

Club Power BI

La communauté Power BI francophone



/Club-Power-BI



@ClubPowerBI



/ClubPowerBI



/ClubPowerBI

Merci à notre hôte

Mobiskill

Power BI en 2018 - Rétrospective

January	February	March	April	May	June
Reporting <ul style="list-style-type: none"> Show and hide pages Control data label background color for Cartesian and maps visuals Increase area used for axis labels in charts Bar/column padding control Show dates as a hierarchy (preview) Add an anchor date for relative date slicer Top N selection in Q&A PowerApps custom visual Analytics <ul style="list-style-type: none"> Correlation coefficient quick measure Data connectivity <ul style="list-style-type: none"> Support for Azure Active Directory authentication for Azure SQL Database & Data Warehouse connectors Other <ul style="list-style-type: none"> Advanced language settings for the Windows store app 	Reporting <ul style="list-style-type: none"> Multi-select data points across multiple charts Sync slicers across multiple report pages Numeric range slicer improvement – whole number snapping (preview) Faster geocoding for Bing maps (web preview) Overflow data labels for bar and column charts Search the formatting and analytics pane Organizational custom visuals (preview) Analytics <ul style="list-style-type: none"> Mark custom date table Quick measures are generally available UTCNOW() and UTCTODAY() Data connectivity <ul style="list-style-type: none"> Improvements to DirectQuery over SAP HANA (preview) Other <ul style="list-style-type: none"> DirectQuery performance improvements Open and save performance improvements "Show items with no data" improvements 	Reporting <ul style="list-style-type: none"> Report page tooltips (preview) Bookmarking is now generally available Display units and precision control for table & matrix columns Turn off the visual header in reading mode for a report Improved default visual placement Mapbox custom visual Data connectivity <ul style="list-style-type: none"> SAP HANA connector enhancements: DirectQuery multidimensional support GA & SSL certificate validation support SAP BW DirectQuery GA Azure Analysis Services connector GA Other <ul style="list-style-type: none"> Improved error reporting View previous errors you've encountered 	Reporting <ul style="list-style-type: none"> Q&A Explorer (preview) Buttons to trigger actions Combo chart line styles Improved default sort for visuals Numeric Slicer GA Organizational custom visuals GA Analytics <ul style="list-style-type: none"> Update your linguistic schema New DAX function: COMBINEVALUES() Data connectivity <ul style="list-style-type: none"> Adobe Analytics connector now generally available SAP HANA connector improvement: Port selection Dynamics 365 Business Central connector 	Reporting <ul style="list-style-type: none"> Conditional formatting by a different field Advanced slicer syncing Log axis improvements Data label options for funnel chart Set line stroke width to zero Analytics <ul style="list-style-type: none"> Measure drill-through Data connectivity <ul style="list-style-type: none"> New "From Web" connector - By example data extraction (preview) Common Data Service for Apps connector (beta) Azure KustoDB connector (beta) Google BigQuery and Azure HDInsight Spark connectors now generally available Adobe Analytics connector update - Support for multiple domain logins (preview) Visual Studio Team Services connector update - Analytics views support OLE DB connector update - Support for alternate Windows credentials SAP BW DirectQuery connector update - Improved technical name support Data preparation <ul style="list-style-type: none"> Improvements to Add Column From Examples 	Reporting <ul style="list-style-type: none"> High contrast support for reports Donut radius control Pie and donut detail labels position control Format data labels separately for each measure in a combo chart Longer phone reports Modeling <ul style="list-style-type: none"> Filtering and sorting in data view Improved locale formatting Data connectivity <ul style="list-style-type: none"> SAP BusinessWarehouse connector improvements - New driver and improved performance & Improved support for hierarchy variables Spark connector now supports Windows Authentication OData V4 connector enhancements ODBC connector improvements Folding support for Top Rows Ability to filter navigation by DSN catalog Other <ul style="list-style-type: none"> National cloud selector



Power BI en 2018 - Rétrospective

July	August	September	October	November	December
Reporting <ul style="list-style-type: none"> New visual header with more flexibility and formatting Wallpaper formatting Theming update – more visual and page control (preview) Tooltips for table and matrix Turn tooltips off for visuals Slicer accessibility Formatting pane improvements Stepped line support for line and combo charts Turn off combo chart data labels for individual series Sorting experience improvement Modeling <ul style="list-style-type: none"> Composite models (preview) Analytics <ul style="list-style-type: none"> Distribution factor insights Data Connectivity <ul style="list-style-type: none"> IBM DB2 DirectQuery connector (preview) Improvements to Web By Example connector - Importing multiple custom tables, Automatic completions for sample values, & exposure of attribute selectors SAP HANA – Default values for variables in Variable Input experience 	Reporting <ul style="list-style-type: none"> Print reports through Export to PDF Create bookmark groups Theming generally available Slicer restatement Modeling <ul style="list-style-type: none"> Data categories for measures Statistical DAX functions Analytics <ul style="list-style-type: none"> Conditional formatting by values Python integration (preview) Q&A improvements Q&A in Power BI Desktop generally available Better matching experience for Power BI Desktop and Power BI Premium Data connectivity <ul style="list-style-type: none"> Spark connector generally available 	Reporting <ul style="list-style-type: none"> Dot plot layout support in scatter charts Copy value and selection from table & matrix Built-in report theme options Report page tooltips generally available & Card support Accessibility improvements for analytics and formatting pane support Analytics <ul style="list-style-type: none"> Aggregations (preview) Q&A support for RLS Data connectivity <ul style="list-style-type: none"> PDF file connector (preview) SAP BW connector – support for measure properties Dataflows connector (beta) Data preparation <ul style="list-style-type: none"> M intellisense (preview) Add Column From Examples – support for text padding 	Reporting <ul style="list-style-type: none"> Search in filter cards Improved accessible authoring experiences Performance improvements for ArcGIS Map Modeling <ul style="list-style-type: none"> DAX editor improvements Analytics <ul style="list-style-type: none"> Composite models & Aggregation support in the Power BI service (preview) Explain the increase for non-additive measures Data connectivity <ul style="list-style-type: none"> Web By Example connector is now generally available SAP BW Connector implementation v2 is now generally available SAP BW Message Server Connector is now generally available Vertica connector is now generally available Dynamics NAV and Dynamics 365 Business Central connectors are now generally available New Dynamics 365 Business Central On-premises connector Data preparation <ul style="list-style-type: none"> Data Profiling in Query Editor (preview) Fuzzy Matching options for Merge Queries (preview) Other <ul style="list-style-type: none"> Control export data options for your reports 	Reporting <ul style="list-style-type: none"> Expand & collapse matrix row headers Copy & paste between Desktop files Updated filtering experience (preview) Report accessibility improvements Analytics <ul style="list-style-type: none"> Color saturation on visuals upgraded to use conditional formatting Related questions in the Q&A explorer Modeling <ul style="list-style-type: none"> New modeling view (preview) Composite models now generally available Modeling accessibility improvements New DAX function: Optional DrilldownFilter argument for the RollupAddIsSubtotal function New DAX function: NonVisual function New DAX function: IsInScope function Data connectivity <ul style="list-style-type: none"> Azure DevOps Server connector PDF Connector support for Start Page and End Page parameters Improved Azure Consumption Insights connector 	Reporting <ul style="list-style-type: none"> Smart guides for aligning objects on a page Fields list accessibility support Set tab order for objects on a page Tooltips for button visuals Updated icons for "Visual interactions" Analytics <ul style="list-style-type: none"> Live connect support for Q&A (preview) Modeling <ul style="list-style-type: none"> DAX formula bar updates Data view accessibility support Data connectivity <ul style="list-style-type: none"> AtScale connector (Beta) Oracle Essbase connector (Beta) Data preparation <ul style="list-style-type: none"> Fuzzy merge - support for Top N best matches Other <ul style="list-style-type: none"> High contrast support for all panes and report footer Updated keyboard shortcut dialog



