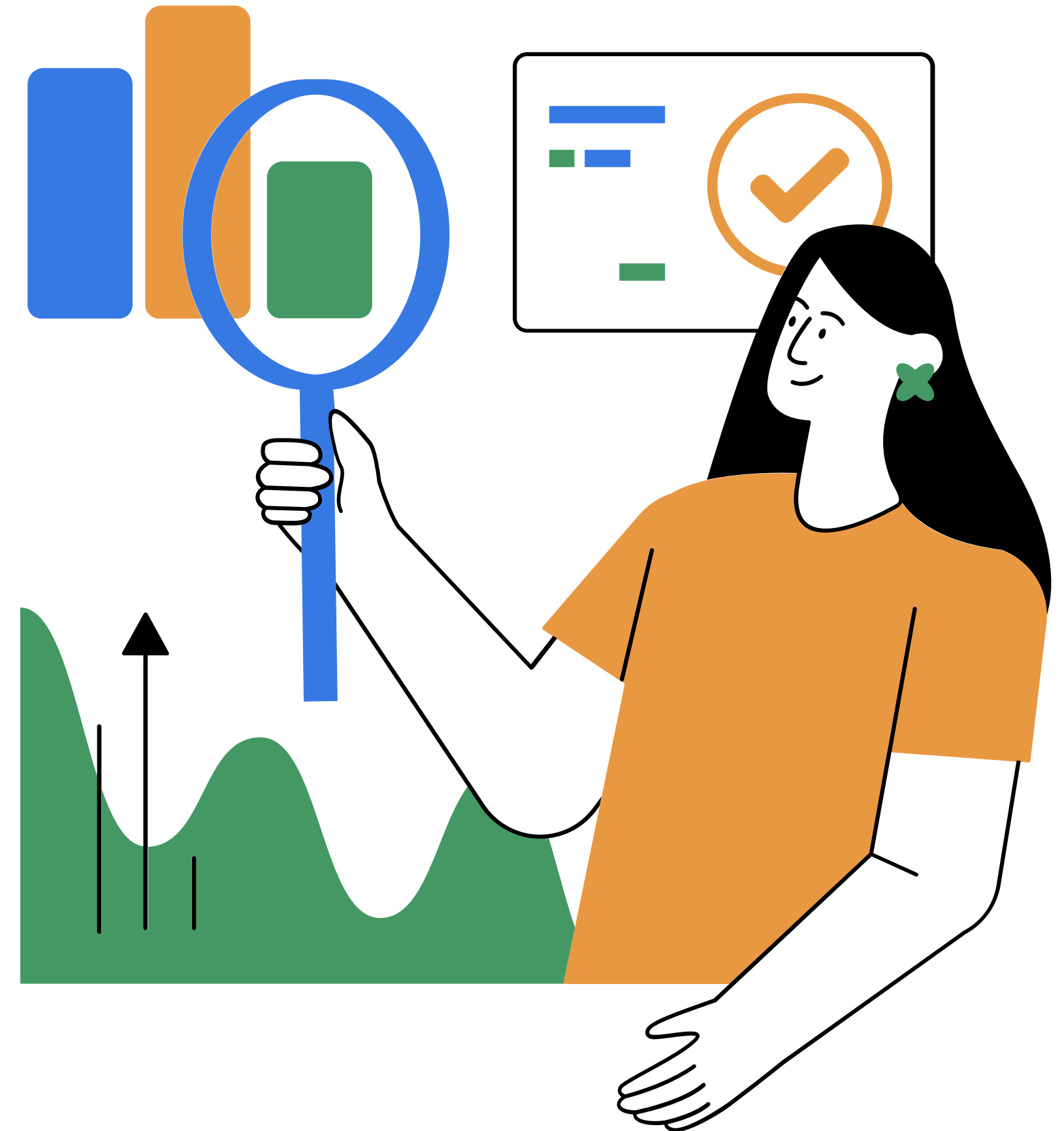


# Initiation a la Data science et Machine Learning avec Python

**Onesime Mbulayi**

PhD Student  
University of Padova | Italy  
Data Scientist & Bioinformatics

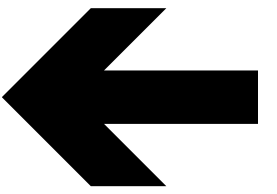


# L'intelligence Artificielle ←

L'intelligence artificielle (IA) est un domaine de l'informatique qui vise à créer des systèmes capables de réaliser des tâches qui, normalement, nécessiteraient l'intelligence humaine.



# AI vs Robotique



l'IA et la robotique sont souvent complémentaire. l'IA apporte les capacités cognitives nécessaires pour que les robots puissent fonctionner de manière autonome et intelligente, tandis que la robotique fournit le cadre physique pour que l'IA puisse interagir avec le monde réel.

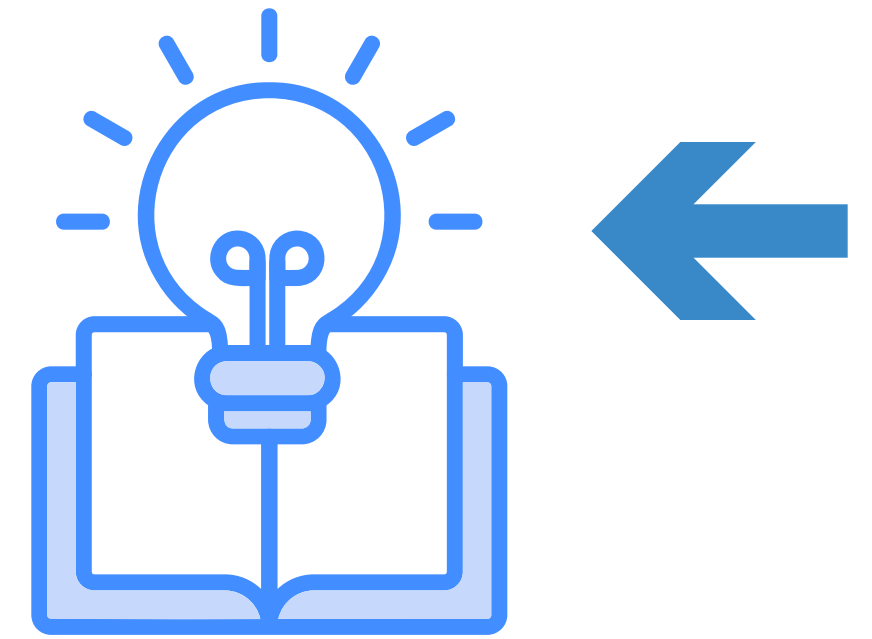
# Qu'est-ce la Data Science



La data science est un domaine interdisciplinaire qui utilise des méthodes scientifiques, des algorithmes, des processus et des systèmes pour extraire des connaissances et des insights à partir de données structurées ou non structurées.

