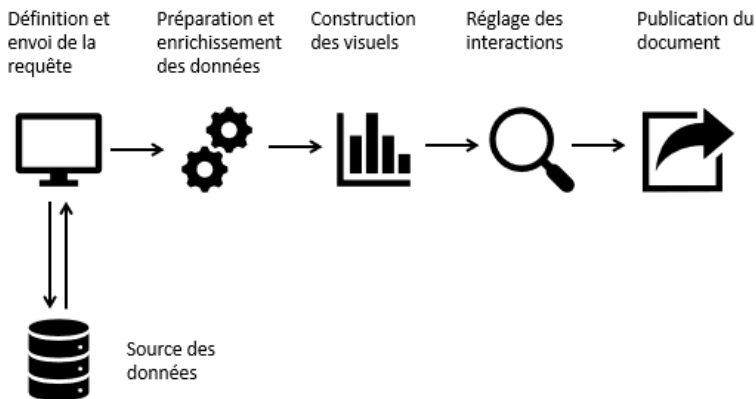


A. Les concepts clés

1. Circulation des données

Comprendre la façon dont circulent et sont stockées les données lorsque vous utilisez Power BI Desktop peut s'avérer utile, notamment lors du choix de la connexion, lors de l'actualisation, de la création des colonnes ou des mesures, et de l'utilisation même du rapport.

Reprenons le schéma du chapitre précédent :

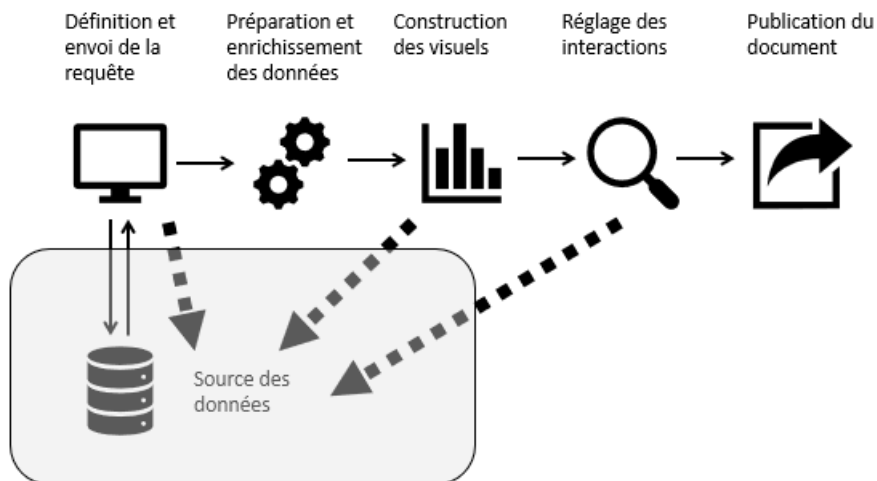


Le premier point clé concerne le type de connexion : la volumétrie des données, ou le besoin de temps (quasi) réel, peuvent vous amener à utiliser une connexion en direct sur votre base de données (mode Direct) – dans ce cas en effet, les données restent dans la base, qui est interrogée à chaque manipulation dans le fichier au prix d'un effort important sur la base.

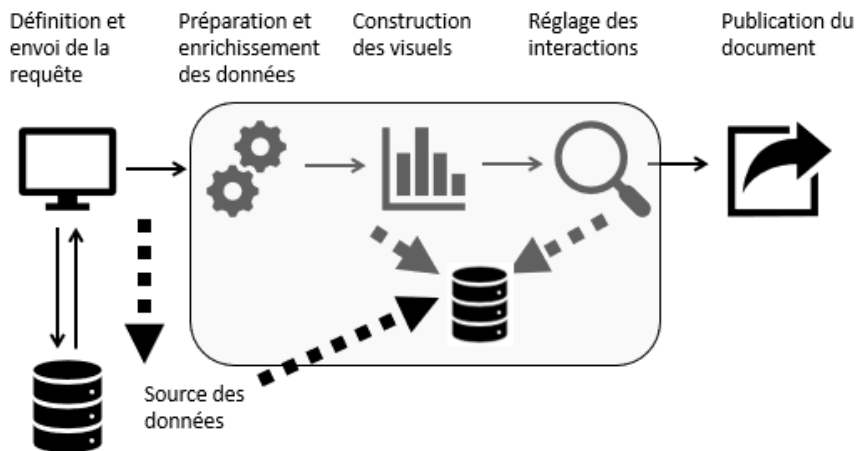


Les types de connexion sont précisés dans la section suivante.