Le Club Power BI en 2018

18 Meetups

1 Power
Saturday

Lille

Aix-En-
Provence

Nantes

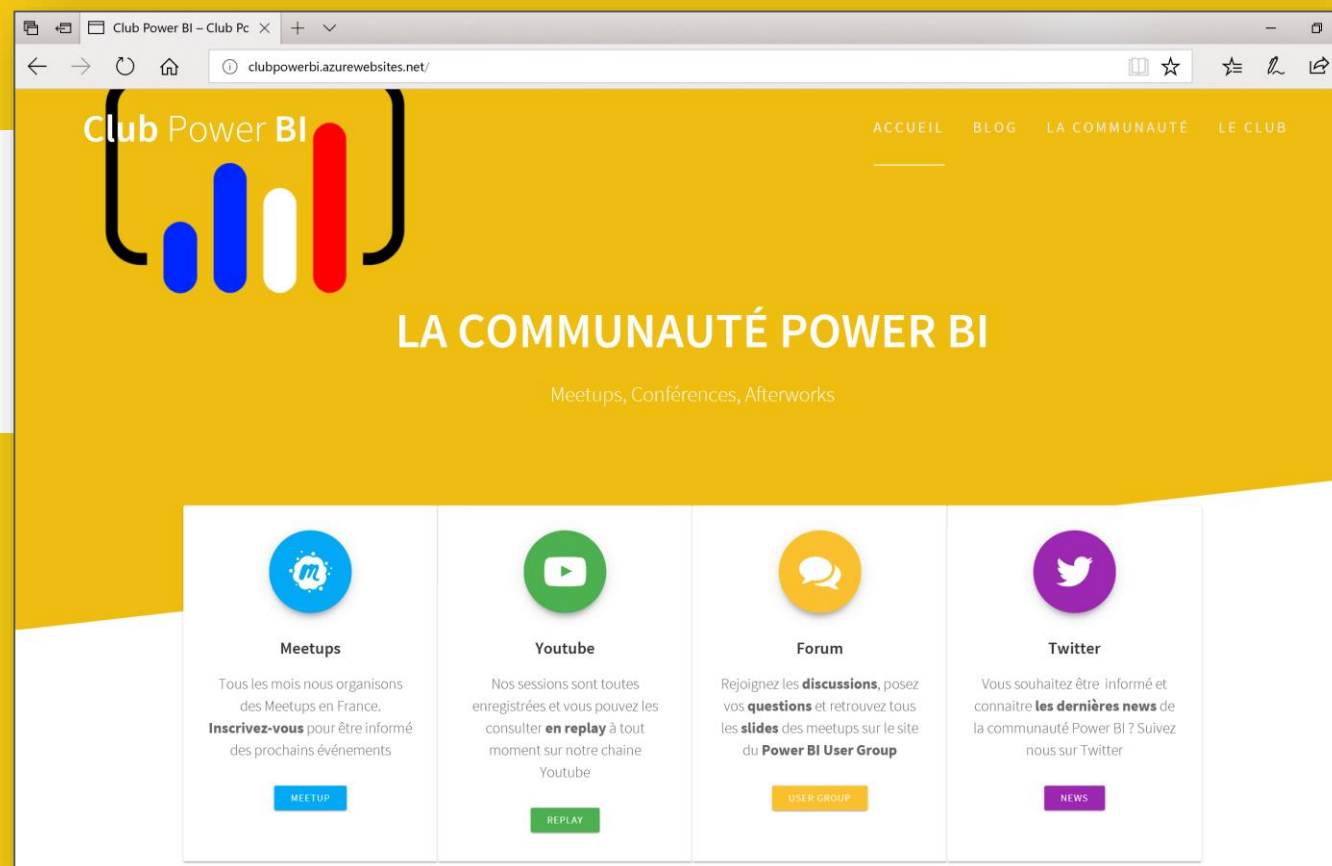
1 000+
membres

25 heures
sur Youtube



Club Power BI

<http://clubpowerbi.com>



Power Platform **Summit** **EUROPE**

*Stay ahead of the business intelligence
demands of today's data driven world.*

REGISTER TODAY

AMSTERDAM | 27 - 29 MARCH



Sondage



Afin de mieux connaître votre utilisation de Power BI et de vous proposer du meilleur contenu en 2019, Le Club vous sollicite pour répondre à ce court sondage disponible ici:

[Bit.ly/sondageclubpowerbi2019](https://bit.ly/sondageclubpowerbi2019)



Vous pouvez remplir ce sondage en moins de 5 minutes sur ordinateurs, téléphones et tablettes. Nous comptons sur vous !