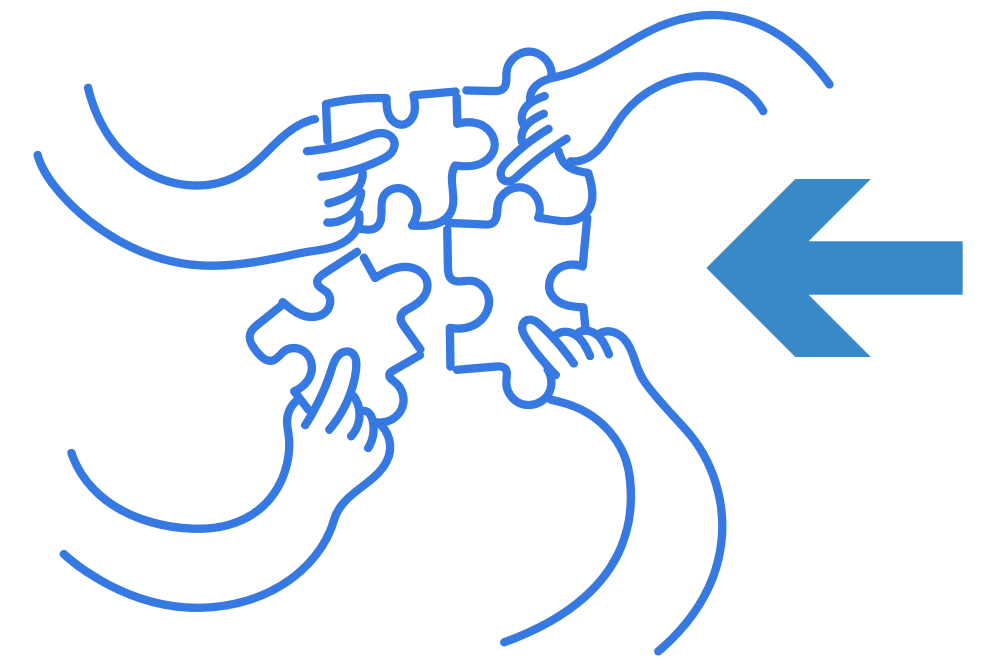
Elle englobe **la collecte, le nettoyage, l'analyse, la visualisation et l'interprétation des données** pour résoudre des **problèmes complexes ou prendre des décisions éclairées.**

# Qu'est-ce Machine Learning



Le machine learning est une sous-discipline de l'intelligence artificielle (IA) qui se concentre sur la création d'algorithmes et de modèles permettant aux machines d'apprendre à partir de données et de faire des prédictions ou des décisions sans être explicitement programmées pour chaque tâche.

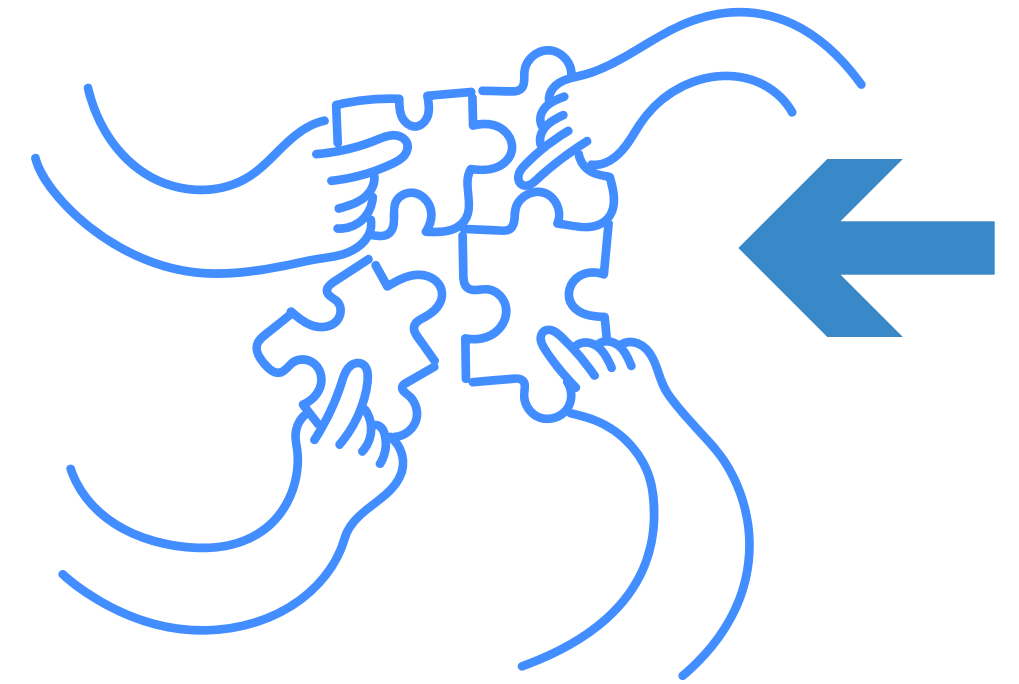
# Méthodes et techniques pour la Data Science



La data science utilise une variété de techniques, notamment :

1. La statistique descriptive et inférentielle.
2. La visualisation de données (tableaux, graphiques, etc.).
3. Le nettoyage et la préparation des données (data wrangling).
4. L'analyse exploratoire des données (EDA).
5. L'utilisation d'algorithmes de machine learning pour l'analyse prédictive.

