

PROCEDURE DE FORMATION



POWER BI

**Création d'un visuel de type scatterplot complexe
avec poignées de zoom sur échelles d'axes**

1. CONTEXTE ET OBJECTIF

Power BI est une puissante plateforme de visualisation de données développée par Microsoft qui permet de créer des graphiques interactifs, des tableaux de bord dynamiques et des rapports analytiques professionnels.

Dans ce guide, nous allons vous montrer comment créer un scatterplot complexe à l'aide de Power BI Desktop, en partant de l'importation des données jusqu'à l'optimisation finale du visuel pour une présentation professionnelle.

Cas pratique : Analyse de la relation entre l'indice de mortalité et l'accès à l'eau potable par région géographique.

2. PREREQUIS

Avant de commencer, vous allez avoir besoin :

- d'une version opérationnelle de Power BI Desktop
- du fichier « **data_pbi00001.csv** » qui contient les données nécessaires à l'élaboration du graphique.

NOTA : ces données sanitaires sur l'accès à l'eau potable dans le monde et à la mortalité sont publiques et conformes au RGPD (source FAO)

3. INSTRUCTIONS

3.1 Etape 1 : Importation des données

3.1.1 Connexion à la source de donnée

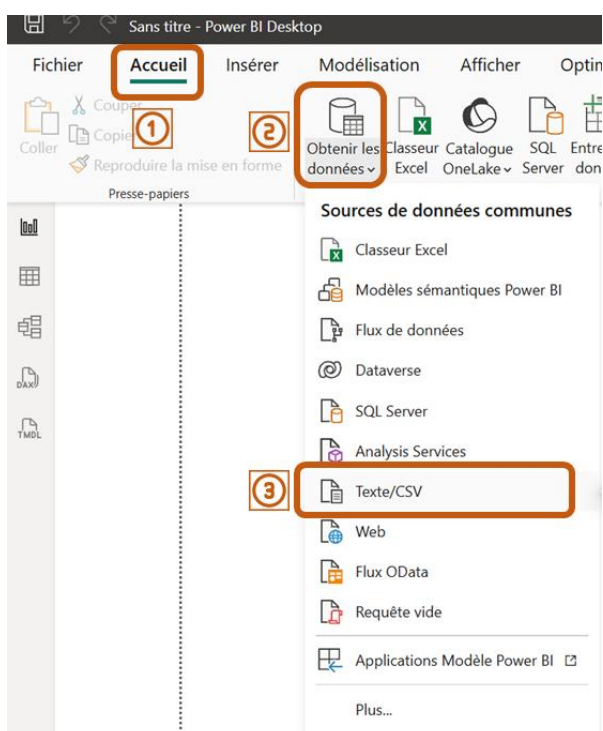
- Ouvrir Power BI Desktop
- Lancer l'application Power BI Desktop
- Cliquer sur **Fichier** puis sur **Nouveau / Créer un nouveau rapport vierge**
- Enregistrer** tout de suite votre fichier avant tout travail !



Rapport vierge

Maintenant nous allons nous connecter au fichier source **data_pbi00001.csv** et importer ses données :

- Dans le ruban **Accueil**, cliquer sur "**Obtenir les données**"
- Sélectionner "**Texte/CSV**"



- Naviguer jusqu'à l'emplacement du fichier source **data_pbi00001.csv**
- Cliquer sur "**Ouvrir**", une fenêtre s'ouvre : il s'agit d'un aperçu des données de votre CSV.
- Cliquer sur « **Transformer les données** ».

data_pbi00001.csv

Origine du fichier: 65001: Unicode (UTF-8) | Délimiteur: Virgule | Détection du type de données: Selon les 200 premières lignes