Les nouveautés sur Power BI Service

Jean-Pierre Riehl – Co Fondateur du Club Power BI



DAX – Récupérer des valeurs du contexte

Tristan Malherbe – Co Fondateur du Club Power BI



Fonctions DAX étudiées

SelectedValue()

HasOneValue()

Concatenex()

Values()

TopN()

démo

Synthèse & Cas d'usage

- ❖ SelectedValue() ou ConcatenX() pour récupérer du **texte**
- ❖ Eventuellement Min() ou Max() pour du **numérique**

Quelques **cas d'usage** métiers:

- « ***Dynamiser*** » l'affichage: Titres, Pages d'atterrissage (Drillthrough) & informations de contexte
- **Switch**: Mensuel / YTD, Current/LY, Actuals/Budget
- « **What-If** » & **segmentation** : Stockage d'un paramètre choisi par l'utilisateur... qu'on réinjecte dans un calcul ailleurs

Python dans Power BI Desktop

Paul PETON – Data Strategist | Microsoft AI MVP

Sommaire

- ❖ Place de Python dans l'offre Microsoft
- ❖ Paramétrage de Power BI Desktop pour utiliser Python
- ❖ Trois utilisations possibles :
 - l'import de données
 - les transformations dans l'éditeur de requêtes
 - la visualisation de données
- ❖ En synthèse :
 - Valeurs ajoutées du langage Python (versus M & DAX) ?
 - Limites de l'utilisation de Python dans Power BI ?
 - Comment faire une analyse prédictive ?

Python ?



- ❖ Première version le 20 février 1991
- ❖ Langage objet, impératif et interprété
- ❖ Typage dynamique
- ❖ Multiplateforme
- ❖ Visuellement épuré
 - Identification des blocs par indentation
- ❖ Avec aussi un groupe Meetup : PyData Paris

The string **02201991** occurs at position 190,623,086 counting from the first digit after the decimal point :
30639226089751607610**02201991**24389287874198947054

```
def factorielle(n):  
    if n < 2:  
        return 1  
    else:  
        return n * factorielle(n - 1)
```

lun. 21 janv., 19:00

PyData Paris - January 2019 Meetup

📍 Institut de Physique du Globe de Paris



Mark your calendar for the next session of the PyData Paris Meetup on January 21st 2019. This Meetup will be hosted by IPGP (Institut de Physique du Globe de Paris), rue Jussieu. The speakers for this session are Joris Van den Bossche and Viviane...



130 participants

Participer

La notion fondamentale de dataframe

❖ **dataframe** : « *in memory dataset* »

❖ Collection de vecteurs

- de même taille
- mais dont les types peuvent être différents

❖ S'écrit DataFrame en Python

- Demande d'appeler la librairie **pandas**

➤ `import pandas as pd`

```
df = pd.DataFrame({'n': [2, 3, 5], 's': ["aa", "bb", "cc"],  
                  'b': [TRUE, FALSE, TRUE]}, columns = ['n', 's', 'b'])
```

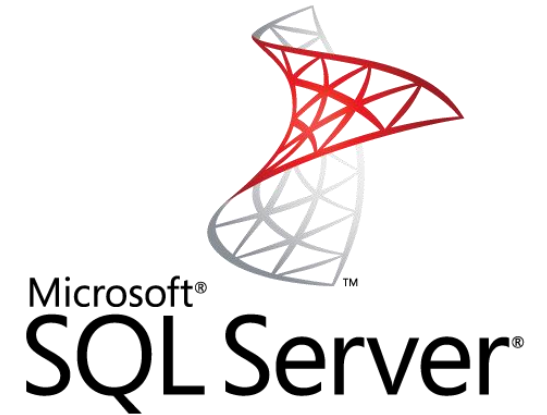
Python dans l'offre Microsoft

❖ On Premise

- SQL Server Machine Learning Services
 - Encapsulé dans un script T-SQL
- Machine Learning Server
 - Travail en « *remote connection* » sur une grappe de VM

❖ Sur Azure

- Azure (Jupyter) Notebooks
- Azure Machine Learning Studio
- Azure DataBricks
 - Notebooks automatisables grâce à Azure Data Factory v2
- Azure Functions (preview)
 - « *serverless* » mais pas de gestion de la capacité de traitement
- A venir dans SQL DataBase (service managé)

