



Quand python, pandas, scikit-learn et jupyter
rencontrent Power BI

Le rêve du data scientist ?



Qui suis-je ?

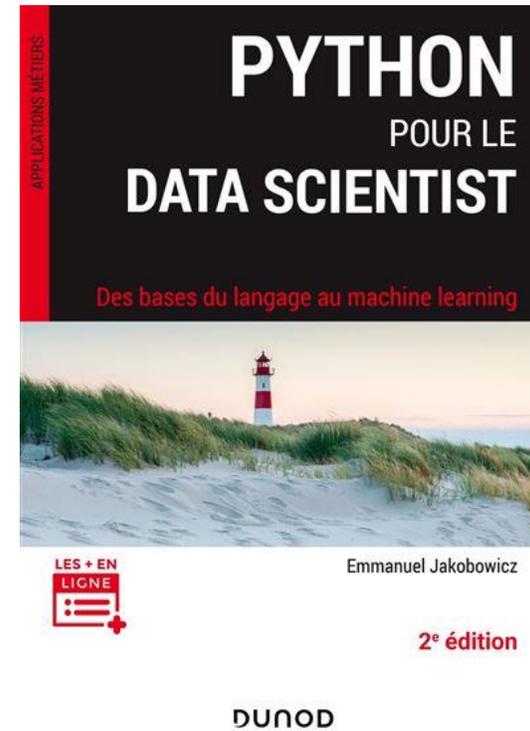


Emmanuel Jakobowicz

Fondateur et CEO de Stat4decision

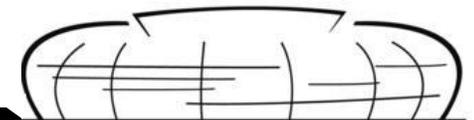
Lead Data Scientist, consultant, formateur, auteur, organisateur de meetups....

  @ejakobowicz



Stat4decision

Data science company





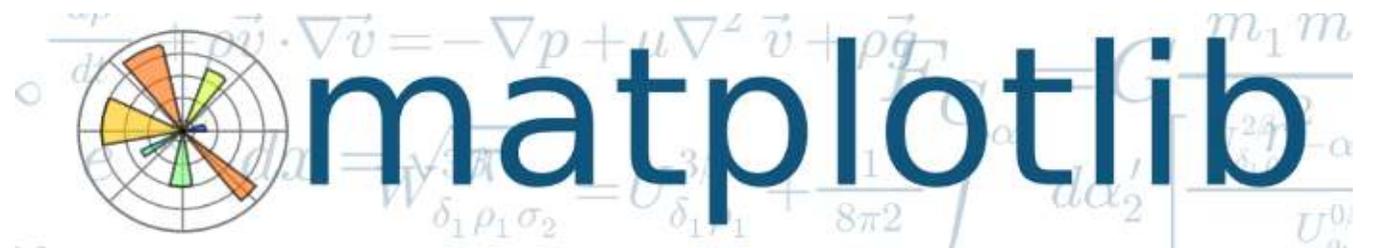
De quoi vais-je parler ?

D'un monde merveilleux peuplé d'animaux, de planètes et de...

Data Scientist

Si il existe encore

L'écosystème du Data Scientist





Et Power BI dans tout ça ?

Premier reflexe

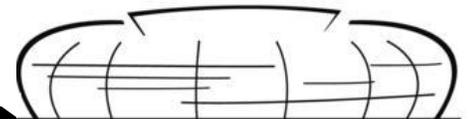


Et Power BI dans tout ça ?

Et si on pouvait l'utiliser
avec nos outils préférés ?

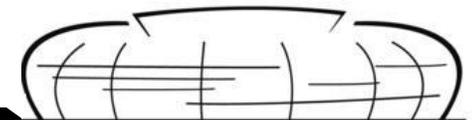


 Power BI



Comment convaincre le data scientist ?

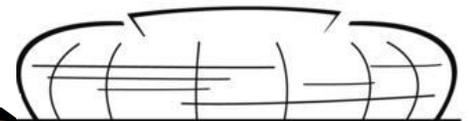
- Et si Power BI te faisait gagner du temps ?
- Et si Power BI te permettait de ne pas faire du Javascript ;-)
- Et si tes codes python fonctionnaient dans Power BI ?
- Et si tu pouvais construire des dashboards sans perdre trop de temps ?





Avant de commencer

- Pour faire du Python pour Power BI, de quoi ai-je besoin ?
 - Power BI Desktop
 - Python installé en local
 - *Facultatif* : un personal Gateway
- Si vous n'aimez pas l'éditeur de script de Power BI (il est moche...), paramétrez votre IDE favorite (VSCode dans mon cas)





Avant de commencer

- Pour utiliser python, on a besoin de :
 - Python en version au minimum 3.7
 - Un gestionnaire d'environnement du type conda (miniconda est possible)
 - Les packages de référence du data scientist installés en local
- Pour préparer Power BI, on a besoin de :
 - De paramétrer les options de scripts Python
 - D'autres choses, mais on verra plus tard !