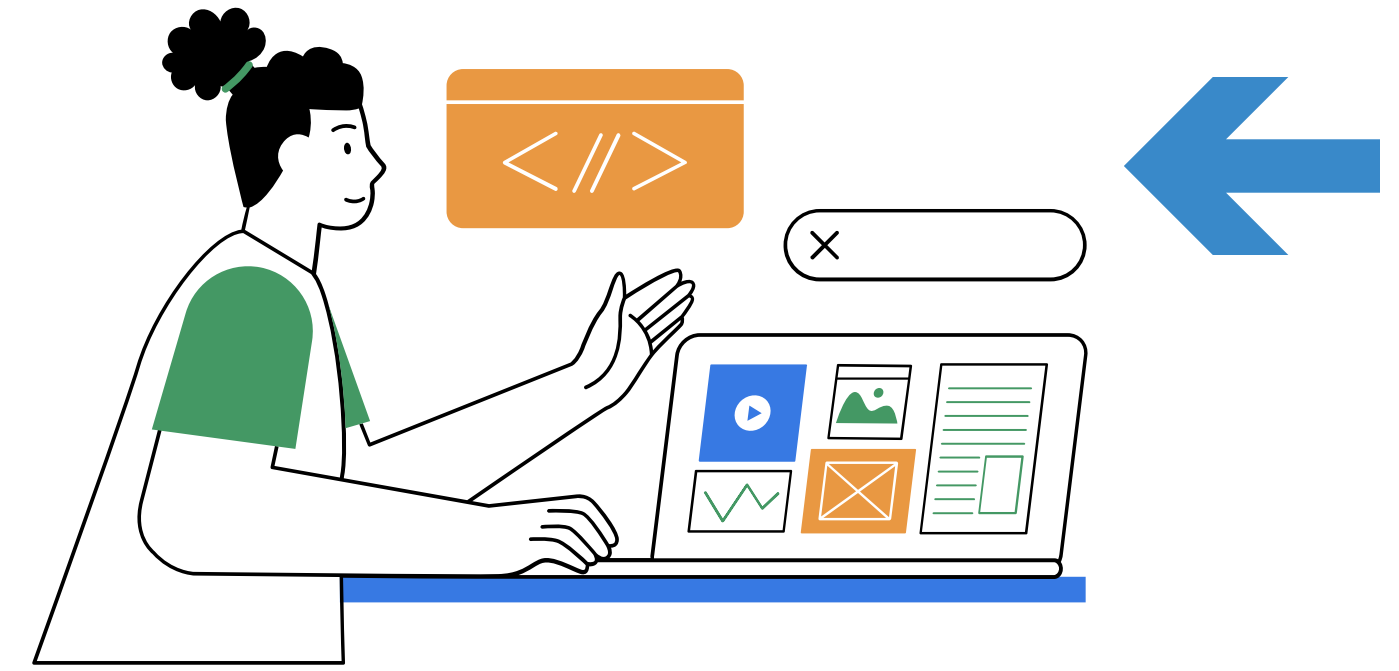
# Méthodes et techniques pour le Machine Learning



Le machine learning se concentre sur des techniques spécifiques pour entraîner des modèles, telles que :

1. L'apprentissage supervisé (classification, régression).
2. L'apprentissage non supervisé (clustering, réduction de dimensionnalité).
3. L'apprentissage par renforcement (reinforcement learning).
4. Les réseaux de neurones et le deep learning

# Outils et technologies pour la Data Science

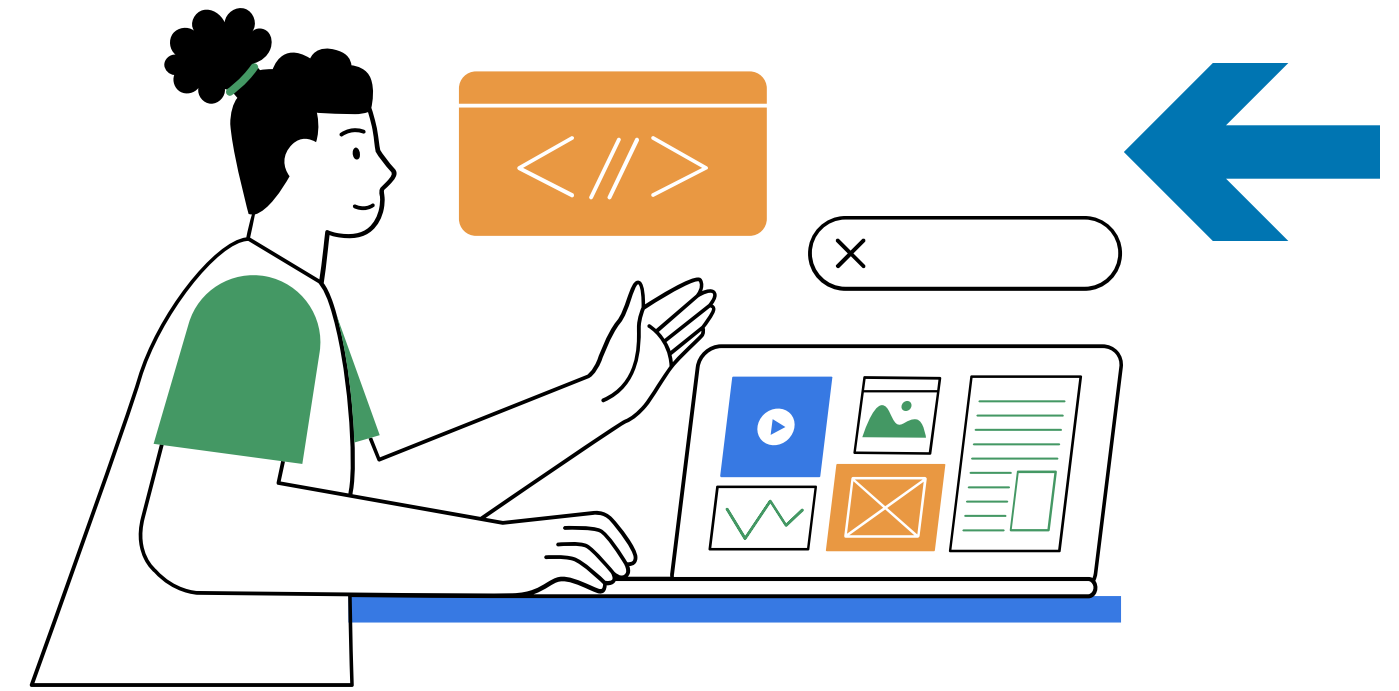


Les data scientists utilisent des outils comme :

1. **Python**, R, SQL pour l'analyse de données.
2. Tableau, Power BI pour la visualisation.
3. **Pandas, NumPy, Matplotlib** pour la manipulation et l'analyse des données.
4. Hadoop, Spark pour le traitement de grandes quantités de données



