



POWER BI, PYTHON & Datavisualisation

(example : Machine Learning)

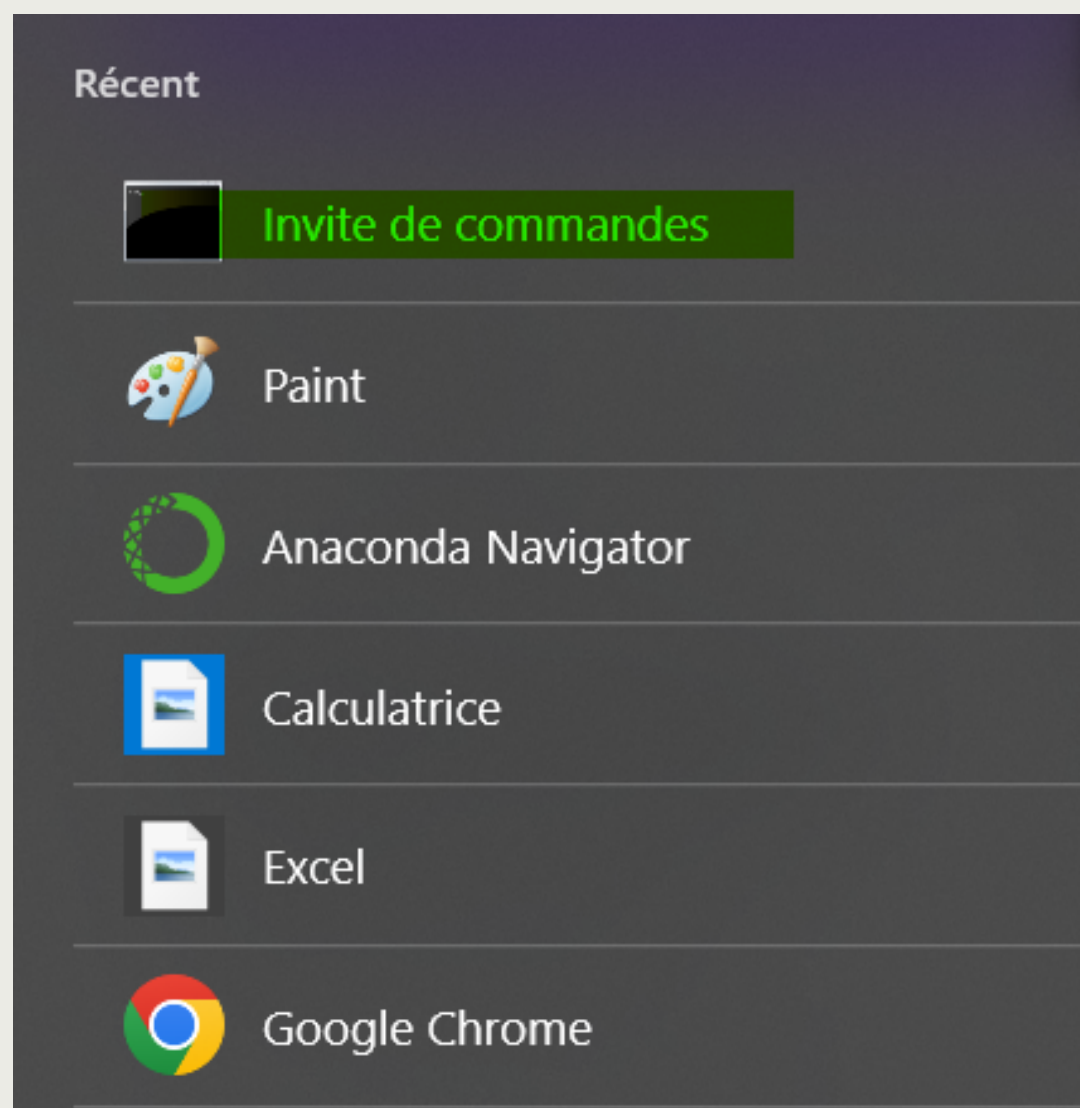
Combiner la puissance du langage Python et Power BI

Vous voulez aller plus loin dans l'analyse de données ? Grâce à l'intégration de Python dans Power BI, vous pouvez visualiser des modèles de Machine Learning et les partager en quelques clics avec vos équipes non techniques.