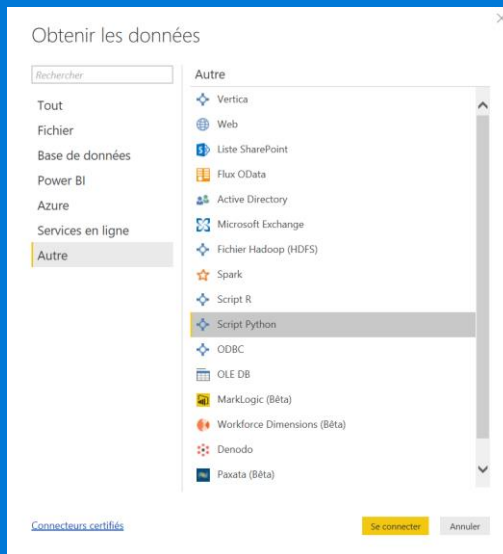


Trois possibilités d'utilisation des scripts Python

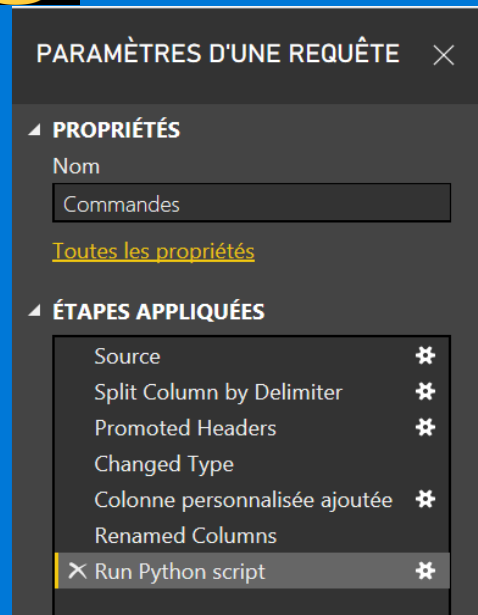
Power BI
Desktop



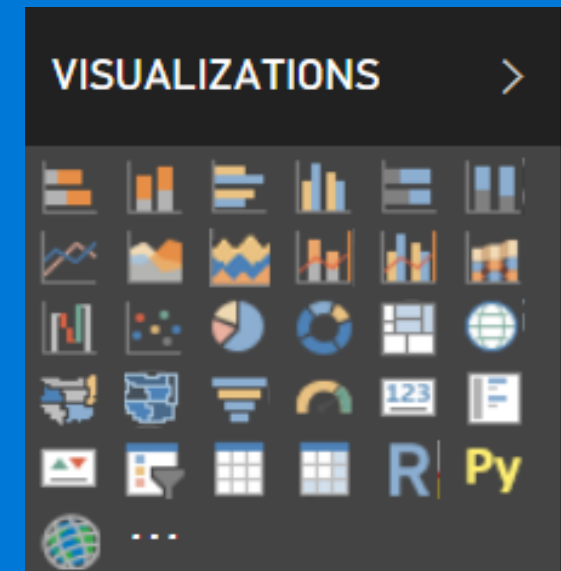
Connector



in Power Query



Visuals



Configurer Power BI Desktop pour Python

- ❖ Sur le poste exécutant Power BI Desktop :
 - Installer une distribution de Python : ANACONDA
 - Adaptée pour l'approche analytique
 - ... dans le PATH !
 - Et un IDE : Visual Studio Code ?
- ❖ Dans les options de Power BI Desktop :
 - Activer la préversion de Python
 - Vérifier la reconnaissance des installations et IDE



Options de Power BI Desktop

Options

GLOBAL

Chargement des données
Éditeur Power Query
DirectQuery
Script R
Création de scripts Python
Sécurité
Confidentialité
Paramètres régionaux
Mises à jour
Données d'utilisation
Diagnostics
Fonctionnalités en préversion
Récupération automatique
Paramètres de rapport

Options de script Python

Pour choisir un répertoire de base pour Python, sélectionnez une installation Python détectée dans la liste déroulante. Vous pouvez également sélectionner Autre et accéder à l'emplacement souhaité.

Répertoires de base Python détectés :

C:\Anaconda

[Comment installer Python](#)

Pour choisir l'environnement de développement intégré (IDE) Python lancé par Power BI Desktop, sélectionnez un IDE détecté dans la liste déroulante ou sélectionnez Autre pour accéder à un autre IDE sur votre ordinateur.

IDE Python détectés :

Visual Studio Code

[En savoir plus sur les IDE Python](#)

[Modifier l'emplacement de stockage temporaire](#)

Remarque : Dans certains cas, les visuels personnalisés Python installent automatiquement des packages supplémentaires. Pour qu'ils fonctionnent, le nom du dossier de stockage temporaire doit être écrit en caractères latins (lettres de l'alphabet).

FICHIER ACTIF

Chargement des données
Paramètres régionaux
Confidentialité
Récupération automatique
Réduction de requête
Paramètres de rapport

OK

Annuler

Options

GLOBAL

Chargement des données
Éditeur Power Query
DirectQuery
Script R
Création de scripts Python
Sécurité
Confidentialité
Paramètres régionaux
Mises à jour
Données d'utilisation
Diagnostics
Fonctionnalités en préversion
Récupération automatique
Paramètres de rapport

Fonctionnalités en préversion

Les fonctionnalités suivantes sont disponibles à l'essai dans cette version. Les fonctionnalités en version préliminaire sont susceptibles d'être modifiées ou supprimées dans les versions futures.

- ✓ Visuel Carte de formes [En savoir plus](#)
- ✓ M Intellisense [En savoir plus](#)
- ☐ Prise en charge de l'espagnol dans Questions et réponses [En savoir plus](#)
- ✓ Obtenir les données des fichiers PDF [En savoir plus](#)
- ✓ Activer le profilage de colonne [En savoir plus](#)
- ✓ Afficher les dates sous forme de hiérarchie dans la liste de champs [En savoir plus](#)
- ✓ Prise en charge de Python [En savoir plus](#)
- ✓ Stratégies d'actualisation incrémentielle [En savoir plus](#)
- ✓ Gérer les agrégations [En savoir plus](#)
- ✓ Activer la fusion approximative [En savoir plus](#)
- ✓ Vue de modélisation [En savoir plus](#)
- ✓ Nouvelle expérience de filtre [En savoir plus](#)
- ✓ Activer Questions et réponses Live Connect [En savoir plus](#)

FICHIER ACTIF

Chargement des données
Paramètres régionaux
Confidentialité
Récupération automatique
Réduction de requête
Paramètres de rapport

OK

Annuler

DEMONSTRATION

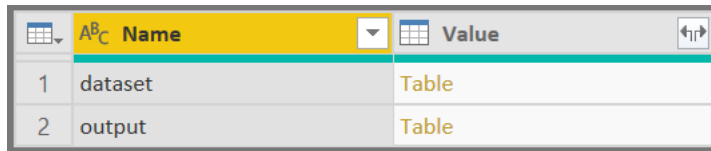
Python dans Power BI Desktop

Importer des données avec Python

- ❖ Nouvelle source de données
- ❖ Ecrire du code pour obtenir un **dataframe** qui sera ensuite interprété comme un objet de requête « table »
- ❖ Importer les librairies nécessaires dans le code
- ❖ **Attention, pas de reconnaissance des types de données**
- ❖ Quelques cas d'usage :
 - Exploitation de code déjà existant
 - Web Scraping (chargement avancé de pages web)
 - Fichiers dans des formats non reconnus par Power BI
 - ZIP, TAR.GZ...
 - SAS (*.sas7bdat), SPSS (*.sav)...

