

A. Objectif

La deuxième partie de cet ouvrage propose dans un premier temps de découvrir et de se familiariser avec l'élaboration d'un modèle de données grâce à Power Pivot dans Excel.

Puis, un cas pratique, la réalisation d'un tableau de bord inspiré d'un projet professionnel, sera proposé afin d'appliquer de manière concrète les éléments abordés.

B. Pourquoi créer un modèle de données dans Excel

Power Pivot était dans sa première version un simple complément d'Excel 2010. Bien que cette version offrait moins de possibilités qu'aujourd'hui, pour la première fois, les « power user » ont pu utiliser Excel de la même manière qu'une base de données relationnelle ; c'est-à-dire créer des relations de tables à l'intérieur d'un fichier Excel sans avoir à utiliser un nombre important de fonctions de recherche type RECHERCHEV ou INDEX/EQUIV.

Les anglo-saxons ont un terme assez explicite pour désigner un fichier Excel qui tente de répliquer les fonctionnalités d'une base de données relationnelles à partir de formules : EXCEL HELL

Malheureusement, durant mon expérience professionnelle, je me suis retrouvé dans cette situation à de nombreuses reprises. À chaque fois, au fur et à mesure que le fichier se développe, il devient impossible à maintenir, lent et à vrai dire, au bout d'un moment, plus personne ne comprend à quoi servent les centaines de formules complexes peuplant chacun des onglets.

La faculté de créer un modèle de données afin de fusionner des sources de données disparates contenant des centaines de milliers de lignes dans un moteur analytique aussi évolué et accessible qu'Excel était révolutionnaire.

Avec la sortie d'Excel 2016, Microsoft a choisi d'incorporer Power Pivot directement dans le ruban d'Excel.

Cependant, contrairement aux concepts traditionnels d'Excel, où l'approche de développement de solutions est relativement intuitive, vous devez avoir une compréhension de base de la terminologie et de l'architecture de base de données afin de tirer le meilleur parti de Power Pivot.

C. Les principes fondamentaux d'un modèle de données

Le modèle de données permet d'organiser les données à la manière d'une base de données relationnelle directement dans Excel. Il s'agit d'une composante de l'outil Power Pivot d'Excel.

Il sera ainsi possible :

- ▶ de gérer et analyser un ensemble de données volumineux qui ne pourrait pas être contenu dans une feuille de calcul Excel traditionnelle ;
- ▶ de créer des relations de tables afin d'afficher et d'agréger les données à la demande ;
- ▶ de créer des tableaux croisés dynamiques non pas à partir d'une table unique mais à partir d'un ensemble de tables organisées et reliées entre elles.

1. La normalisation

D'une manière générale, la normalisation consiste à organiser les tables et les colonnes dans un modèle de données structuré afin de réduire les redondances et de préserver l'intégrité des données.

Les objectifs de la normalisation sont :

- ▶ d'éliminer les données redondantes pour réduire la taille des tables et améliorer la vitesse et l'efficacité du traitement ;
- ▶ de minimiser les erreurs et les anomalies dues aux modifications de données (insertion, mise à niveau ou suppression d'enregistrement) ;
- ▶ de simplifier la mise en place de requêtes et de structurer la base de données pour une analyse significative.

Dans un modèle de données normalisé, chaque table doit avoir un objectif distinct et spécifique (informations sur les clients ou les fournisseurs, enregistrements d'une transaction, etc.).

Exemple

Vous retrouverez les données de cet exemple dans le fichier `modèle_données.xlsx`, la résolution de cet exemple se trouve dans le fichier `modèle_données_résolu.xlsx`.

Dans l'onglet **table**, le tableau suivant retrace les emprunts de livres de lecteurs d'une bibliothèque :