Exemple : Création d'un modèle d'arbre de décision pour prédire des résultats, puis visualisation dans Power BI.

Prérequis :

- Python doit être installé sur votre machine.
- Disponibilité des librairies nécessaires comme pandas, scikit-learn, et matplotlib.



Saisir “pip list” et vérifier les bibliothèques installées

```
Invite de commandes

Microsoft Windows [version 10.0.19045.4894]
(c) Microsoft Corporation. Tous droits réservés.

C:\Users\nickz>pip list
Package                                Version
-----
altair                                  5.0.1
asttokens                              2.2.1
attrs                                  23.1.0
backcall                               0.2.0
blinker                                1.6.2
cachetools                             5.3.1
certifi                                2023.5.7
charset-normalizer                     3.1.0
click                                   8.1.3
colorama                               0.4.6
comm                                    0.1.3
contourpy                              1.0.7
cyclor                                 0.11.0
debugpy                                1.6.7
decorator                              5.1.1
executing                              1.2.0
fonttools                              4.39.3
gitdb                                   4.0.10
GitPython                              3.1.31
```

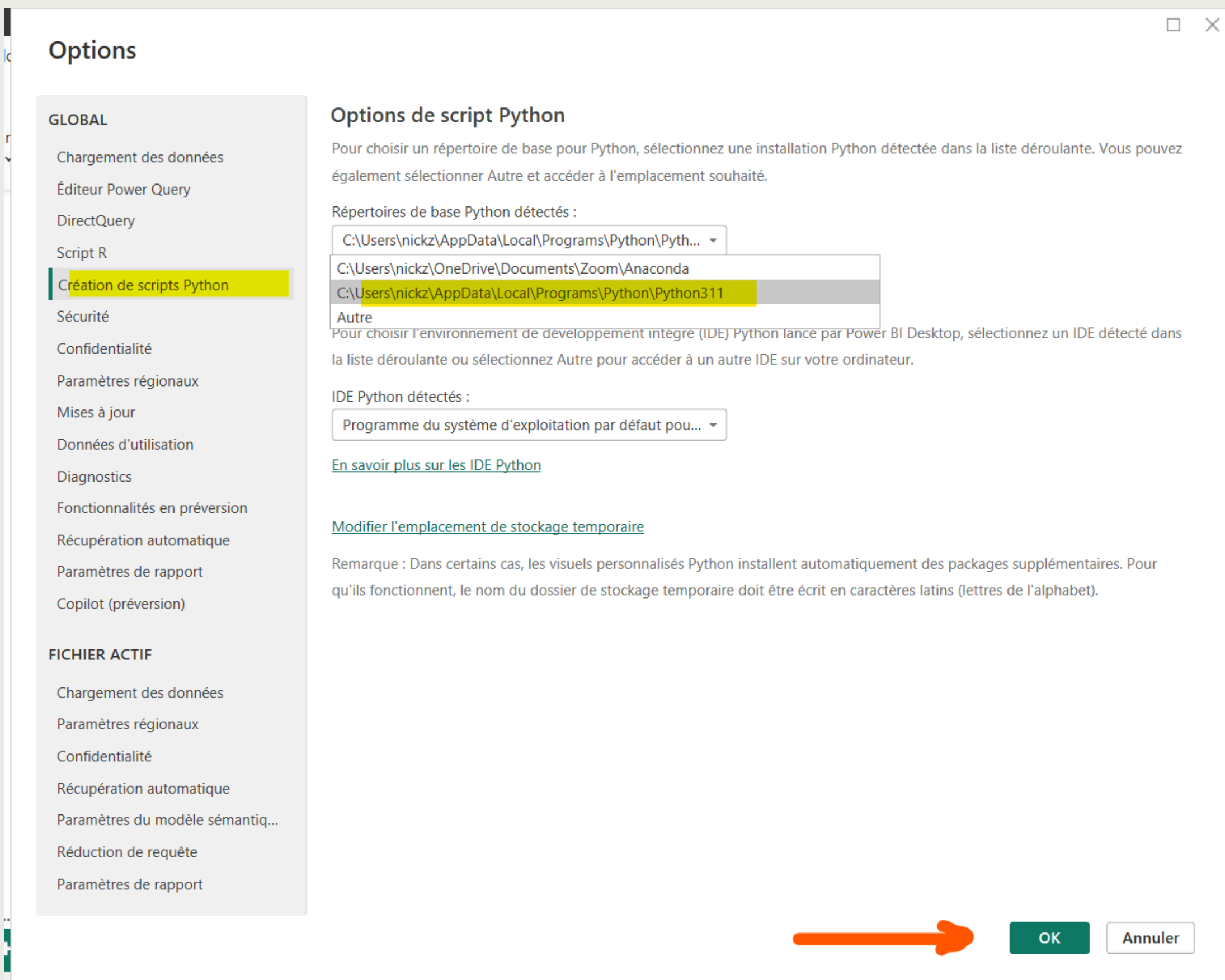
Le jeu de données utilisé pour ce
“pas à pas” peut être téléchargé ici :

