifier Supprimer</div>				
Active	Table 1	Cardinalité	Direction du filtre	Table 2
Oui	Clients [idgeo]	Plusieurs à un (*:1)	<< Vers Clients	Géographie [idgeo]
Oui	Ventes [Date de vente]	Plusieurs à un (*:1)	<< Vers Ventes	DATE_Calendrier [Cledate]
Oui	Ventes [idclient]	Plusieurs à un (*:1)	<< Vers Ventes	Clients [Idclient]
Oui	Ventes [idproduit]	Plusieurs à un (*:1)	<< Vers Ventes	Produit [Idproduit]
Oui	Ventes [idvendeur]	Plusieurs à un (*:1)	<< Vers Ventes	Vendeurs [Idvendeur]

Votre schéma final devra être celui-là :



b. Optimisation du CUBE

1. Définir table des dates ou créer tables de tables,
2. Trier le mois (calendarmonthlabel) en fonction du numéro de mois.
3. Trier aussi les tranches d'âges par le numéro de tranche d'âge créé dans Power Query
4. Créer la hiérarchie de temporalité dans la table date = année, trimestre, mois.
5. Créer la hiérarchie dans la table géographie = continent, pays, état, ville.
6. Créer la hiérarchie dans la table produit = catégorie, sous-catégorie, nom produit.
7. Regarder ensuite ce que cela produit sur un tableau croisé dynamique si vous y insérez une hiérarchie ?

IV. Power Pivot – Langage DAX – 3^{ème} jour (mieux de commencer plus tôt)

a. Modifications simples

1. *Sur la table Factsales (si chargée via Power Pivot),*
 - *Utiliser la fonction LEFT(colonne ; 10) afin de récupérer que la date de la colonne DATEKEY.*
 - *Ensuite, changer le type de données en DATE.*
 - *Pour finir, refaire la relation avec la nouvelle colonne DATEKEY au bon format vers la colonne DateKey de la table Date_calendrier.*
 - *Que serait-il mieux de faire ? charger directement sur Power Query malgré les millions de lignes car n'a pas les limitations d'Excel.*

b. Partie création de colonnes calculées

1. Ajout d'une colonne calculée dans la table Fact Sales : Faire la différence entre Sales Amount et TotalCost et renommer la nouvelle colonne « Marge »,
2. Voir le SUMX : plus performant pour faire le PROFIT

c. Créations de mesures et KPI (count, distinct count, nb couleurs, produits achetés)

3. Création d'une mesure pour connaître le montant total des ventes (comme sur Excel mais en anglais donc SUM)
4. Création d'une mesure SUMX qui va calculer la différence entre les montants des ventes au total et les coûts (Permet d'éviter de créer une colonne sur une table centrale / d'indicateurs / de transactions contenant de grosses volumétries)
5. Création d'une mesure qui me donne les montants des ventes seulement aux Etats-Unis
6. Création d'une mesure qui me donne les montants des ventes seulement quand le prix unitaire (table Sales) est supérieur à 10 et dont le canal de vente est Store,
7. Création d'une mesure pour avoir la part du montant des ventes pour les Etats Unis dans les ventes au total puis changer le format.
8. Création d'une mesure qui va calculer le profit de l'année précédente. Il faut mettre la mesure dans un tableau avec les années pour voir comment elle fonctionne.
9. Rajout d'un KPI (sur le total des ventes) : je souhaite que mes vendeurs fassent 10% de plus sur les ventes cette année par rapport à l'année dernière.... Juste une multiplication car j'ai déjà le profit de l'année dernière ([Profit N-1]*1,1

V. Création d'une visite guidée sur Power Map : 3D MAPS

1. Type de carte et montrer que l'on peut bouger la carte sur les côtes mais aussi l'aplatir,

2. Création d'une première scène avec un premier type de graphique (symbole carré avec des ronds à l'intérieur : visualisation sur bulles). Mettre Pays puis ville et cliquer sur ville, taille : montant des ventes, catégorie : canal de vente
3. Ajout d'annotations et de commentaires,
4. Voir les paramétrages de la scène : durée de la transition, type de transition, etc.
5. Création d'une seconde scène avec un calque en plus et un type de graphique (symbole carré : visualisation sur région) avec pays en localisation et total des coûts en valeur.
6. Création d'une seconde scène qui copie la précédente mais on change de continent,
7. Enplacement Cityname et graphique histogramme empilé et heure avec datekey

VI. Créations de rapports : Power BI

a. Importer les tables récupérées et refaire les relations – en groupe

Pour mes tableaux de bords, j'ai besoin de modéliser un schéma qui comportera ces différentes sources de données :

1. « **Ventes et géographie** » → Document Excel contenant 5 feuilles à importer : « Géographie », « Géographie2 », « VentesAVANT2023 », « VentesDEPUIS2023 » et « DATE_Calendrier ».
2. « **Produits** » → document Excel contenant 1 feuille MAIS il faut importer le tableau intitulé Produit ».
3. « **Clients** » → fichier texte à importer,
4. « **Vendeur** » : Connexion Web → Aller chercher le tableau sur la page du site internet suivant : <http://berenguer-formation-conseil.fr/tableau-vendeur-pbi/>
Le tableau sera chargé sous le nom de table1. Il faut le renommer en « Vendeur ».