Dans ce premier cas, les données circulent « à la demande » et ne sont pas stockées localement. Une opération aussi simple qu'afficher un total au bas d'un tableau ou filtrer un graphique amène à réinterroger la source. Les performances dépendent des performances du système de base de données :



Dans les autres cas, l'application sera plus performante si les données sont rapatriées localement (mode Import) : la source de données n'est interrogée qu'une fois, lors de l'exécution ou de l'actualisation de la requête, puis les données sont stockées localement, à l'intérieur du fichier.



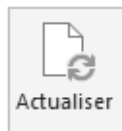
Dans ce cadre, Power BI repose sur un outil très performant : le moteur de stockage tabulaire, VertiPaq, accélère considérablement la mise en place des visuels par la sélection des champs, les interrogations du modèle à chaque utilisation et à chaque calcul d'une mesure, la mise à jour des visuels lors de l'utilisation des filtres et des interactions en général.

Il est également possible de mettre en place des modèles composites, dont une partie des tables (ou des sources), sont exploitées en mode Import, et les autres en mode Direct. Par exemple, des tables Produits ou Sites géographiques, qui évoluent peu, peuvent être importées, mais les tables des transactions, ou des chaînes de production, qui ont parfois besoin d'être interrogées en quasi-temps réel, peuvent rester en mode Direct.

2. Actualisation

Dans le cadre d'une connexion avec importation des données, l'actualisation dans Power BI Desktop se fait exclusivement à l'aide de l'outil **Actualiser**.

Dans le cadre d'une connexion directe à la source de données, nous l'avons vu, l'actualisation dans Power BI Desktop se fait à chaque action sur un visuel. Elle peut aussi être déclenchée manuellement à l'aide de l'outil **Actualiser** dans l'onglet Accueil :



Enfin, c'est dans le cadre de la publication sur Power BI Service (le serveur Power BI) que les possibilités d'actualisation sont les plus complètes.

Lorsque vous publiez un rapport sur Power BI Service, un jeu de données est également envoyé sur le serveur, et disponible aux côtés des rapports et des tableaux de bord (composés par l'utilisateur à partir des visuels présents dans les rapports).

Ce jeu de données peut faire l'objet d'une actualisation immédiate (manuelle), automatique (notamment si le fichier source est stocké dans SharePoint ou OneDrive) ou programmée (planifiée).



Le plus souvent, il faut avoir mis en place une passerelle entre le serveur Power BI et le serveur de base de données, pour que la mise à jour puisse se faire.

Si l'utilisateur a créé un tableau de bord, en épinglant différents visuels disponibles dans les rapports auxquels il a accès, une fonctionnalité **Actualiser les vignettes du tableau de bord** est disponible, pour forcer l'actualisation (en dehors de cette actualisation manuelle, une actualisation automatique a lieu toutes les 15 ou 30 minutes).



Pour certains types de sources – fichiers stockés sur OneDrive, SharePoint, bases de données en ligne Salesforce, base de données Azure, l'actualisation est le plus souvent automatique (toutes les heures par exemple pour des fichiers stockés sur OneDrive).



Il est également possible d'aller chercher uniquement les données ajoutées ou modifiées depuis la précédente actualisation : l'actualisation incrémentielle repose sur l'existence d'une date témoin, et accélère considérablement le temps d'actualisation d'un rapport. Pour plus d'information sur ce point, je vous invite à lire l'article que je lui ai consacré : <https://daxone.fr/un-choix-parmi-les-nouveautes-de-fevrier-2020/>.

B. Se connecter

Power BI propose une vaste bibliothèque de connecteurs – et encore celle-ci évolue avec chaque nouvelle livraison du logiciel : les décrire tous serait fastidieux, mais il est possible de dégager les grandes lignes et le fonctionnement des principaux connecteurs.



Vous trouverez sur le site de l'éditeur une description très complète de la plupart des connecteurs : <https://docs.microsoft.com/fr-fr/power-bi/desktop-data-sources>

1. Les trois types de connexion

Un premier choix essentiel consiste à définir le type de connexion : import ou direct.