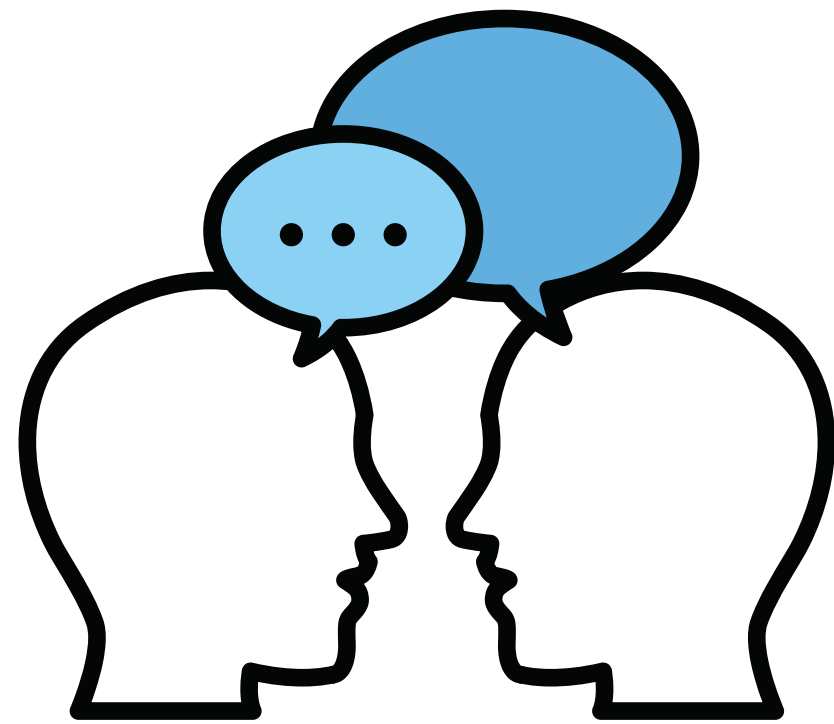
# Outils et technologies pour le Machine Learning



Les spécialistes du machine learning utilisent des outils comme :

1. TensorFlow, PyTorch pour le deep learning.
2. **Scikit-learn** pour les algorithmes de ML classiques.
3. Keras, XGBoost pour des modèles spécifiques.

# Data Science vs Machine Learning



- La **data science** se concentre sur l'extraction de sens et de valeur à partir des données, en utilisant une variété de techniques, y compris le machine learning.
- Le **machine learning** est une discipline technique qui se concentre sur la création de modèles prédictifs et l'apprentissage automatique à partir de données.

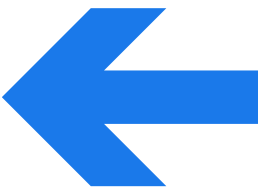
**En d'autres termes, le machine learning est une composante de la data science, mais la data science va bien au-delà du machine learning en intégrant d'autres aspects liés à la gestion et à l'analyse des données.**

# Objectifs du Cours



1. Acquérir une compréhension approfondie des concepts de la DS.
2. Maîtriser les techniques de manipulation de données avec Pandas
3. Apprendre à visualiser les données de manière efficace avec ,matplotlib et seaborn
4. Comprendre les fondamentaux de ML avec Scikit-Learn
5. Concevoir les projets avec les données réelles

# Compétences du Cours



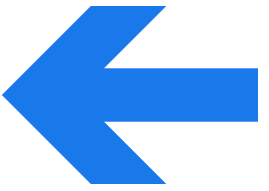
A la fin de la formation, les participants seront capable de :

1. Effectuer les operation de base sur des dataset avec  
Pandas
2. Visualiser et interpréter les données
3. Appliquer les technques ML simple pour resoudre des  
problèmes predictifs.
4. Travailler de maniere autonome sur des projets de  
Data Science en utilisant Python.



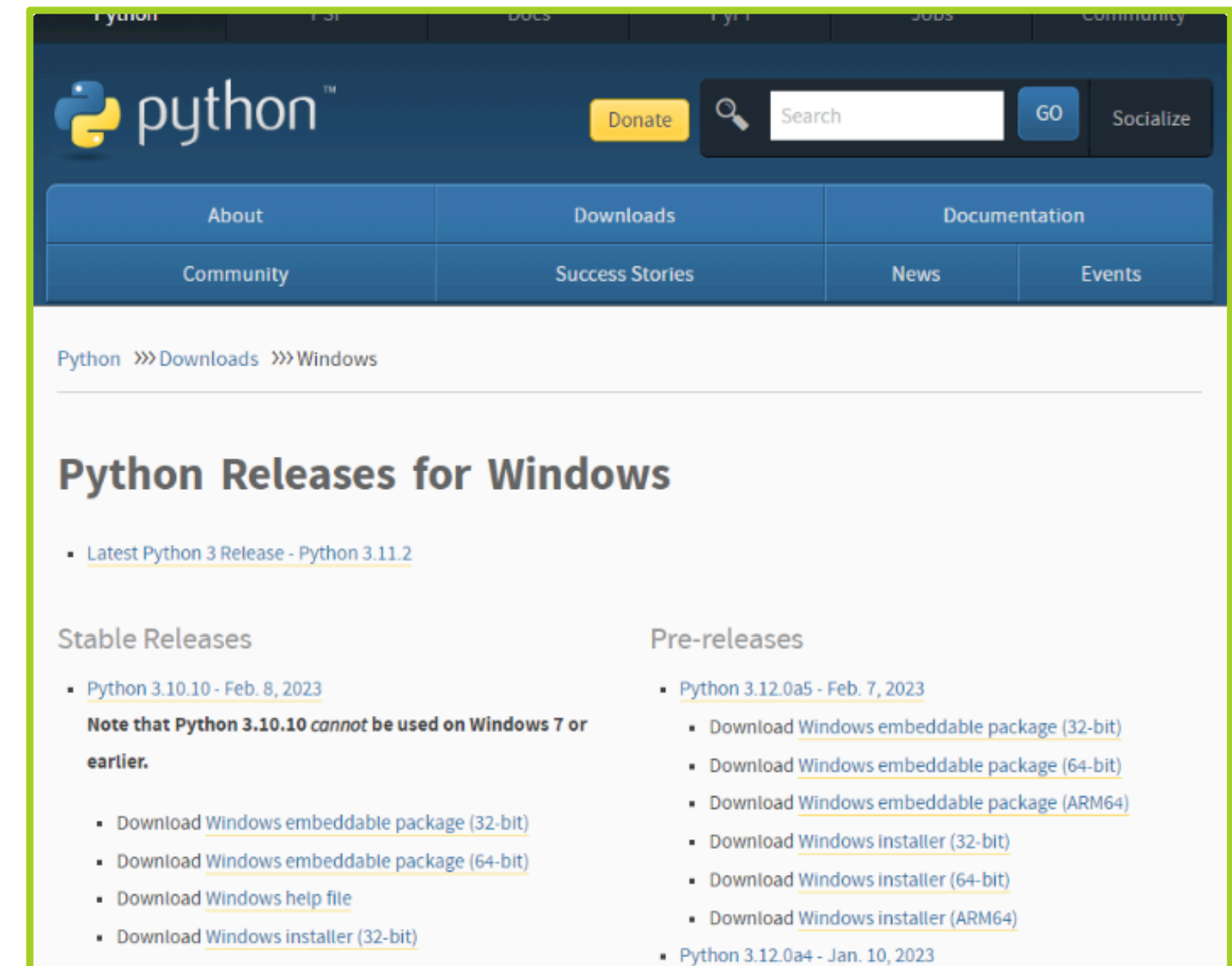
# Environnements & Outils

# Environnement: Installation Python

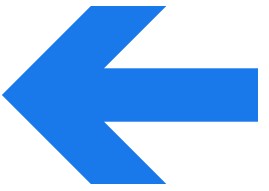


## Étape 1 : Téléchargement du programme d'installation Python

1. Accéder à l'officiel Python download page for Windows.
2. Trouvez une version stable de Python
3. Cliquez sur le lien approprié à votre système pour télécharger le fichier exécutable :  
programme d'installation Windows (64 bits) ou  
programme d'installation Windows (32 bits).