b. Aller dans Power Query :

1. Renommer les colonnes « Nom » dans les *tables* « client », « vendeur » avec la bonne valeur : « nom du client », « nom du vendeur ».
2. Supprimer la colonne « Pays Test » sur la *table* **GEOGRAPHIE** mais en gardant l'assistant de suppression dans les étapes de modification (Choisir les colonnes).
3. Je souhaite remplacer les valeurs vides (dû au CSV) par des valeurs Null reconnus par Power BI dans les colonnes « genre » et « situation familiale » de la *table client*.

Je remplace : « rien » par « null » en minuscule 😊

4. Fractionner la colonne « Pays et ville » dans la *table géographie2*,
5. Dans la *table client*, supprimer les lettres qui apparaissent dans la colonne idclient (fractionnement spécifique) et renommer la colonne idclient proprement (sans le 1).
6. Supprimer les doublons sur la colonne « idvendeur » de la *table vendeur*
7. Filtrer les lignes sur la *table* « **VentesAVANT2023** » : sur la colonne « Date de ventes », il faut seulement les ventes à partir de 2019.
8. Concaténer les lignes de la *table VentesAVANT2023* avec celles de la *table VentesDEPUIS2023* : en utilisant « Ajouter des requêtes » puis renommez-là « Ventes ».
9. Fusionner la *table geographie* avec la *table geographie2* afin de pouvoir récupérer toutes les colonnes de la *table geographie2* dans la *table géographie* : en utilisant « Fusionner des requêtes »
10. Créer groupes de requêtes : créer groupe « intermédiaire » avec les requêtes *geographie2* et *VentesDEPUIS2023*. Cliquer droit sur la requête : mettre dans autre groupe – créer.
11. Cliquez-droit sur les requêtes dans le groupe intermédiaire et cliquer sur « Activer le chargement »

12. Sur la *table Ventes*. Créer une colonne personnalisée qui nous permette de calculer la part de bénéfices par ventes. On a besoin donc de soustraire le montant des ventes par celui des coûts totaux avant de diviser par le montant des ventes global.

→ Penser à renommer la colonne, renommer l'étape appliquée à droite et modifier le type de données de votre nouvelle colonne.

$$=([\text{MontantVentes}] - [\text{CoûtTotal}]) / [\text{MontantVentes}]$$

13. Sur la *table Produit* : Dans l'onglet, « ajouter une colonne », fusionner les colonnes Idproduit et nom de produit en les séparant par un point-virgule.
14. Sur la *table clients* : Créer une nouvelle colonne Genre (à partir de la colonne genre 0, avec Femme à la place de F et Homme à la place de H) en utilisant une colonne conditionnelle.
15. Sur la *table clients* : Créer une nouvelle colonne Situation familiale (à partir de la colonne situation familiale0, avec Célibataire à la place de C, Mariée à la place de M et Pacsée à la place de P) en utilisant une colonne conditionnelle.
16. Concaténer les deux nouvelles colonnes Genre et situation familiale ensemble dans une nouvelle colonne en fusionnant les colonnes avec un espace en guise de séparateur !
17. Création de l'âge du client en 3 étapes
 - Ajouter une colonne : Date = Age
 - Transformer : Durée – Nombre d'années
 - Transformer : Arrondir à l'entier inférieur
18. Création du numéro du mois dans la *table Ventes* en utilisant le champ *Date de vente*
19. Création des tranches d'âges des **CLIENTS** = Moins de 20 ans, 20-39 ans (moins de 40), 40 à 59 ans, 60 ans et plus.

c. Fenêtre Power Bi (onglets correspondants à power pivot)

1. Définir table des dates ou créer tables de tables,
2. Trier le mois (calendarmonthlabel) en fonction du numéro de mois.
3. Trier aussi les tranches d'âges par le numéro de tranche d'âge créé dans Power Query
4. Créer la hiérarchie de temporalité dans la table date = année, trimestre, mois.
5. Créer la hiérarchie dans la table géographie = continent, pays, état, ville.
6. Créer la hiérarchie dans la table produit = catégorie, sous-catégorie, nom produit.
7. Regarder ensuite ce que cela produit sur un tableau croisé dynamique si vous y insérez une hiérarchie ?
8. Création de la mesure Montant des Ventes
9. Création de la mesure Montant des ventes USA
10. Création du Profit
11. Création du Profit de l'année dernière

d. Création d'un rapport ensemble

1. Création d'un segment date
2. Création d'un segment classe
3. Création d'un graphique avec les montants des ventes par BrandName, seulement pour les BrandName supérieur à 1 milliard
4. Création d'un graphique histogramme et courbe avec en histogramme les montant des ventes par catégories et en courbe, la part des ventes aux USA. Filtre sur les 5 meilleures catégories en fonction de leur montant des ventes.