Mortalité 2016 (%)	Accès eau potable (%)	Nombre de décès 2016	Région	Pays
0.01	96.68	0	Europe	Estonia
0.01	97.935	0	Eastern Mediterranean	Qatar
0.01	100	1	Eastern Mediterranean	Kuwait
0.02	96.775	0	Europe	Latvia
0.02	98.69	0	Europe	Slovenia
0.03	95.05	0	Europe	Montenegro
0.03	97.41	3	Eastern Mediterranean	United Arab Emirates
0.03	100	0	Europe	Malta
0.04	99.49	1	Eastern Mediterranean	Bahrain
0.04	99.745	2	Europe	Slovakia
0.04	99.815	2	Europe	Finland
0.04	100	4	Europe	Greece
0.05	90.92	2	Eastern Mediterranean	Oman
0.05	95.485	5	Europe	Belarus
0.05	99.84	0	Europe	Luxembourg
0.05	99.9	0	Western Pacific	Brunei Darussalam
0.07	87.68	2	Europe	North Macedonia
0.07	97.345	4	Europe	Ireland
0.07	98.89	0	Americas	Bahamas
0.07	100	4	Western Pacific	Singapore

Les données dans l'aperçu ont été tronquées en raison de limites de taille.

Extraire une table avec des exemples | Charger | **Transformer les données** | Annuler

NOTA : il est toujours préférable de choisir cette option car il est rare, même si l'aperçu semble correct, de ne pas avoir de transformations à faire sur les données avant de les charger.

3.1.2 Transformations des données

La fenêtre Editeur de Power Query s'ouvre. Nous allons maintenant faire les vérifications sur les données.

Il est en effet recommandé de s'assurer que les points suivants sont conformes avant de les charger complètement.

- **Vérifier** que :
 - Les en-têtes de colonnes sont correctement détectés.
 - Les types de données sont appropriés (nombre, texte, etc.)
 - Les séparateurs sont bien reconnus.

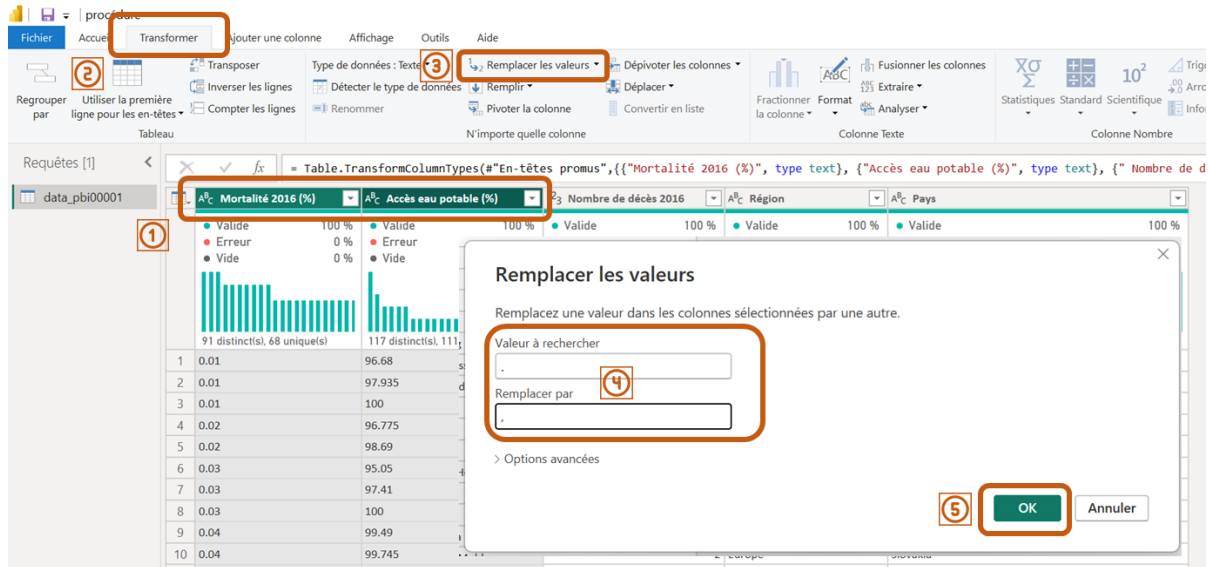
Dans notre cas, des anomalies de type et de séparateurs sont à corriger :

