

RER

Python : Fichiers - Fonctions - Modules

Contexte : Maitriser l'utilisation des modules, des fonctions originales et des fichiers (lecture, écriture, copie, suppression, renommage).

Mots clés :

- **Module** : une bibliothèque ou une librairie.
- **Librairie / bibliothèque** : un fichier contenant des fonctions, des variables, des constantes, des objets, des classes, des méthodes mise à disposition afin d'être utilisé sans avoir de les réécrire.
- **Fichier** :
 - **.csv : Comma-Separated Values** (valeurs séparées par des virgules) : contient des données tabulaires, en général, séparés les uns des autres par un délimiteur : ",", mais on peut utiliser des ";"
 - **.xls** : fichier Excel, xlsx (consomme moins de mémoire), créé par Microsoft
 - **.xml** : Extensible Markup Language (XML), méta langage de balises
 - **.json** : JavaScript Object Notation, stock des données sous forme d'un dictionnaire (version développée de .xml)
- **Random** : est un module natif de Python, qui permet de générer des nombres aléatoires et de faire une sélection aléatoire dans une liste.
- **Math** : un module natif de Python, qui fournit l'accès aux fonctions mathématiques définies par la norme C. (logarithme, sin., cos., fonctions trigonométriques, etc.)
- **NumPy** : Numerical Python est une bibliothèque pour langage de programmation Python, destinée à manipuler des matrices ou tableaux multidimensionnels ainsi que des fonctions mathématiques opérant sur ces tableaux.
- **Pandas** : Python Data Analysis est une bibliothèque de Python écrite pour le langage de programmation Python permettant la manipulation et l'analyse des données. Elle propose en particulier des structures de données et des opérations de manipulation de tableaux numériques et de séries temporelles.
- **Matplotlib** : est une bibliothèque du langage de programmation Python destinée à tracer et visualiser des données sous forme de graphiques, diagrammes, nuages de points, histogrammes etc. .

- **DataFrames** : est une structure de données qui organise les données dans un tableau bidimensionnel de lignes et de colonnes, un peu comme une feuille de calcul. (ex. : tableau NumPy 1D + axe d'index). Les DataFrames sont l'une des structures de données les plus couramment utilisées dans l'analyse de données moderne, car elles constituent un moyen flexible et intuitif de stocker et de travailler avec des données. Une association des séries.
- **Package** : est un regroupement des modules.
- **Fonction Built-In** : les fonctions natives de Python.
- **Fonction originale** : les fonctions créées par l'utilisateur.
- **Import** : on importe le module en entier,
- **From import** : si on n'a pas besoin d'un module en entier
- **as** : pour renommer ce qu'on a importé (méthode...)

Exemple : **from** Mon_module **import** Ma_méthode **as** mt

Problématique(s) :

- Comment peut-on manipuler les différents types de fichiers avec les modules.
- Comment les modules facilitent les instructions dans le langage de programmation Python.

Hypothèses :

1. Les modules Random, Math, NumPy, Pandas, Matplotlib proposent les mêmes fonctions (Adeline) **Faux**
2. Les modules peuvent contenir des nouveaux objets (Adrien) **Vrai**
3. Pandas est le seul module pour Python permettant de traiter les fichiers (modifier, créer, renommer, supprimer) (Loïc) **Faux**
4. Les modules sont des fonctions importés (Solenn) **Faux**
5. Le module Pandas ne permet pas de traiter tous les types de fichiers (Axel) **Vrai**
6. Si on supprime un fichier importé par le module – le programme ne fonctionne plus (Briand) **Faux**
7. Un module est un objet (Solenn) **Faux**
8. L'ouverture d'un fichier dans Python nécessite le concours du contexte-manager. (Etienne) **Faux**
9. Une fonction est une méthode (Aude) **Vrai**
10. Pour utiliser un module – il faut l'installer d'abord. (Seydou) : **Dépend du module : natif ou pas**
11. Il y a des librairies qui permettent de créer et de manipuler les DataFrames (Nicolas) **Vrai**
12. Les modules mises à disposition font que Python est un langage par excellence de l'analyse des données. (Adeline) **Vrai**
13. Les opérations : lecture, écriture, copie, suppression, renommage sont les seules opérations possibles (Jean Paul) **Faux**
14. Les modules et les librairies est un gain de temps en programmation (Tetyana) **Vrai**

Plan d'action :

- Explorer les ressources

- Définir les mots-clefs
- Répondre aux hypothèses
- Utiliser Jupyter Notebook pour tester toutes les notions abordées
- Répondre à la problématique
- Restitution du RER
- Livrable