N° opération	date	nom	prénom	adresse	Livre	auteur	prix d'achat	date achat	ISBN
1	01/01/2020	Bob	Morane	19 rue des jasmins	Alcool	Poppy Z.	18	01/05/2007	2846261326
2	02/01/2020	Clark	Burger	1 Grand'rue	99 francs	Frédéric B.	20	02/06/2000	1864786260
3	03/01/2020	Bob	Morane	19 rue des jasmins	Galaxie watch	Comics D.	15	15/06/2015	9864157823
4	04/01/2020	Bob	Morane	19 rue des jasmins	Red & Black	Stend H.	15	15/06/2000	8792518573
5	05/01/2020	Bob	Morane	19 rue des jasmins	Les contes	Poppy Z.	18	05/06/2004	8972159735
6	06/01/2020	Bruce	Beauchef	7 rue du Bourg	Miss Wonderful	Marvel	15	16/08/2011	1045284738
7	07/01/2020	Harley	Britchu	8 rue de l'université	Galaxie watch	Comics D.	15	15/06/2015	9864157823
8	08/01/2020	Bob	Morane	19 rue des jasmins	Lost Souls	Poppy Z.	18	06/07/1999	1486489726
9	09/01/2020	Bruce	Beauchef	7 rue du Bourg	Galaxie watch	Comics D.	15	15/06/2015	9864157823

Les cellules grisées représentent les redondances présentes dans la table.



Cela peut sembler anodin, mais des inefficiences mineures peuvent devenir des problèmes majeurs à mesure que la taille de la base de données augmente.

La normalisation consiste à séparer les données en plusieurs tables comportant des éléments de même nature.

Emprunts		Adhérents			Livres				
N° opération	date	nom	prénom	adresse	Livre	auteur	prix d'achat	date achat	ISBN
1	01/01/2020	Bob	Morane	19 rue des jasmins	Alcool	Poppy Z.	18	01/05/2007	2846261326
2	02/01/2020	Clark	Burger	1 Grand'rue	99 francs	Frédéric B.	20	02/06/2000	1864786260
3	03/01/2020	Bob	Morane	19 rue des jasmins	Galaxie watch	Comics D.	15	15/06/2015	9864157823
4	04/01/2020	Bob	Morane	19 rue des jasmins	Red & Black	Stend H.	15	15/06/2000	8792518573
5	05/01/2020	Bob	Morane	19 rue des jasmins	Les contes	Poppy Z.	18	05/06/2004	8972159735
6	06/01/2020	Bruce	Beauchef	7 rue du Bourg	Miss Wonderful	Marvel	15	16/08/2011	1045284738
7	07/01/2020	Harley	Britchu	8 rue de l'université	Galaxie watch	Comics D.	15	15/06/2015	9864157823
8	08/01/2020	Bob	Morane	19 rue des jasmins	Lost Souls	Poppy Z.	18	06/07/1999	1486489726
9	09/01/2020	Bruce	Beauchef	7 rue du Bourg	Galaxie watch	Comics D.	15	15/06/2015	9864157823

La table initiale serait scindée en trois tables :

- ▶ La table T_Emprunt

ID emprunt	date	ID Adhérent	ID livre
1	01/01/2020	1	1
2	02/01/2020	2	2
3	03/01/2020	1	3
4	04/01/2020	1	4
5	05/01/2020	1	5
6	06/01/2020	3	6
7	07/01/2020	4	3
8	08/01/2020	1	7
9	09/01/2020	3	3

- ▶ La table T_Adhérent

ID Adhérent	nom	prénom	adresse
1	Bob	Morane	19 rue des jasmins
2	Clark	Burger	1 Grand rue
3	Bruce	Beauchef	7 rue du Bourg
4	Harley	Britchu	8 rue de l'université
5	Bob	Morane	19 rue des jasmins

- ▶ La table T_Livres

ID livre	Livre	auteur	prix d'achat	date achat	ISBN
1	Alcool	Poppy Z.	18	01/05/2007	2846261326
2	99 francs	Frédéric B.	20	02/06/2000	1864786260
3	Galaxie watch	Comics D.	15	15/06/2015	9864157823
4	Red & Black	Stend H.	15	15/06/2000	8792518573
5	Les contes	Poppy Z.	18	05/06/2004	8972159735
6	Miss Wonderful	Marvel	15	16/08/2011	1045284738
7	Lost Souls	Poppy Z.	18	06/07/1999	1486489726

Vous retrouverez ces trois tables dans l'onglet **tables normalisées**.

2. Importation des tables dans le modèle de données

Si l'onglet **Power Pivot** n'est pas affiché dans le ruban, réalisez la manipulation suivante :

☞ Sélectionnez l'onglet **Fichier** puis **Options**.