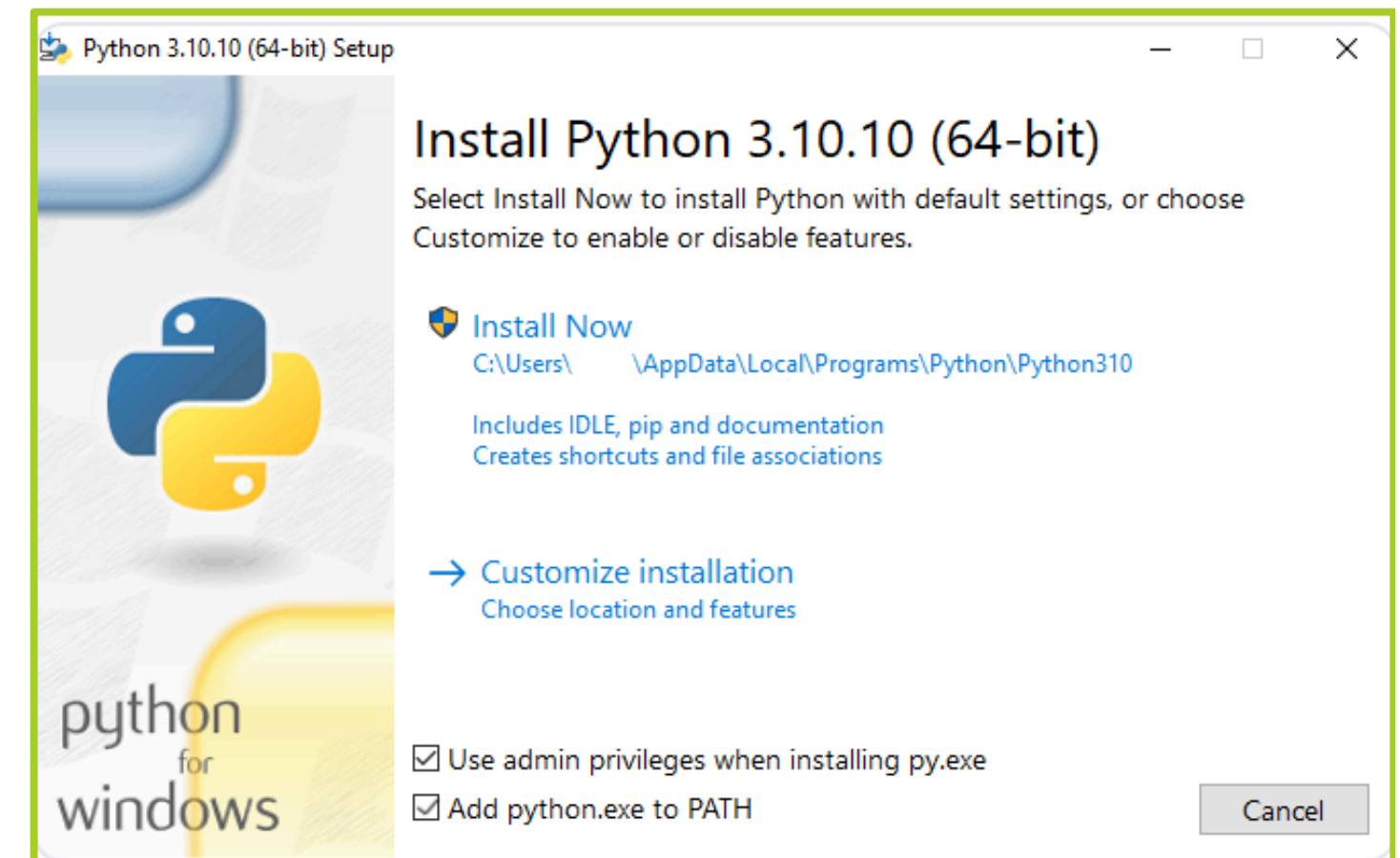


# Environnement: Installation Python



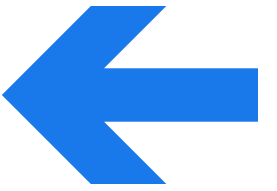
## Étape 2: Exécution du programme d'installation

1. Une fois le programme d'installation téléchargé, double-cliquez sur le fichier **.exe**, par exemple **python-3.10.10-amd64.exe**, pour exécuter le programme d'installation Python.
2. Cochez la case Installer le lanceur pour tous les utilisateurs, qui permet à tous les utilisateurs de l'ordinateur d'accéder à l'application de lancement Python.
3. Cochez la case Ajouter python.exe au PATH, qui permet aux utilisateurs de lancer Python à partir de la ligne de commande.