<https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset>


Etape 1

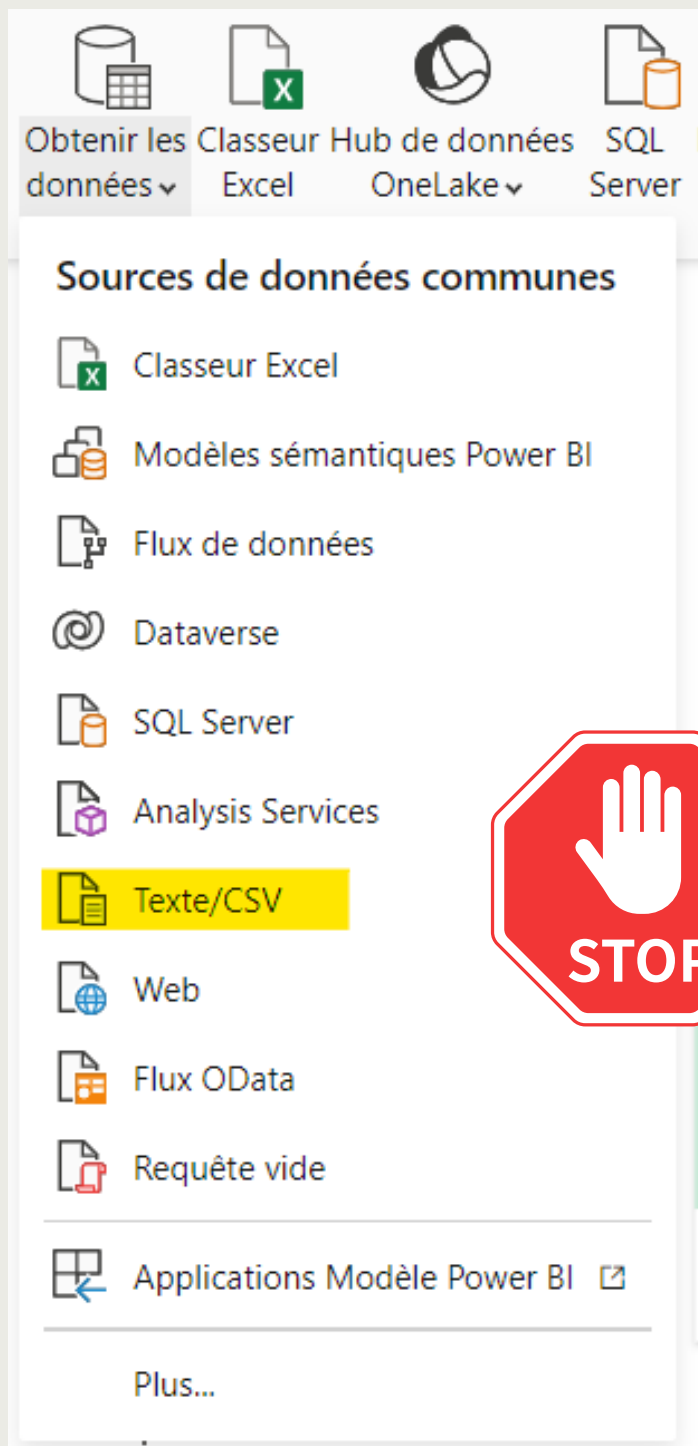
- Sélectionnez une installation Python dans Power BI















- Importez à présent le jeu de données.
Dans l'exemple, le fichier est au format CSV



 On ne choisira pas “format csv” dans le menu déroulant :



La bonne méthode est la suivante :


Sources de données communes

-  Classeur Excel
-  Modèles sémantiques Power BI
-  Flux de données
-  Dataverse
-  SQL Server
-  Analysis Services
-  Texte/CSV
-  Web
-  Flux OData
-  Requête vide

 Applications Modèle Power BI 

Plus...