Mortalité 2016 (%)	Accès eau potable (%)	Nombre de décès 2016
<p>Valide 100 %</p> <p>Erreur 0 %</p> <p>Vide 0 %</p> <p>91 distinct(s), 68 unique(s)</p> <p>0.01</p> <p>0.01</p>	<p>Valide 100 %</p> <p>Erreur 0 %</p> <p>Vide 0 %</p> <p>117 distinct(s), 111 unique(s)</p> <p>96.68</p> <p>97.935</p>	<p>Valide 100 %</p> <p>Erreur 0 %</p> <p>Vide 0 %</p> <p>84 distinct(s), 71 unique(s)</p> <p>0</p> <p>0</p>

« . » Au lieu de « , »

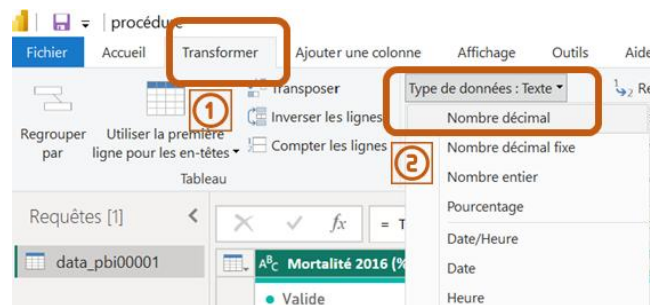
Pour changer le type de données dans ce cas présent :

- **Sélectionner** les 2 colonnes successivement, en maintenant la touche CTRL de votre clavier.
- Cliquer sur le menu **Transformer**
- Puis **Remplacer les valeurs**
- Et enfin dans la fenêtre qui apparaît, **renseigner les paramètres tel qu'indiqué (4) et valider**



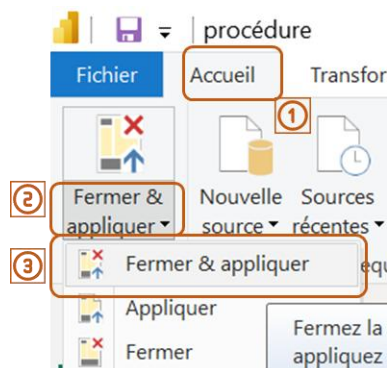
Il ne nous reste plus qu'à modifier le typage :

- Menu **Transformer**, puis sur **Type de données**
- Choisir **Nombre décimal**



Dernière étape dans cette partie, valider les transformations :

- Menu **Accueil**, bouton **Fermer & Appliquer** et dans le menu sélectionner **Fermer & Appliquer**



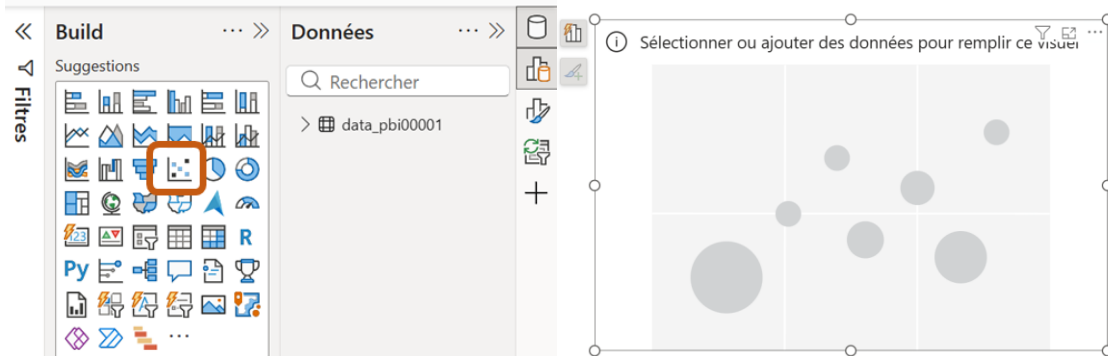
➔ Après quelques secondes, l'ensemble des transformations ont été appliquées aux données sources et la table nettoyée est désormais prête pour créer nos 1^{er} visuels.

3.2 ÉTAPE 2 : CRÉATION DU VISUEL DE BASE

Nous voulons visualiser la relation entre le taux de mortalité (dû à l'eau insalubre) et l'accès à l'eau potable. Chaque point correspond à un pays, la taille étant proportionnel au nombre de décès (dû à l'eau insalubre). Nous souhaitons par ailleurs avoir une légende permettant de regrouper la visualisation par région.