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:\Users\s4d-asus-14\anaconda3-bis

[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](#)



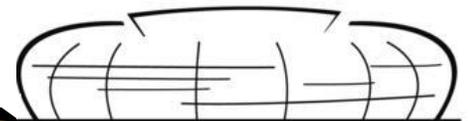
NIVEAU 1



Que fait le data scientist 90% du temps avec python ?

Il manipule des données pour les préparer

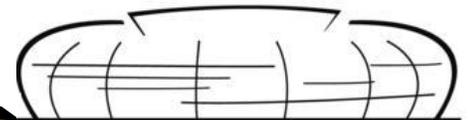
Avec quoi ?





1^{ère} problématique- la préparation des données

- L'outil de référence : pandas
- Force :
 - super facile !
 - Python
 - Super concis
- La question métier :
 - J'ai une base de données et j'ai construit un traitement en python pour dédoublonner et faire une grosse préparation. Que faire pour intégrer le résultat dans mon dashboard Power BI ?
- La réponse :
 - Option 1 : traduis-moi ça !!!
 - Option 2 : stocke ta base de données transformée avant de la lire avec Power BI
 - Option 3 : lance le traitement depuis un script power BI





DEMO



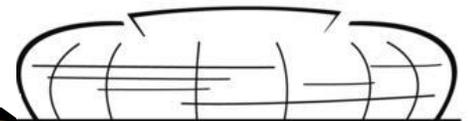
NIVEAU 2



Autre tâche du data scientist ?

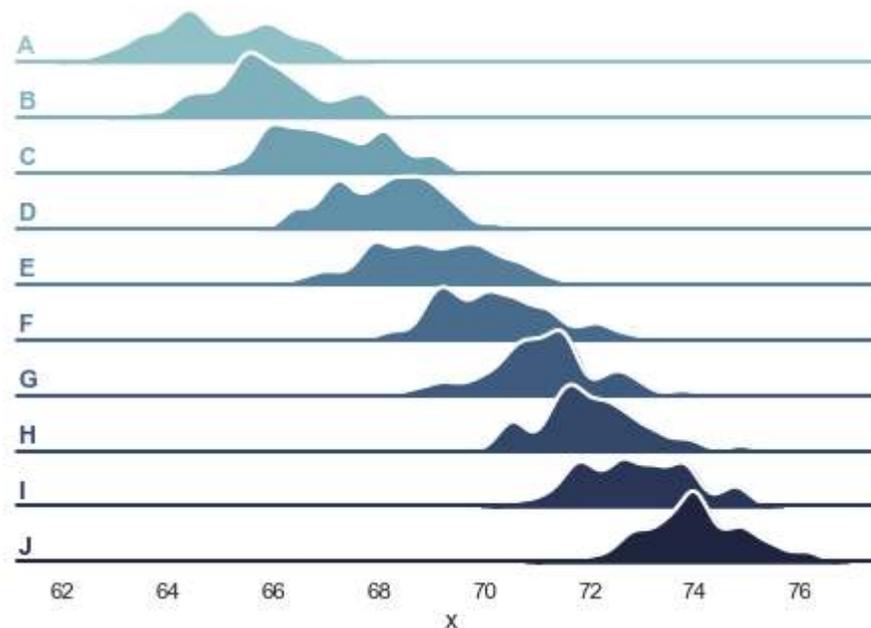
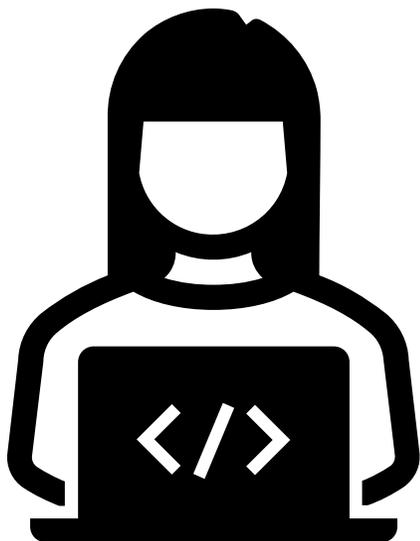
Créer des représentations impactantes

Avec quoi ?



2^{ème} question – niveau 2

- J'ai construit un beau graphique avec mon package préféré en python et je veux le voir dans mon dashboard Power BI
 - Il y a rien d'équivalent dans les visuels Power BI
 - Je veux qu'il soit aussi beau





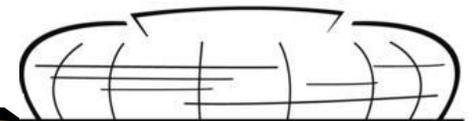
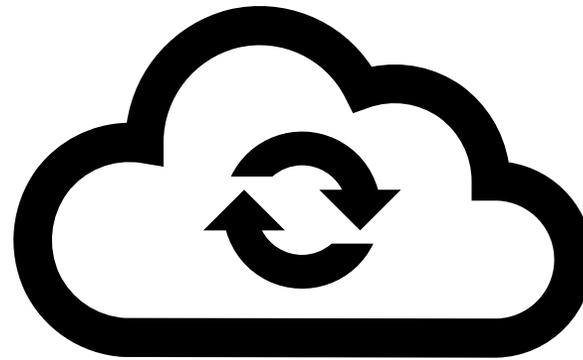
DEMO





2^{ème} question – niveau 2+

- Et si je veux passer en mode Power BI Service ?