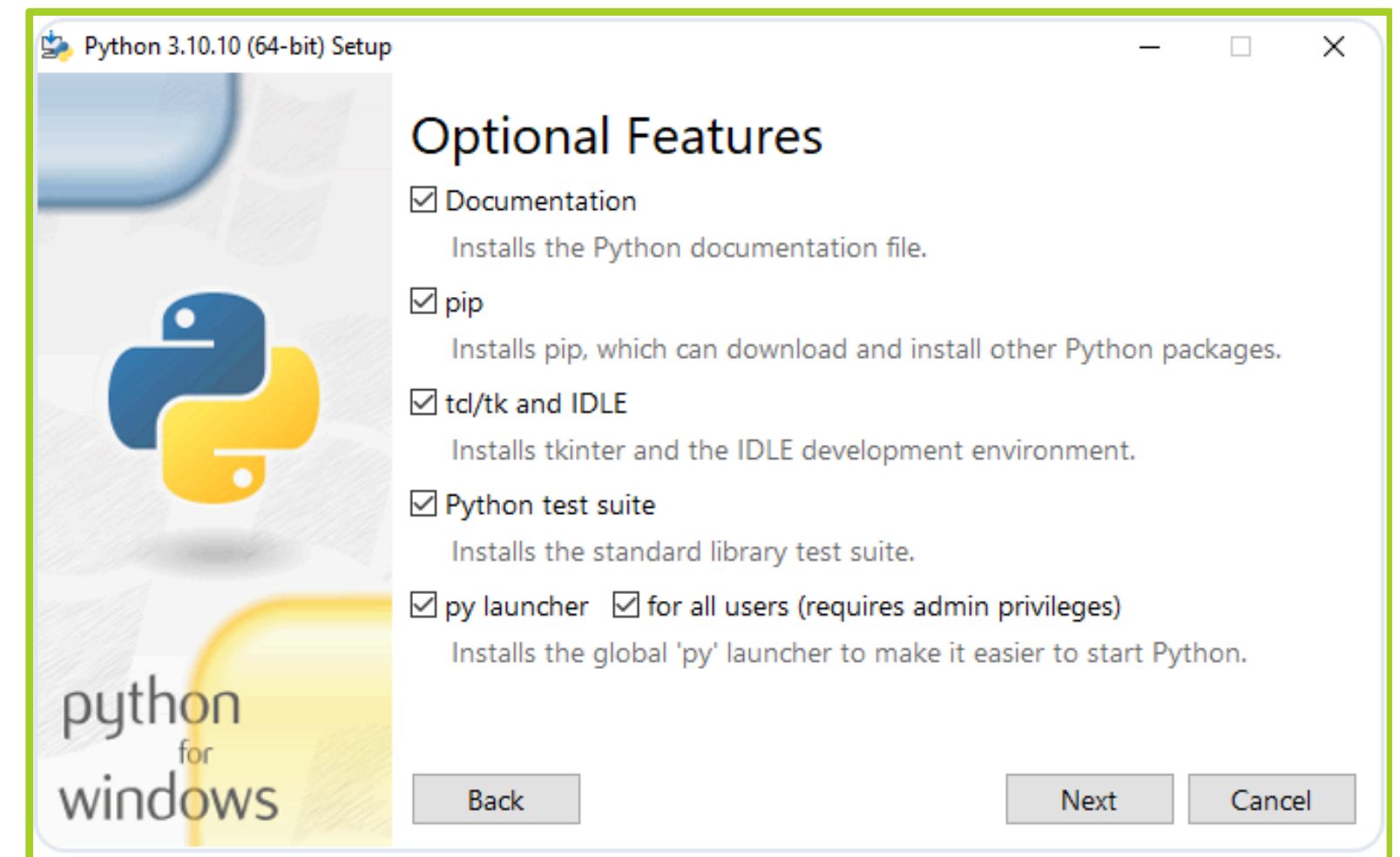




# Environnement: Installation Python

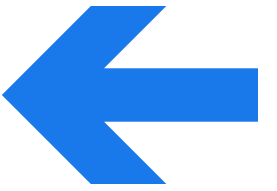


1. Si vous débutez avec Python et que vous souhaitez l'installer avec les fonctionnalités par défaut décrites dans la boîte de dialogue, cliquez sur Installer maintenant et passez à l'étape 4. Pour installer d'autres fonctionnalités facultatives et avancées, cliquez sur Personnaliser l'installation et continuez.
2. Les fonctionnalités facultatives incluent des outils et des ressources communs à Python et vous pouvez les installer tous, même si vous ne prévoyez pas de les utiliser.





# Environnement: Installation Python

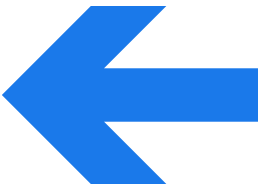


Sélectionnez les options qui correspondent à vos besoins :

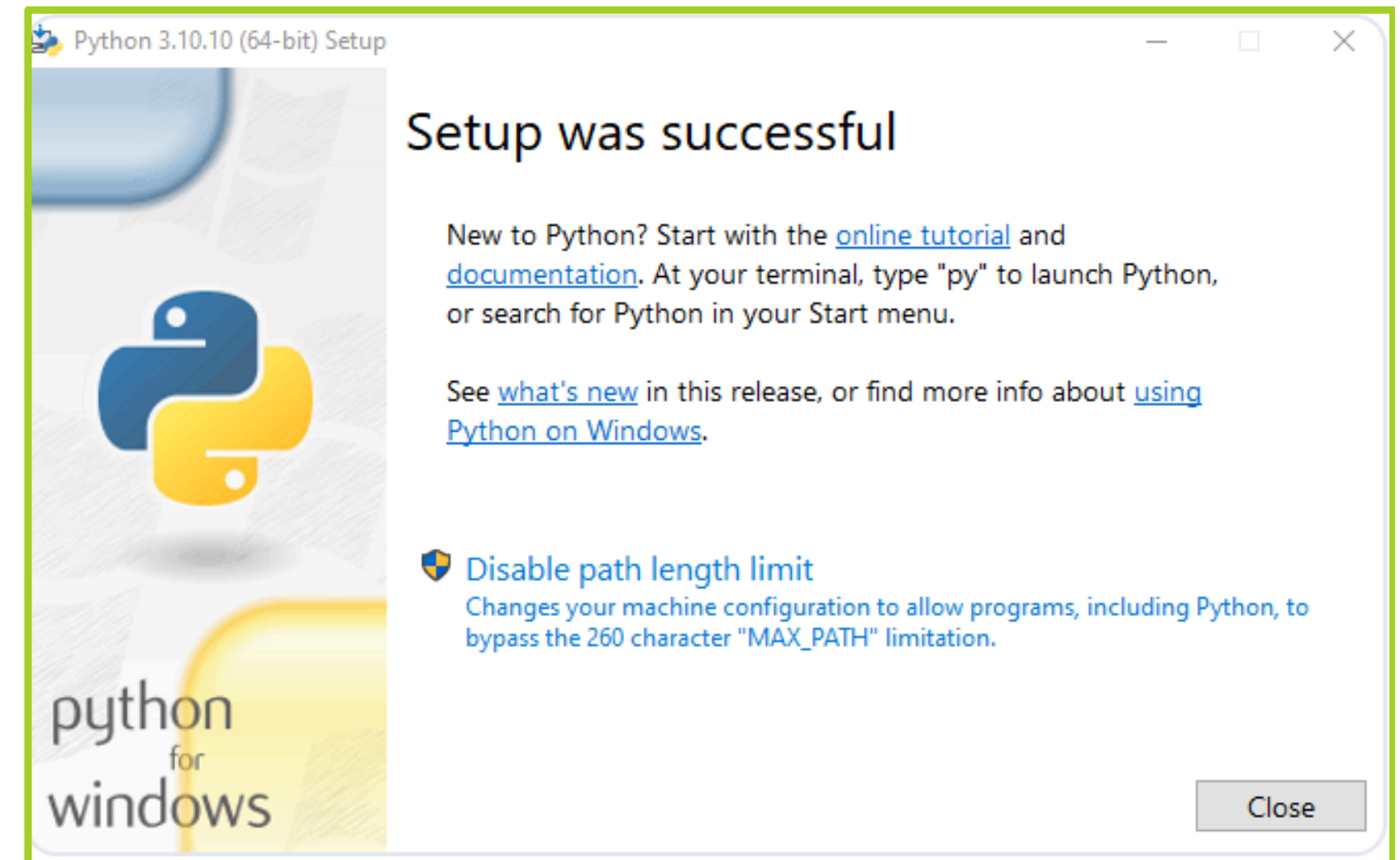
- **Installer pour tous les utilisateurs** : recommandé si vous n'êtes pas le seul utilisateur sur cet ordinateur
- **Associer les fichiers à Python** : recommandé, car cette option associe tous les types de fichiers Python au lanceur ou à l'éditeur
- **Créer des raccourcis pour les applications installées** : recommandé pour activer les raccourcis pour les applications Python
- **Ajouter Python aux variables d'environnement** : recommandé pour activer le lancement de Python
- **Précompiler la bibliothèque standard** : non requis, cela pourrait ralentir l'installation
- **Télécharger les symboles de débogage et télécharger les binaires de débogage** : recommandé uniquement si vous prévoyez de créer des extensions C ou C++