3.2.1 Sélection du type de visuel

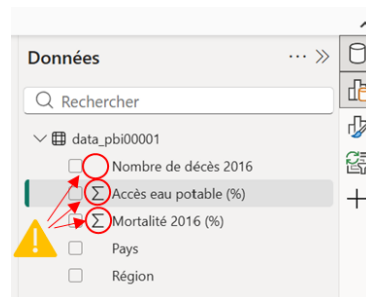
- Dans le volet "Visualisations" (panneau de droite), **cliquer** sur l'icône "**Graphique de nuage de points**".
- Dans la zone du rapport, un espace réservé vide apparaît sur le canevas.



3.2.2 Ajustement du paramétrage des champs avant utilisation dans le visuel

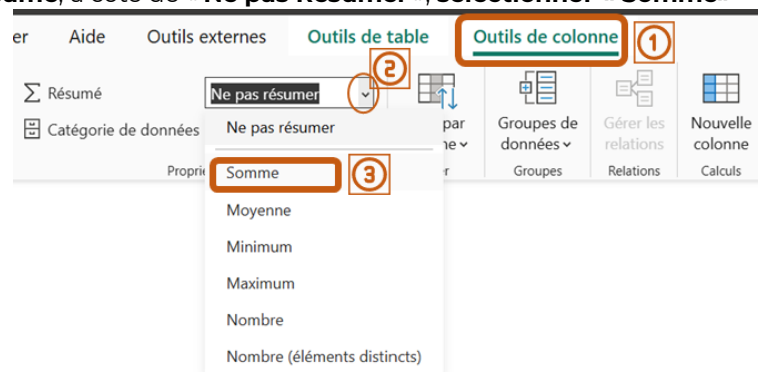
Par défaut, Power BI réalise un résumé des valeurs numériques décimales afin de faciliter la réalisation de certains types de visuels.

Dans notre cas, les colonnes accès eau potable et Mortalité seront utilisées comme axes, et non comme champs valeurs. A l'inverse nous souhaitons utiliser le champs Nombre de décès comme valeurs, il faudra donc lui aussi appliquer une modification afin de faire la somme.



➤ Pour le champs « Nombre de décès » : ajout d'une somme automatique.

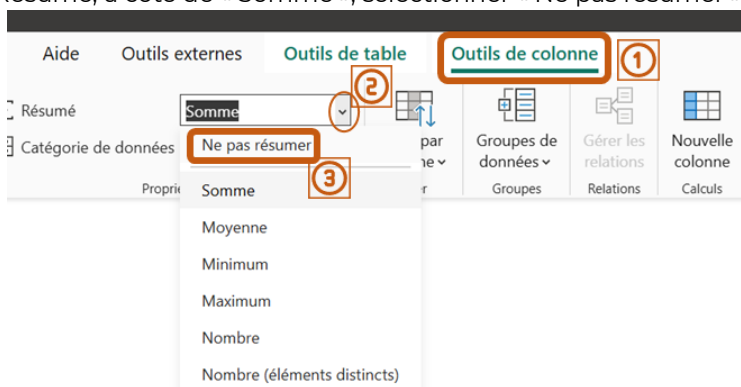
- Aller dans le menu **Outils de colonne**
- Dans la zone **Résumé**, à côté de « Ne pas Résumer », sélectionner « Somme »



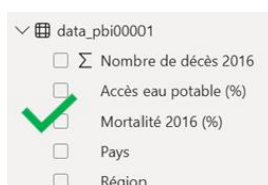
- Pour les champs « Accès eau potable » et « Mortalité » : retrait des sommes automatiques.