Transformer des données avec Python

- ❖ Etape de transformation dans l'éditeur de requête
- ❖ L'étape précédente est interprétée comme un
 - **pandas dataframe**
 - `# 'dataset' holds the input data for this script`
- ❖ Le code doit générer un ou plusieurs nouveaux dataframes
 - L'un de ces dataframes sera ensuite ouvert par une étape de navigation



	ABC Name	Value
1	dataset	Table
2	output	Table

- ❖ Le code peut réaliser toutes les transformations souhaitées
 - En appelant les librairies nécessaires
- ❖ Idée : réaliser un export des données !
 - `pandas.DataFrame.to_csv`

Générer un visuel en partant de zéro

- ❖ Choisir les champs nécessaires à la construction du visuel
- ❖ Ecrire le code complet (appel aux packages inclus) avec Python

- matplotlib
- seaborn

- ❖ Générer un objet graphique
 - Le code doit se terminer par :
 - `plt.show()`

Impossible d'afficher cet élément visuel. ✕

Aucune image n'a été créée. Le code Python n'a créé aucun visuel. Vérifiez que votre script Python génère un tracé sur l'appareil Python par défaut.

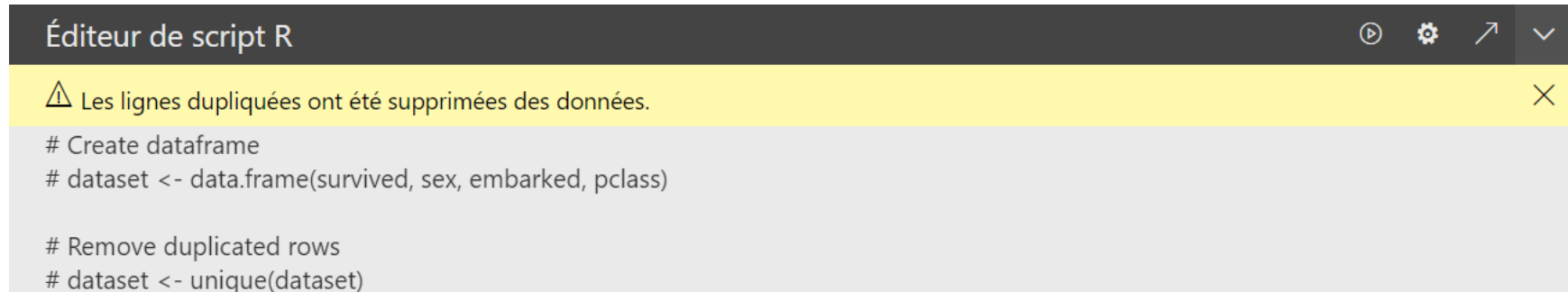
[Copier les détails](#)

Envoyer un smiley mécontent

Fermer

- ❖ Astuce 1 : observer les données sélectionnées dans un fichier temporaire
- ❖ Astuce 2 : afficher les valeurs d'indicateurs dans la légende du graphique
 - `plt.title(« La valeur de l'indicateur... »)`
- ❖ Astuce 3 : sélectionner dynamiquement les variables

« Les lignes dupliquées ont été supprimées des données »



The screenshot shows an R script editor window titled "Éditeur de script R". A yellow message bar at the top displays a warning icon and the text "Les lignes dupliquées ont été supprimées des données." with a close button. Below the message, the R script contains the following code:

```
# Create dataframe
# dataset <- data.frame(survived, sex, embarked, pclass)

# Remove duplicated rows
# dataset <- unique(dataset)
```

- ❖ Le jeu de données est dédoublonné automatiquement.
- ❖ Pour éviter cela, il faut créer une colonne d'index dans l'éditeur de requête (*Ajouter une colonne > Colonne d'index*) puis ajouter cette colonne dans la sélection.
- ❖ Si besoin, faire ensuite un regroupement des données avec la fonction Python `groupby()`

Sélectionner dynamiquement les champs du visuel

- ❖ Créer deux tables contenant la liste des variables
 - S'appuyer sur la table de résumé
 - `V1 = VALUES('df (summary)'[Variable])`
- ❖ Ajouter deux segments portant sur chaque de ces deux tables
- ❖ Créer deux mesures donnant la valeur sélectionnée dans les segments
 - `V1_selection = SELECTEDVALUE(V1[Variable])`
- ❖ Créer un visuel Python et ajouter tous les champs du dataset
 - Plus les deux mesures de sélection
- ❖ Coder...

```
import numpy as np
selection = np.unique(dataset[['V1_selection',
'V2_selection']].values)
selection = np.append(selection, 'Index')
df=dataset[selection]
```

Exploiter les « subplots »

- ❖ Objet figure : peut regrouper plusieurs graphiques
 - `pyplot.figure(1)` # ouvre la figure numéro 1
- ❖ Objet subplot : partage la figure et sélectionne une zone
 - `pyplot.subplot(2,1,1)` # partage la figure en 2 x 1 emplacements de graphes (2 lignes et 1 colonne) et sélectionne le 1er emplacement pour les instructions graphiques suivantes

```
pyplot.figure(1)
```

```
pyplot.subplot(1, 2, 1)
```