DEMO





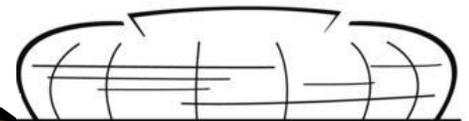
NIVEAU 3



Obsession du data scientist ?

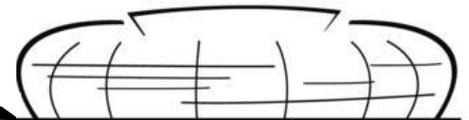
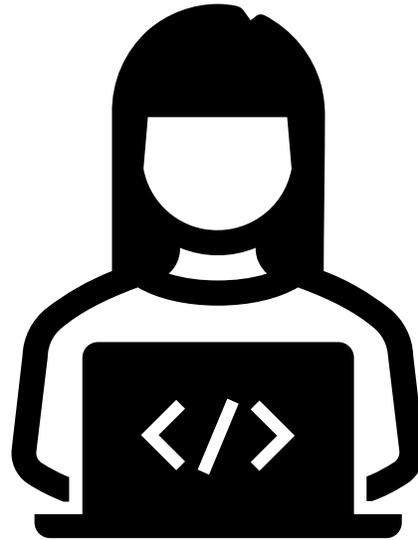
Faire du ML !!!

Avec quoi ?



3^{ème} question

- Je veux construire un modèle de machine learning pour segmenter mes clients afin de mieux piloter mon activité
 - J'aime développer dans un Notebook ou dans VSCode
 - Je veux pouvoir voir le résultat dans mon dashboard Power BI





DEMO

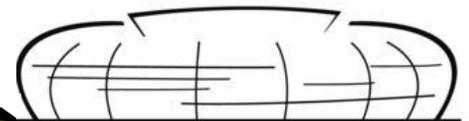




Outil du data scientist ?

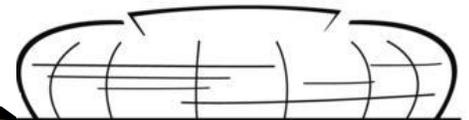
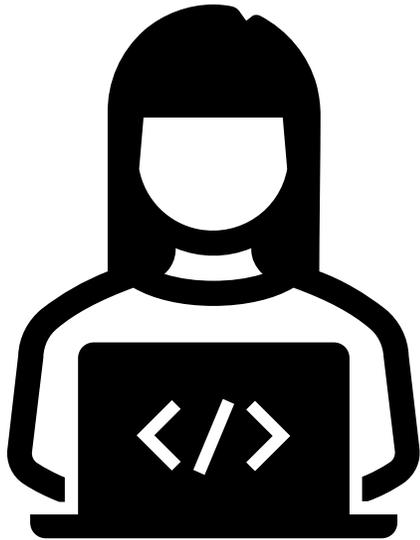
Travailler dans un Notebook afin de visualiser les résultats de son travail

Avec quoi ?



3^{ème} question – niveau 2

- J'adore les Jupyter Notebook pour partager mes travaux
- J'adore mes dashboards Power BI et je ne veux pas me lancer dans les packages de dashboarding (Dash, Bokeh...)
- Comment j'intègre mon dashboard Power BI dans mon Jupyter Notebook ?





1. Install the Power BI Client package in your Jupyter Notebook environment

Here's how



Install Power BI Jupyter Client from PyPI



Open a Jupyter notebook in your browser



Import the Power BI Client into Jupyter notebook



2. Embed your Power BI Embedded analytics report

Here's how



Authenticate to Power BI



Get report and workspace ID



(Optional) Set report event handlers



Embed the report in an output cell



3. Interact with the report

Here's how



Apply bookmarks and filters to the embedded report



Get the page and visual name



Export summarized data from a report visual



Explore Power BI data in-depth in Jupyter notebook



DEMO



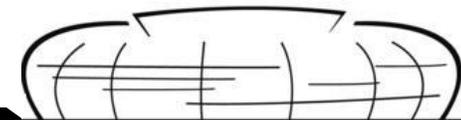
Alors ?





Attention quand même, tout n'est pas parfait ! Les limites

- Power BI Service est trop limité pour aller plus loin (manque de package et de puissance)
- Possibilité d'utiliser le personal Gateway pour faire tourner des scripts python mais pas le enterprise Gateway...
- Limitation dans la taille des données : les données utilisées par le visuel Python pour le traçage sont limitées à 150 000 lignes ou 250Mo
- Résolution : tous les visuels Python sont affichés dans une résolution de 72 ppp.
- Temps de calcul : si le calcul d'un visuel Python prend plus de cinq minutes, le délai d'exécution est dépassé, ce qui provoque une erreur.





Python, c'est facile et ça marche bien !!!
Essayez-le !
Ca n'est pas que pour l'IA !!!



Stat4decision
Data science company

Merci !

Laissez nous votre avis sur
cette session :



P  **OWER** 365