Obtenir les données

python 

Tout


Autre

Tout

 Script Python

Connecteurs certifiés

Applications modèles

 Se connecter

Annuler

Script Python

Script

Le script sera exécuté avec l'installation de Python suivante :

C:\USERS\NICKZ\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON311.

Pour configurer vos paramètres et changer l'installation de Python à exécuter, accédez à Options et paramètres.

OK

Annuler

- Saisissez dans la fenêtre de script qui vient de s'ouvrir la commande nécessaire pour récupérer les données. Dans notre exemple :

```
import pandas as pd
df = pd.read_csv("nomdufichier.csv")
print(df)
```

Le nom du fichier dans ce cas est :

C:\Users\nickz\OneDrive\Bureau\breast-cancer.csv

Mais les \ entraînent ce message d'erreur :

```
df = pd.read_csv("C:\Users\u202a\nickz\OneDrive\Bureau\breast-
cancer.csv")
```

^

SyntaxError: (unicode error) 'unicodeescape' codec can't decode
bytes in position 2-3: truncated \UXXXXXXXX escape

</pi> »

Réessayer

Modifier

Annuler

Le code corrigé :

Script Python

Script

```
import pandas as pd

df = pd.read_csv("C:/Users/nickz/OneDrive/Bureau/breast-cancer.csv")

print(df)
```

Le script sera exécuté avec l'installation de Python suivante :

C:\USERS\NICKZ\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON311.

Pour configurer vos paramètres et changer l'installation de Python à exécuter, accédez à Options et paramètres.

OK

Annuler

Navigateur

Options d'affichage ▾

Python [1]