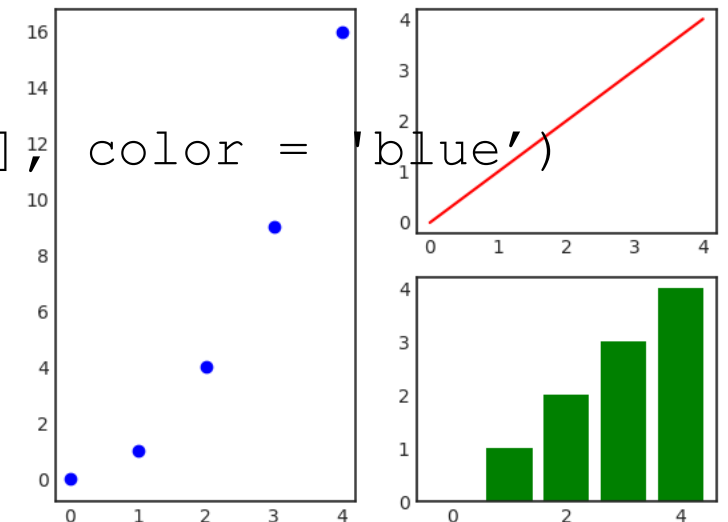
```
pyplot.scatter(range(5), [x ** 2 for x in range(5)], color = 'blue')
```

```
pyplot.subplot(2, 2, 2)
```

```
pyplot.plot(range(5), color = 'red')
```

```
pyplot.subplot(2, 2, 4)
```

```
pyplot.bar(range(5), range(5), color = 'green')
```



Mettre en place une approche analytique

A la recherche des relations entre les données

Un peu de vocabulaire...

- ❖ Colonne
 - Pour les utilisateurs d'Excel
- ❖ Champ
 - Pour les DBA
- ❖ Variable
 - Pour les statisticien.ne.s
- ❖ Deux grandes familles de variables statistiques :
 - Quantitative (numérique)
 - Qualitative (alphanumérique)
 - Donc aussi les codes postaux, numéros de sécu, etc.

Les bases de la Statistique descriptive

❖ Pour les variables quantitatives :

Raisonner centrage et dispersion

■ Centrage

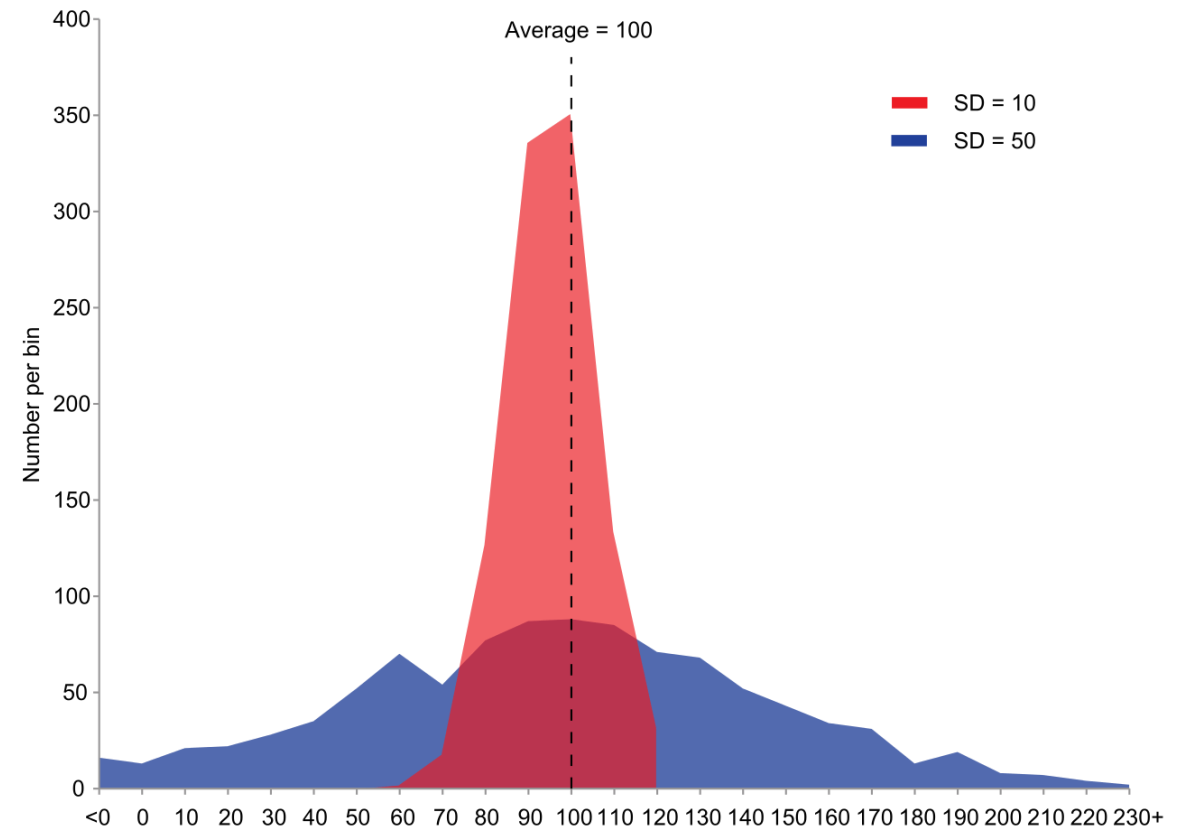
- Moyenne
- Médiane
- Mode

■ Dispersion

- Etendue = Max – Min
- Variance, écart-type
- Quartiles

❖ Comment les calculer rapidement dans Power BI ?

- Pas en DAX 😞



Exemple de deux échantillons ayant la même moyenne mais des écarts types différents

Approche statistique bivariable

- ❖ Croisement de deux variables
 - On recherche la **relation** statistique entre ces deux variables
 - 2 types donc 3 cas possibles

V1 / V2	Quantitative	Qualitative
Quantitative	Coefficient de corrélation	Test de moyennes
Qualitative	Test de moyennes	Test du Khi Deux

- ❖ Attention, les tests statistiques
 - demandent la vérification d'hypothèses
 - ne sont pas robustes sous grande volumétrie !