Sur le même principe, pour chaque colonne, l'une après l'autre :

- Aller dans le menu Outils de colonne
- Dans la zone Résumé, à côté de « Somme », sélectionner « Ne pas résumer »

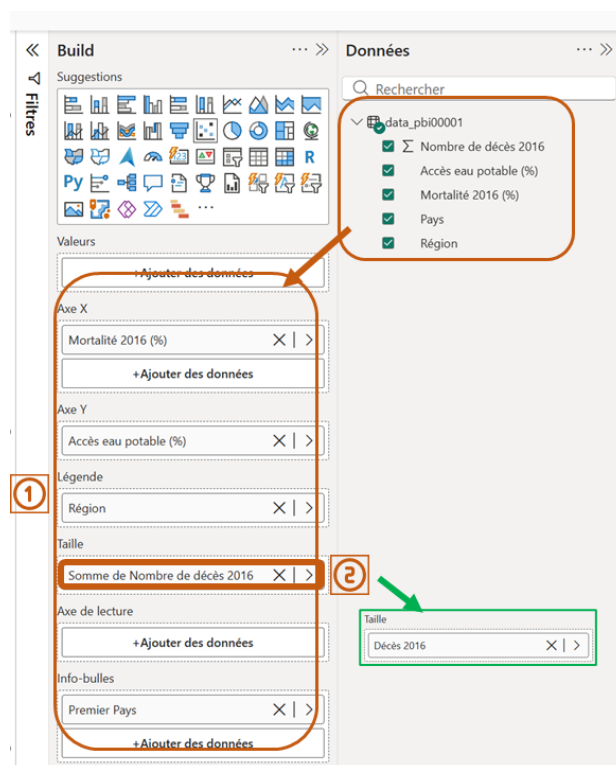


Nous sommes prêts à réaliser notre visuel !



3.2.3 Ajout des champs au visuel

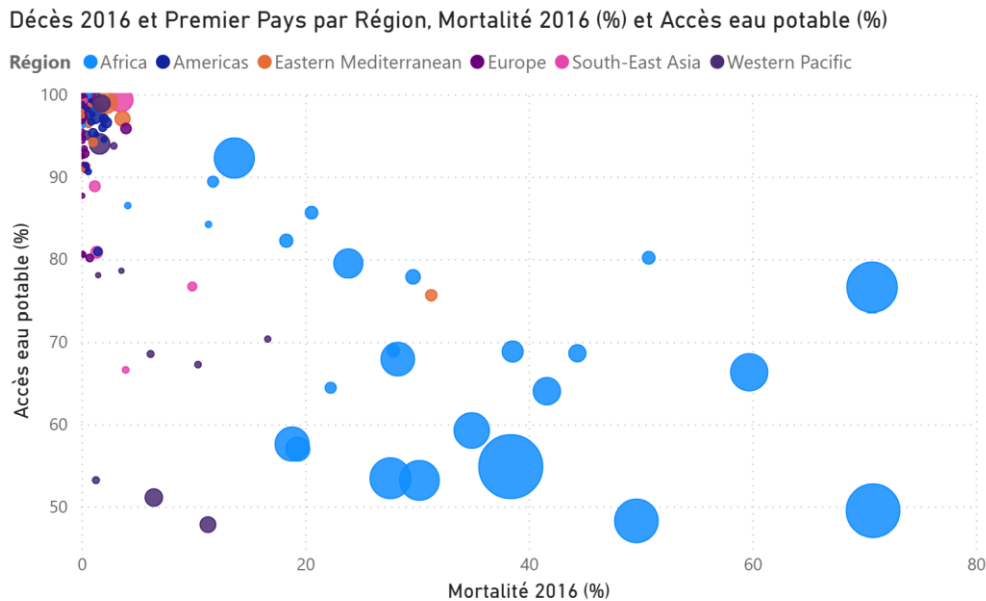
- Cliquer sur votre visuel vide pour l'activer.
- Dans le volet « Visualisations » (panneau de droite), ajouter les champs tels qu'indiqué sur la figure ci-dessous



NOTA : N'hésitez pas à renommer les colonnes en double cliquant sur un champs. C'est ce que j'ai fait à l'étape 2 en vert pour clarifier le contenu.

3.2.4 Visuel obtenu

À ce stade, un nuage de points basique doit apparaître. Les points doivent avoir des tailles différentes. Les couleurs doivent correspondre aux catégories de la légende.