df

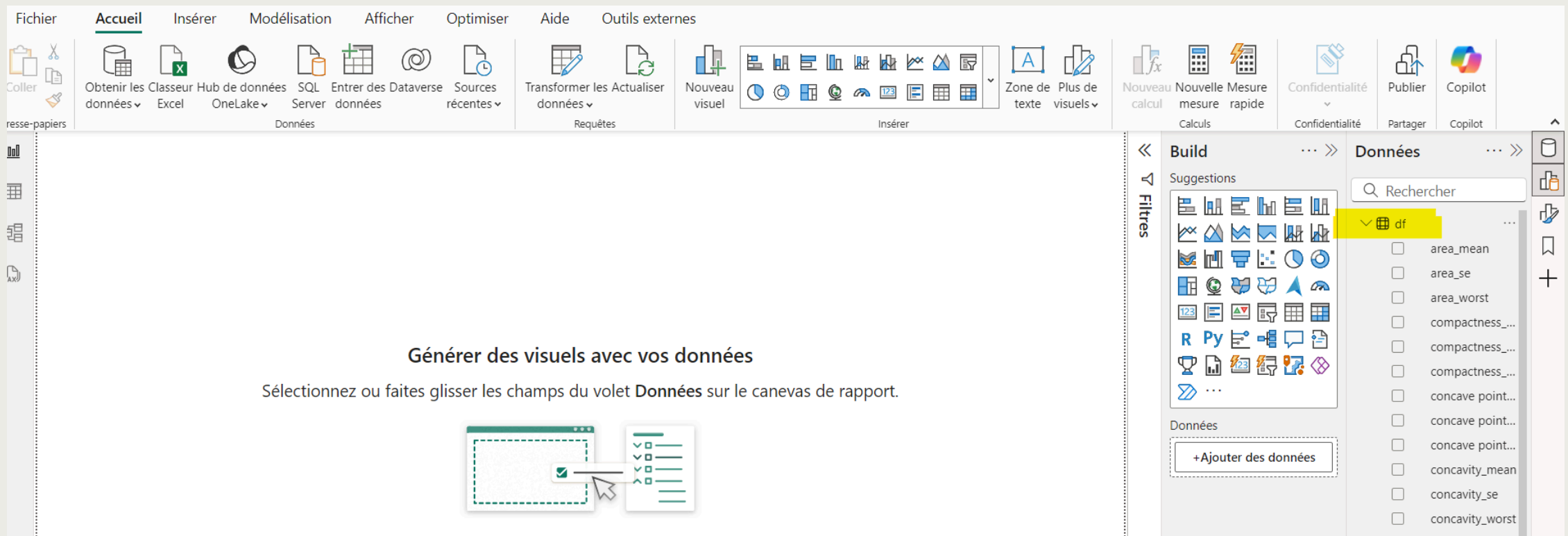


df

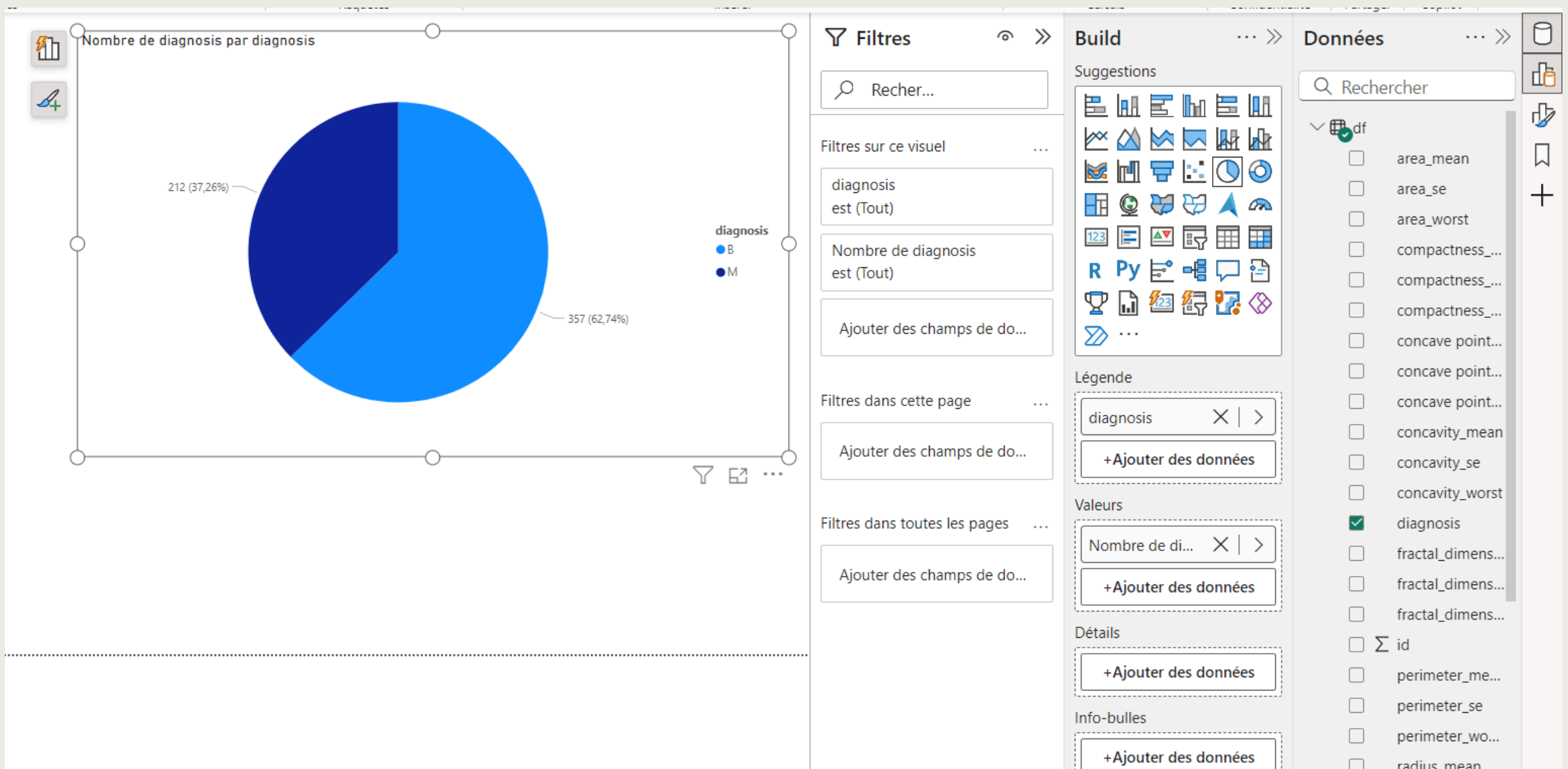
id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean
842302	M	17.99	10.38	122.8	101.3
842517	M	20.57	17.77	132.9	132.6
84300903	M	19.69	21.25	130.0	128.3
84348301	M	11.42	20.38	77.58	38.1
84358402	M	20.29	14.34	135.1	127.1
843786	M	12.45	15.7	82.57	47.4
844359	M	18.25	19.98	119.6	101.8
84458202	M	13.71	20.83	90.2	57.1
844981	M	13.0	21.82	87.5	51.0
84501001	M	12.46	24.04	83.97	47.4
845636	M	16.02	23.24	102.7	79.0
84610002	M	15.78	17.89	103.6	78.4
846226	M	19.17	24.8	132.4	111.5
846381	M	15.85	23.95	103.7	78.7
84667401	M	13.73	22.61	93.6	57.1