- ▶ L'option la plus fréquente consiste à **importer** les données : une fois la requête lancée, Power BI récupère et stocke en local l'ensemble des données.
 - ▶ L'intérêt majeur de ce type de connexion repose sur les performances de Power BI lors de la création de rapports, les possibilités de modélisation et la réactivité aux actions de filtrage de l'utilisateur.
 - ▶ L'inconvénient tient à ce que le poste sur lequel est conçu ou utilisé le rapport doit avoir une puissance suffisante, en termes de RAM (pour le stockage des données) et de CPU (pour les calculs).
 - ▶ Un autre inconvénient possible est la limite de taille du fichier : 1 Go pour la version standard, 10 Go pour les versions Pro et Premium.
 - ▶ Enfin, ce type de connexion implique une actualisation manuelle ou programmée des données – par opposition à la visualisation de données en temps réel.

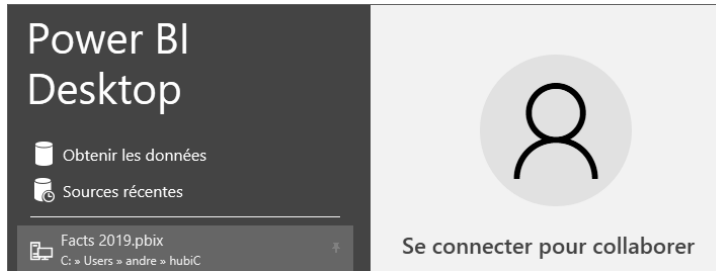
Malgré ces quelques inconvénients, l'import des données est le mode à privilégier.

- ▶ Il est également possible de se connecter en **direct** à certaines sources (bases de données).
 - ▶ L'intérêt ici est d'avoir en permanence des rapports à jour, sans avoir besoin d'actualiser les données. Une simple action (filtre ou sélection dans le visuel) suffit à mettre à jour les données. Par ailleurs, la puissance du poste n'est plus un critère limitant.
 - ▶ En revanche, un inconvénient tient à la capacité de la source : il faut qu'elle soit suffisamment puissante pour accepter une connexion directe. Les performances de l'application peuvent s'en ressentir.
 - ▶ Par ailleurs, toutes les sources ne sont pas disponibles.
 - ▶ Les possibilités de modélisation des données sont limitées (puisqu'elles ont déjà été définies dans la source).
 - ▶ Enfin d'autres limitations incluent la restriction de certaines fonctionnalités de transformation et de création de formule (notamment les fonctions Time Intelligence).

Il existe deux types de connexion en direct : la connexion DirectQuery et la connexion directe. La première est plus générale, mais aussi moins performante, la seconde concerne exclusivement SSAS, Azure ou Power BI Service – elle est plus performante dans la mesure où les données sont déjà stockées sur le modèle qu'utilise Power BI.

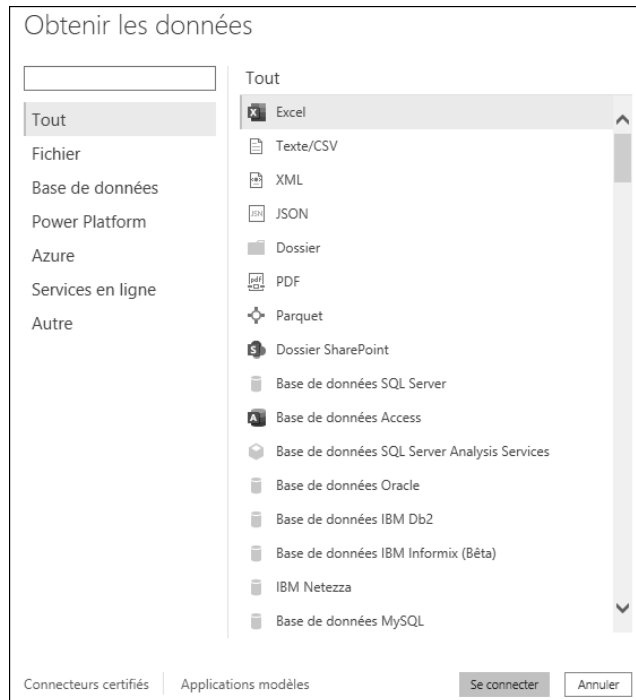
Découvrir les types de connexion

☞ Lancez Power BI Desktop et cliquez sur **Obtenir les données**.



