

# Retour sur Python et introduction à SciPy

Atelier théorique 2





# Python

Vous devriez déjà avoir une bonne base avec la programmation Python suite aux cours sur **CodeAcademy**. On va faire une révision en accéléré des concepts de bases importants applicables au domaine de l'apprentissage automatique.



**EXERCICE** - Révision de Python



# Qu'est-ce que SciPy?

C'est un écosystème composé d'outils computationnels pour les domaines des mathématiques, sciences et ingénierie.

À l'intérieur de l'écosystème, on retrouve une collection de logiciels "open-source" entretenue par la communauté scientifique elle-même.

## **Open-source**

Développer par la communauté et généralement disponible gratuitement.



# Pourquoi SciPy?

1. Pour être efficace en apprentissage machine avec Python, il est primordial de bien connaître et maîtriser l'écosystème SciPy.
2. Les bibliothèques de SciPy sont utilisées **tous les jours** par les analystes et scientifiques de données à travers le monde!!
3. Pour plus d'informations sur SciPy : <https://scipy.org/index.html>

## **Qu'est-ce qu'une bibliothèque Python**

Une bibliothèque Python (nommée “library” ou “package” en anglais) est un ensemble de codes réutilisables qui permet d'effectuer différentes manipulations ou analyses efficacement et rapidement.

# Librairies principales de SciPy



Librairies importantes





# Numpy



**Utilisation :** Manipulation de données numériques

**Avantages :**

- Facilité d'utilisation
- Rapidité d'exécution
- Formats compatibles avec les modèles d'apprentissage automatique

**Lien web :** <https://numpy.org/>



**EXERCICE** - Cours accéléré sur NumPy



# Pandas



**Utilisation :** Manipulation de jeux de données tabulaires.

**Avantages :**

- Structuration et organisation des données
- Modification, restructuration et combinaison de jeux de données
- Création de visualisations de bases

**Lien web :** <https://pandas.pydata.org/>



**EXERCICE** - Cours accéléré sur Pandas



# Matplotlib



**Utilisation :** Production de visualisations simples et complexes

**Avantages :**

- Grand inventaire de visualisation pouvant être utilisé immédiatement (“out-of-the-box”) directement dans la librairie
- Plusieurs attributs disponibles pour personnaliser les figures
- Flexibilité pour créer des visualisations très complexes
- Possibilité de créer des visualisations interactives et animées

**Lien web :** <https://matplotlib.org/stable/index.html>



**EXERCICE** - Cours accéléré sur Matplotlib





# Ressource = Stack Overflow

Stack Overflow est une plateforme en ligne qui permet de poser des questions et d'obtenir des réponses sur des sujets techniques variés, dont la programmation avec Python.

Pourquoi utiliser Stack Overflow?

- Entraide entre les programmeurs
- Plusieurs exemples sont offerts
- Permet de trouver des solutions partielles de code

Lien : <https://stackoverflow.com/>