Le coefficient de corrélation LINEAIRE

❖ Entre deux variables numériques

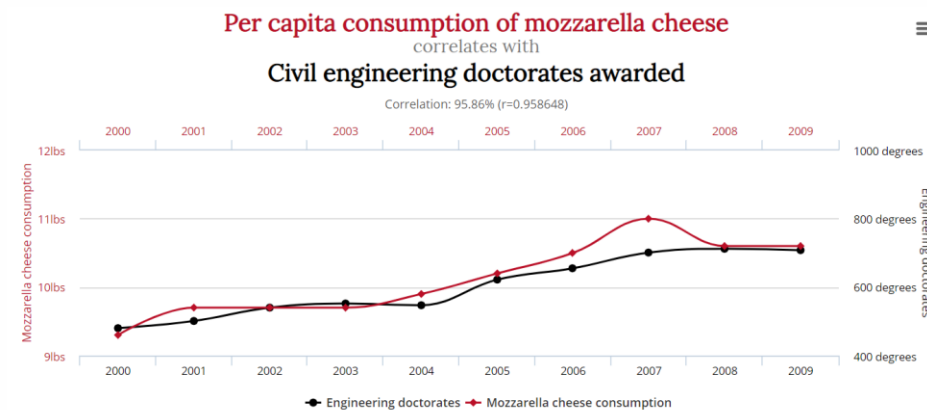
- Dite corrélation de Pearson

- Rapport (sans unité) de variances
- Entre -1 et 1

- Mais toutes les corrélations ne sont pas linéaires...

- On peut aussi travailler sur les rangs (Spearman)

- Certaines sont *fallacieuses*...



Data sources: U.S. Department of Agriculture and National Science Foundation

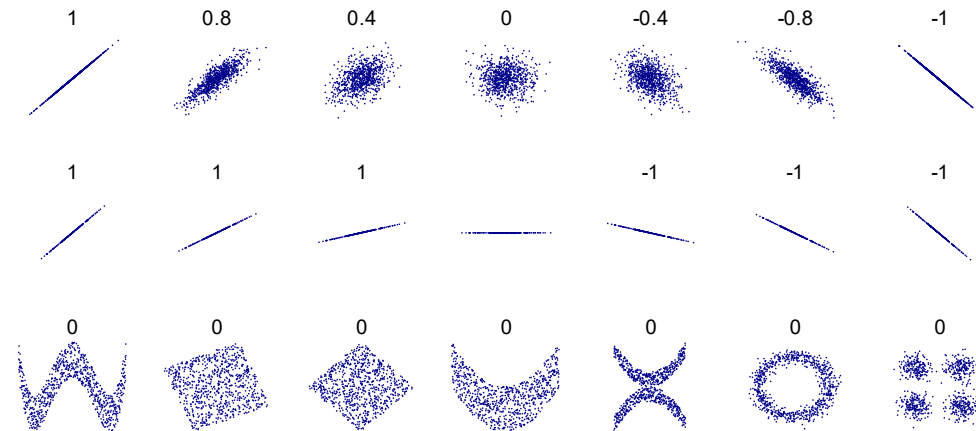
$$\hat{r}_p = \frac{\hat{\sigma}_{XY}}{\hat{\sigma}_X \hat{\sigma}_Y}$$

avec

$$\hat{\sigma}_{XY} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})$$

$$\hat{\sigma}_X = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \text{ et } \hat{\sigma}_Y = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \text{ et } \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$



Affichage des statistiques descriptives

❖ Résumé statistique :

- `import pandas as pd`
- `# conversion en dataframe à la volée`
- `output = pd.DataFrame(dataset.describe())`

❖ Repérage et correction des valeurs manquantes ou aberrantes

- `df['varQT'].fillna((df['varQT'].mean()), inplace=True)`
- `from scipy import stats`
- `df[(np.abs(stats.zscore(df)) < 3).all(axis=1)]`

DEMONSTRATION

Calcul d'une corrélation linéaire entre deux variables numériques

Calcul du coefficient de corrélation linéaire

❖ En DAX : *quick measure*

■ Par catégorie

➤ L'index pour travailler au détail des individus

■ « Figé » sur les variables sélectionnées

❖ En Python : librairie scipy

```
■ from scipy.stats.stats import pearsonr  
■ coef= pearsonr(df['V1'],df['V2'])
```

```
■ from scipy.stats.stats import spearmanr  
■ coef= spearmanr(df['V1'],df['V2'])
```

❖ Astuce : on affiche le résultat en légende d'un graphique

Mesures rapides

Calcul

Coefficient de corrélation ▼

Calcule le coefficient de corrélation entre deux valeurs sur une catégorie. Initialement suggéré par Daniil Maslyuk dans la galerie de mesures rapides.

[En savoir plus](#)

Catégorie ⓘ

Faites glisser ici les champs de données

Mesure X ⓘ

Faites glisser ici les champs de données

Mesure Y ⓘ

Faites glisser ici les champs de données

Quick measure : coefficient de corrélation

```
Correlation Coefficient := VAR Correlation_Table = FILTER ( ADDCOLUMNS (
VALUES ( {Category} ), "Value_X", CALCULATE ( {Measure X} ), "Value_Y",
CALCULATE ( {Measure Y} ) ), AND ( NOT ( ISBLANK ( [Value_X] ) ), NOT (
ISBLANK ( [Value_Y] ) ) ) ) VAR Count_Items = COUNTROWS (
Correlation_Table ) VAR Sum_X = SUMX ( Correlation_Table, [Value_X] ) VAR
Sum_X2 = SUMX ( Correlation_Table, [Value_X] ^ 2 ) VAR Sum_Y = SUMX (
Correlation_Table, [Value_Y] ) VAR Sum_Y2 = SUMX ( Correlation_Table,
[Value_Y] ^ 2 ) VAR Sum_XY = SUMX ( Correlation_Table, [Value_X] *
[Value_Y] ) VAR Pearson_Numerator = Count_Items * Sum_XY - Sum_X *
Sum_Y VAR Pearson_Denominator_X = Count_Items * Sum_X2 - Sum_X ^ 2
VAR Pearson_Denominator_Y = Count_Items * Sum_Y2 - Sum_Y ^ 2 VAR
Pearson_Denominator = SQRT ( Pearson_Denominator_X *
Pearson_Denominator_Y ) RETURN DIVIDE ( Pearson_Numerator,
Pearson_Denominator )
```

SYNTHESE

Est-il pertinent d'utiliser Python dans Power BI Desktop ?