3.3 ÉTAPE 3 : FORMATAGE ET PERSONNALISATION

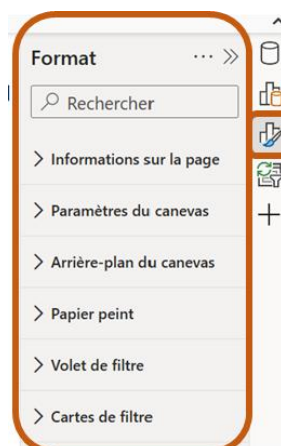
On va maintenant améliorer le visuel en travaillant sur le formatage, modifier les titres (graphique et axes) et améliorer l'expérience utilisateur.

En effet, on voit assez clairement que ce sont majoritairement certains pays d'Afrique qui sont les plus touchés par la mortalité dû à l'eau insalubre, en part de la population (Taux de mortalité) et en nombre de décès.

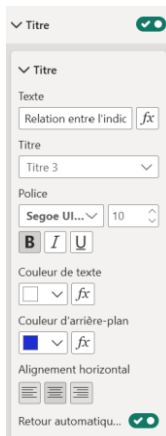
D'un point de vue visualisation, en l'état, l'utilisateur ne peut pas visualiser les pays ayant une mortalité faible et un accès à l'eau fort qui regroupe la grande majorité des pays représentés.

3.3.1 Accès au options de formatage

- Avec le visuel sélectionné, **cliquer sur l'icône "Format"** (pinceau)
- Les sections de formatage s'affichent



3.3.2 Configuration du titre



- **Dérouler la section "Titre"**
- **Modifier le texte** du titre, par exemple : *"Relation entre indice de mortalité et accès à l'eau potable (Taille = Nombre de décès)"*
- **Ajuster** la police, la taille et l'alignement **selon vos préférences**

3.3.3 Personnalisation de la légende



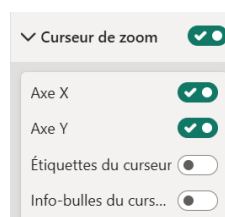
- **Dérouler la section "Légende"** et **configurer** :
 - Position : **En haut ou en bas selon la place disponible**
 - Titre : **ici à supprimer**, car suffisamment explicite
 - Taille du texte : **Ajuster** pour la lisibilité
 - Vérifier que toutes les catégories sont visibles

3.3.4 Ajout de poignées de zoom sur les axes

L'objectif ici est de permettre à l'utilisateur de zoomer sur la zone dense de points (ici en haut à gauche).

Plutôt que de modifier les minimum et maximum manuellement dans les options de chaque axe, une section **«Curseur de zoom»** permet d'ajouter cette fonctionnalité.

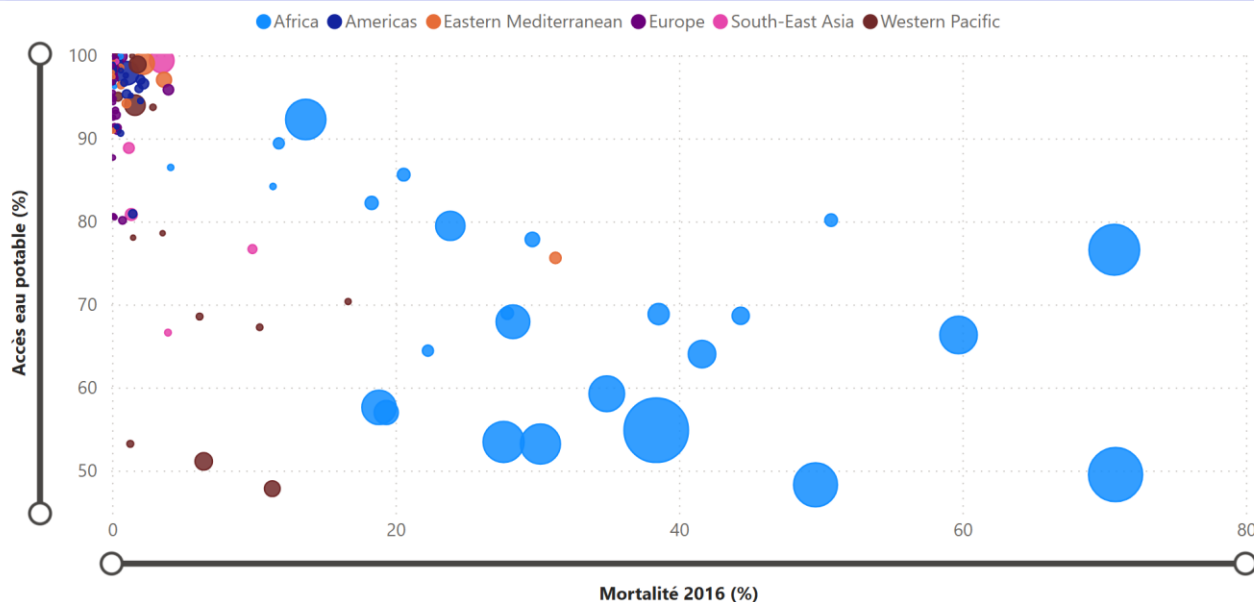
- **Activer le curseur de zoom**
- **Activer les boutons pour les axes Axe X et Axe Y.**