La boîte de dialogue **Options Excel** s'affiche à l'écran.

☞ Sélectionnez **Complément**.

☞ Dans la partie inférieure de la boîte de dialogue, au niveau de la liste déroulante **Gérer**, sélectionnez **Complément COM** puis cliquez sur le bouton **Atteindre**.

La boîte de dialogue **Complément COM** apparaît à l'écran.

- ☞ Cochez la case **Microsoft Power Pivot for Excel**.
- ☞ Cliquez sur le bouton **OK**.
- ☞ Sélectionnez une cellule contenue dans la table **T_emprunt**.
- ☞ Dans l'onglet **Power Pivot** - groupe **Tables**, cliquez sur le bouton **Ajouter au modèle de données**.



Une nouvelle fenêtre nommée **Power Pivot** s'ouvre :

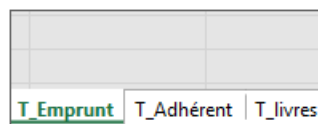
 The screenshot shows the Power Pivot interface in Excel. At the top, there are tabs: "Fichier", "Accueil", "Conception", and "Avancé". Below the tabs is a ribbon with various icons for data operations like "Coller", "Copier", "Actualiser", and "Tableau croisé dynamique". Below the ribbon is a data table with the following content:

ID_emprunt	date	ID_Adhérent	ID_livre	Ajouter une colonne
1	01/01/2020 00:00:00	1	1	
2	02/01/2020 00:00:00	2	2	
3	03/01/2020 00:00:00	1	3	
4	04/01/2020 00:00:00	1	4	
5	05/01/2020 00:00:00	1	5	
6	06/01/2020 00:00:00	3	6	
7	07/01/2020 00:00:00	4	3	
8	08/01/2020 00:00:00	1	7	
9	09/01/2020 00:00:00	3	3	

Cette fenêtre présente les données sous forme de tableau, de la même manière que dans un tableau Excel.

- ☞ En répétant la manipulation précédente, importez les tables **T_Adhérent** et **T_Livres** dans Power Pivot.

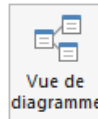
Il est possible de naviguer entre les tables importées dans Power Pivot grâce aux onglets situés au bas de l'interface.



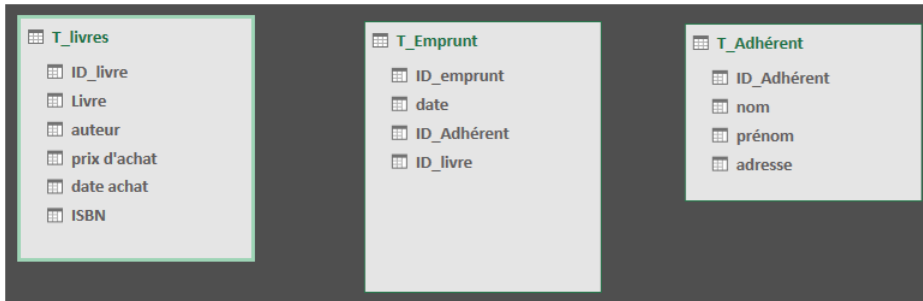
La vue de diagramme

La vue de diagramme est utile pour organiser et créer des relations entre les tables importées dans Power Pivot.

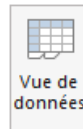
- ☞ Dans l'onglet **Accueil** du ruban de **Power Pivot**, au niveau du groupe **Affichage**, cliquez sur le bouton **Vue de diagramme**.



Le résultat est le suivant :



- ☞ Pour revenir à la vue de données, dans le menu **Accueil** du ruban de **Power Pivot**, au niveau du groupe **Affichage**, cliquez sur le bouton **Vue de données**.



3. Les clés primaires

La clé primaire est le moyen d'identifier une ligne dans une table de manière unique. La plupart du temps la clé primaire est un numéro **unique** généré par un logiciel de base de données.

Nous parlerons alors de **clé primaire synthétique**.

Cependant, il peut exister une colonne dans vos données qui identifie déjà de manière unique chaque ligne, comme par exemple un numéro de sécurité sociale ou un numéro ISBN, dans ce cas, elle pourra faire office de clé primaire.

Nous parlerons alors de **clé primaire naturelle**.