Avantages de Python dans Power BI Desktop

- ❖ Python (beautifulsoup) plus fort pour le web scraping
- ❖ Plus intuitif que le langage M pour manipuler des datasets
 - Avec une approche statistique
 - ... si l'on connaît Python !
- ❖ Structures natives de boucle (for, while) et de condition (switch)
 - lambda functions de Python
 - En M, travailler sur des listes
- ❖ Pour la restitution visuelle :
 - Totale liberté de paramétrage des couleurs, police, libellés...
 - Utiliser les *subplots*

Limites de Python dans Power BI Desktop

- ❖ Installation des packages en local
 - A refaire sur chaque poste
- ❖ Pas de reconnaissance des types de données lors de l'import
- ❖ Exécution complète au moment de la création d'une étape de requête
 - Et non lorsqu'on clique sur « *close & apply* »
- ❖ Limite du nombre de points dans un nuage (150 000?)
- ❖ Fonctionnalités des scripts parfois limitées ou incompatibles
 - ex. : export graphique
- ❖ Noms des tables et des champs (bonnes pratiques à appliquer)
 - Pas d'accent, pas d'espace...

Limites de Python dans Power BI Service

- ❖ Python non supporté pour l'instant
 - A la visualisation
 - A la mise à jour



Les visuels Python sont pris en charge uniquement dans Power BI Desktop pour le moment.

① Vous ne pouvez pas planifier l'actualisation de ce jeu de données, car les sources de données suivantes ne prennent pas en charge l'actualisation pour l'instant :

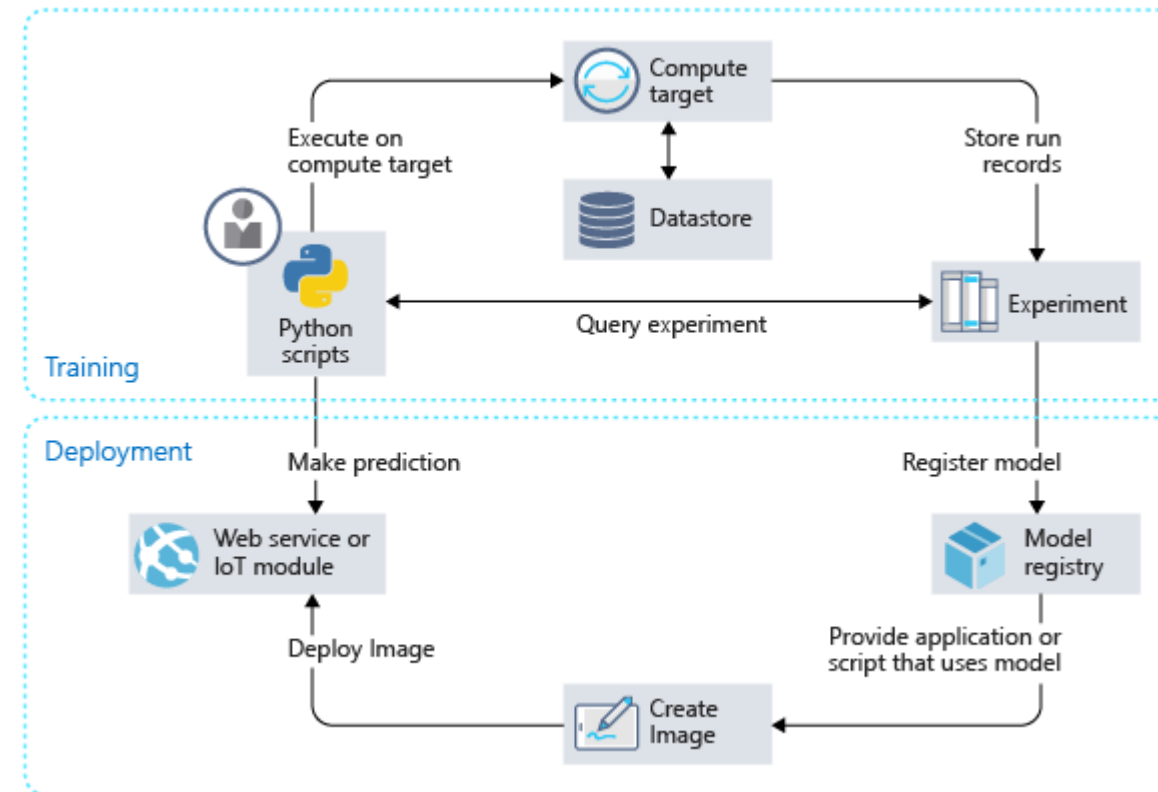
- Source de données pour df
- Source de données pour "df (changed type)"
- Source de données pour "df (summary)"

Découvrir les sources de données

Query contains unknown function name: Python.Execute.

Et pour l'analyse prédictive ?

- ❖ Machine Learning dans Power BI Service
 - meetup de décembre 2018
- ❖ Ou coder (en Python) :
 - Dans un **notebook**
 - En entrainant le modèle avec **Azure Databricks**
 - Et en l'exportant (ex.: pickle)
 - Pour l'enregistrer **sur Azure Machine Learning Service**
 - Puis le déployer sur des images **Docker** !



Merci !

Des questions ?

Place au networking