Notez le répertoire d'installation de Python au cas où vous auriez besoin de vous y référer ultérieurement.

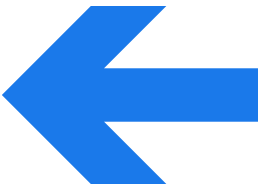
# Environnement: **Installation Python**



1. Cliquez sur **Installer** pour démarrer l'installation.
2. Une fois l'installation terminée, un **message indiquant que l'installation a réussi s'affiche.**



# Environnement: Installation Python



## Étape 3 Ajout de Python aux variables d'environnement (facultatif)

Ignorez cette étape si vous avez sélectionné Ajouter Python aux variables d'environnement lors de l'installation.

Si vous souhaitez accéder à Python via la ligne de commande mais que vous n'avez pas ajouté Python à vos variables d'environnement lors de l'installation, vous pouvez toujours le faire manuellement.

Avant de commencer, recherchez le répertoire d'installation de Python sur votre système. Les répertoires suivants sont des exemples de chemins de répertoire par défaut :

- **C:\Program Files\Python310** : si vous avez sélectionné Installer pour tous les utilisateurs lors de l'installation, le répertoire sera alors à l'échelle du système
- **C:\Users\Onesime\AppData\Local\Programs\Python\Python310** : si vous n'avez pas sélectionné Installer pour tous les utilisateurs lors de l'installation, le répertoire sera alors dans le chemin d'accès de l'utilisateur Windows.

Notez que le nom du dossier sera différent si vous avez installé une version différente, mais démarrera toujours avec Python.

1. Allez dans **Démarrer** et saisissez les **paramètres système avancés** dans la barre de recherche.
2. Cliquez sur **Afficher les paramètres système avancés**.
3. Dans la **boîte de dialogue Propriétés système**, cliquez sur **l'onglet Avancé**, puis sur **Variables d'environnement**.

# Environnement: Installation Python



1. En fonction de votre installation :

- Si vous avez sélectionné **Installer pour tous les utilisateurs** lors de l'installation, sélectionnez **Chemin** dans la liste des **variables système** et cliquez sur **Modifier**.
- Si vous n'avez pas sélectionné **Installer pour tous les utilisateurs** lors de l'installation, sélectionnez **Chemin** dans la liste des variables utilisateur et cliquez sur **Modifier**.
- Cliquez sur **Nouveau** et entrez le chemin du répertoire Python, puis cliquez sur **OK** jusqu'à ce que toutes les boîtes de dialogue soient fermées.

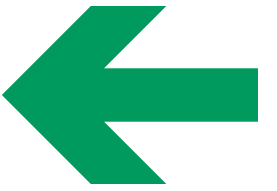
## Étape 4 Vérifier l'installation de Python

- Vous pouvez vérifier si l'installation de Python a réussi soit via la ligne de commande, soit via l'application Environnement de développement intégré (IDLE), si vous avez choisi de l'installer.
- Allez dans **Démarrer** et saisissez cmd dans la barre de recherche.  
**Cliquez sur Invite de commandes.**
- Saisissez la commande suivante dans l'invite de commandes :

```
python --version
```

```
Output  
Python 3.10.10
```

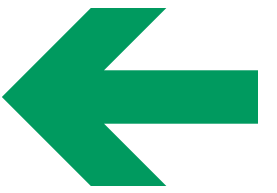
# Environnement: Installation Anaconda



**Anaconda** est un gestionnaire de paquets, un gestionnaire d'environnement et une distribution Python qui contient une collection de paquets open source. C'est un avantage car lorsque vous travaillez sur un projet de Data Science, vous constaterez que vous avez besoin de nombreux paquets différents (numpy, scikit-learn, scipy, pandas pour n'en citer que quelques-uns), avec lesquels une installation d'Anaconda est préinstallée.