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

- Votre jeu de données est à présent utilisable

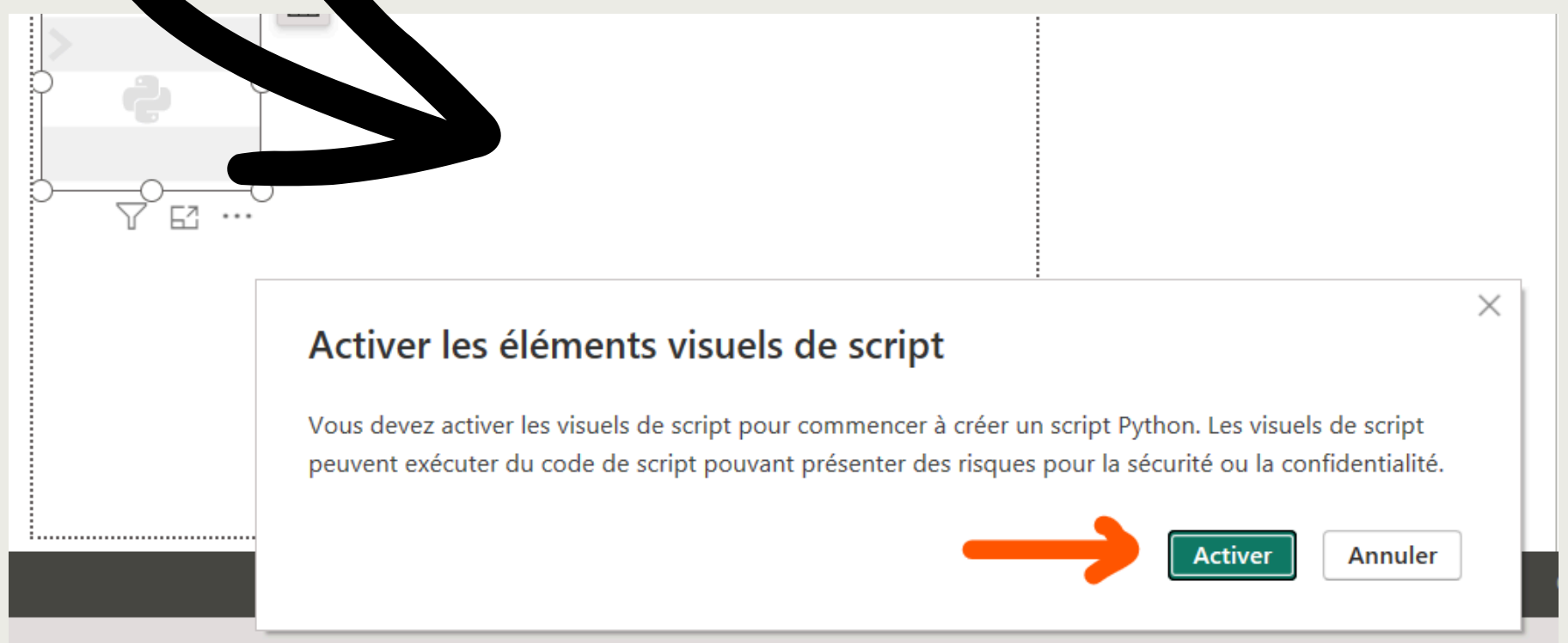
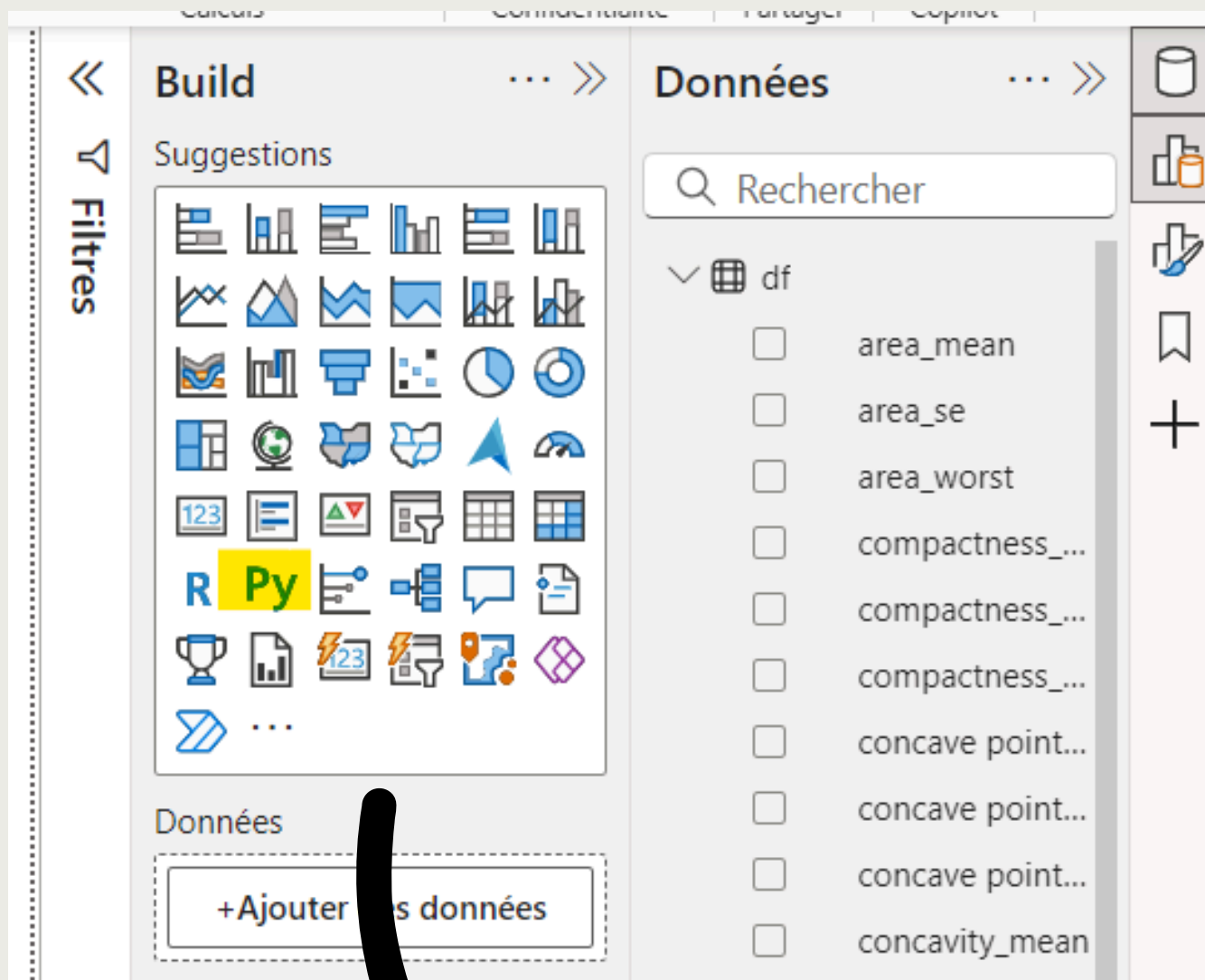


- Vous pouvez créer les visuels souhaités. Exemple :