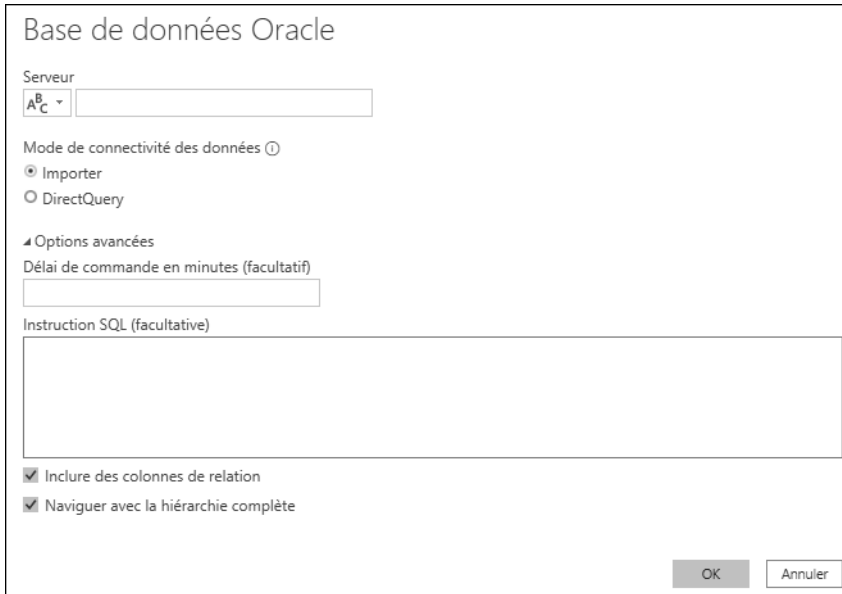
*Vous trouvez aussi la fonctionnalité **Obtenir les données** sur l'onglet **Accueil**.*

☞ Pour accéder à une source, sélectionnez le connecteur dans la liste, puis cliquez sur **Se connecter**.



En fonction du connecteur choisi, des options apparaissent, le nom de la base, le mode de connectivité.

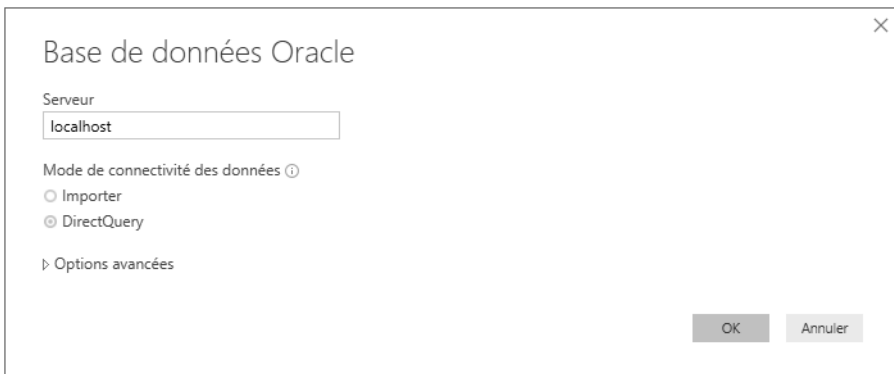
Voici quelques exemples de connecteurs à des bases de données Oracle et SQL Server :



The screenshot shows a dialog box titled "Base de données Oracle". It contains the following elements:

- Serveur:** A text input field with a dropdown menu showing "A B C".
- Mode de connectivité des données:** Two radio buttons: "Importer" (selected) and "DirectQuery".
- Options avancées:** A collapsed section containing:
 - Délai de commande en minutes (facultatif):** A text input field.
 - Instruction SQL (facultative):** A large text area.
 - Inclure des colonnes de relation
 - Naviguer avec la hiérarchie complète
- Buttons:** "OK" and "Annuler" at the bottom right.

Connexion avec import sur une base Oracle : dans les options avancées, remarquez la possibilité de saisir directement du code SQL.



The screenshot shows a dialog box titled "Base de données Oracle". It contains the following elements:

- Serveur:** A text input field containing "localhost".
- Mode de connectivité des données:** Two radio buttons: "Importer" and "DirectQuery" (selected).
- Options avancées:** A collapsed section with a right-pointing arrow.
- Buttons:** "OK" and "Annuler" at the bottom right.

Connexion DirectQuery sur une base Oracle