Cela va faire apparaître des poignées sur chaque axe et ainsi faciliter l'analyse des données pour l'utilisateur.

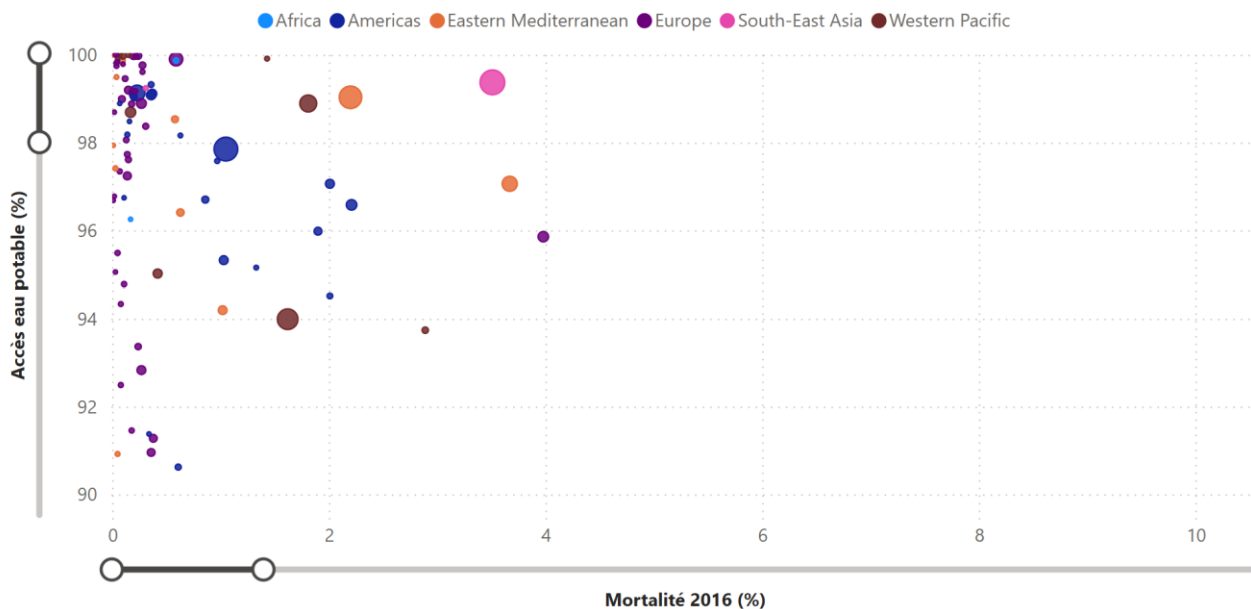
Voici maintenant un exemple du visuel avec optimisation

Relation entre l'indice de mortalité et l'accès à l'eau potable par pays et par zone : taille du point = nombre de décès du pays



⇒ Exemple avec fonction de zoom présente mais non utilisée

Relation entre l'indice de mortalité et l'accès à l'eau potable par pays et par zone : taille du point = nombre de décès du pays



⇒ Exemple avec fonction de zoom active (= analyse du cluster possible)