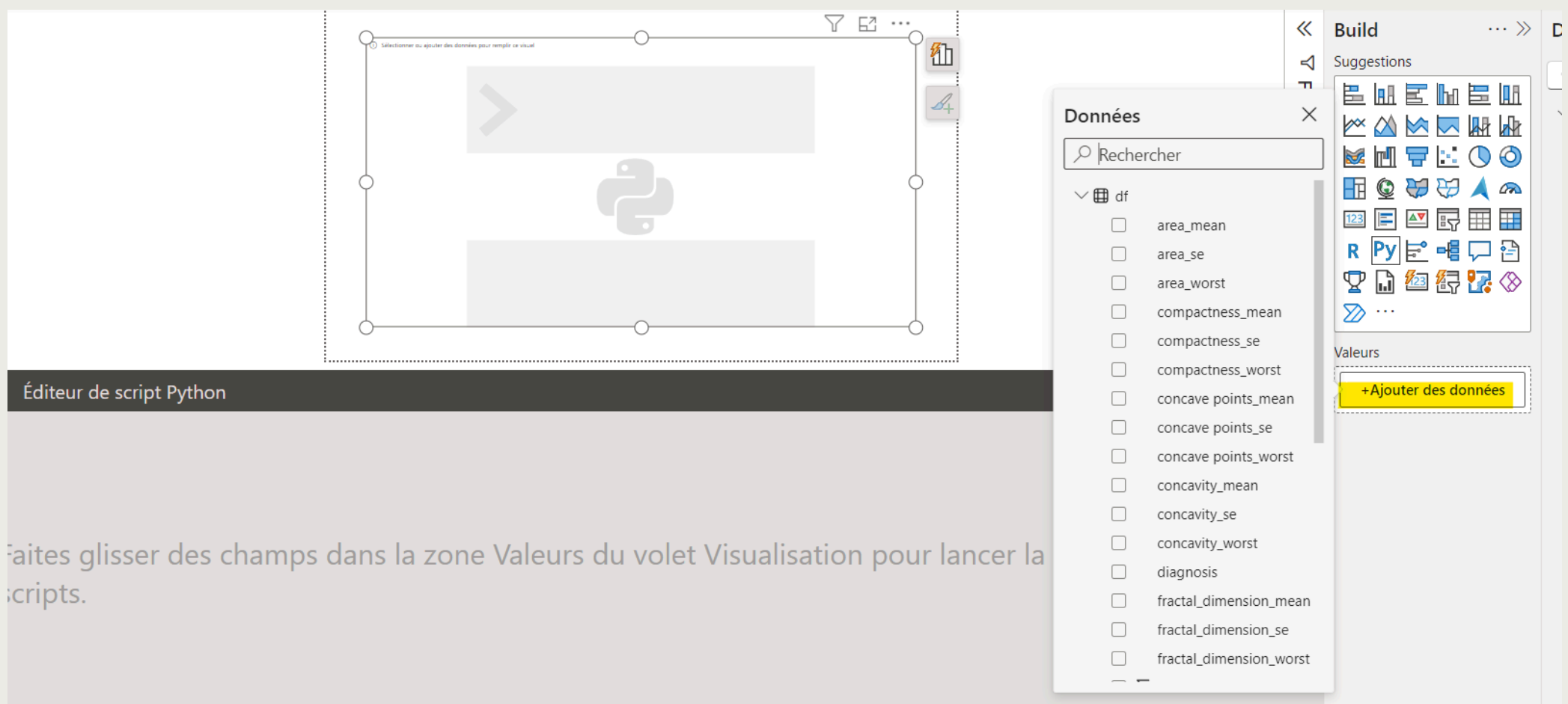


Etape 2

- Rédaction de script Python dans Power BI



- Sélectionnez les données du dataset df nécessaire à la création de vos futurs graphiques, visuels, lignes de code ...



- Dans notre exemple, nous allons tout sélectionner avant de rédiger le code pour obtenir le résultat souhaité. Au fur et à mesure que les cases sont cochées, votre “dataset power BI” est complété

L'utilisation de scripts Python dans Power BI vous ouvre encore d'autres horizons pour toujours plus de visuels et d'analyses.

Besoin du code pour vous entraîner?
Ne pas hésiter à me contacter!