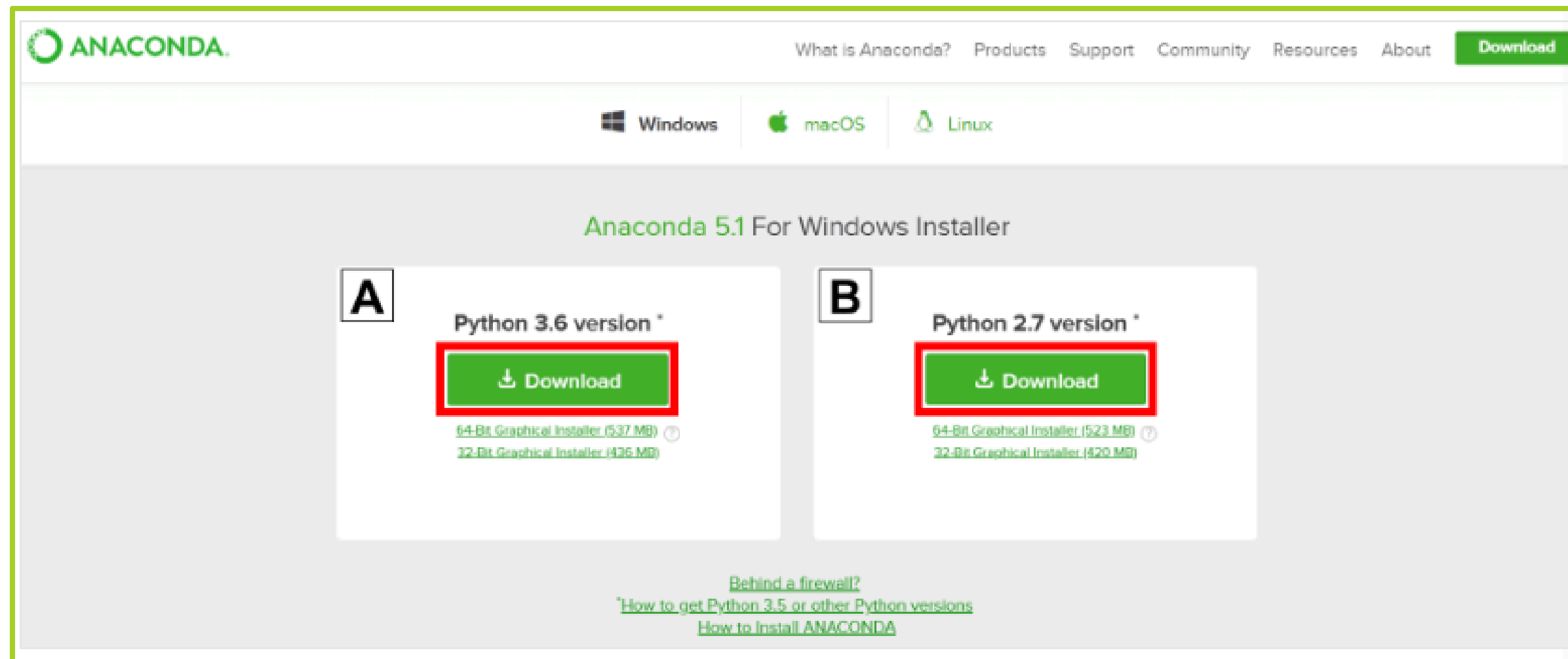
Si vous avez besoin de paquets supplémentaires après l'installation d'Anaconda, vous pouvez utiliser le gestionnaire de paquets d'Anaconda, **conda** ou **pip** pour installer ces paquets. C'est très avantageux car vous n'avez pas à gérer vous-même les dépendances entre plusieurs paquets. Conda facilite même le basculement entre Python 2 et 3. En fait, l'installation d'Anaconda est également la méthode recommandée pour installer Jupyter Notebooks.

# Environnement: Installation Anaconda



## Comment télécharger et installer Anaconda

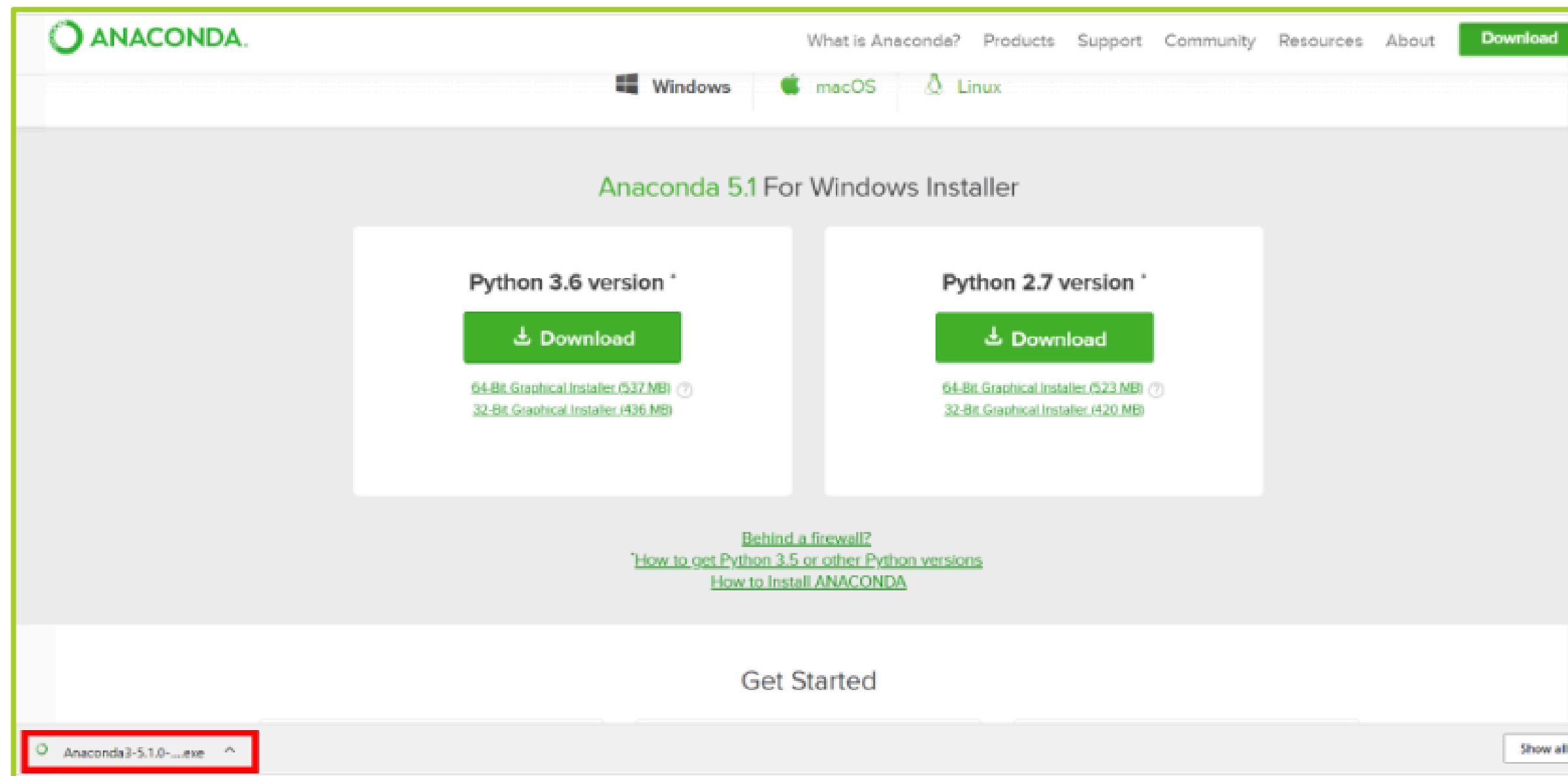
1. Accédez au site Web d'Anaconda et choisissez un programme d'installation graphique Python 3.x (A) ou un programme d'installation graphique Python 2.x (B).



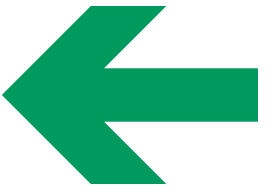
# Environnement: **Installation Anaconda**



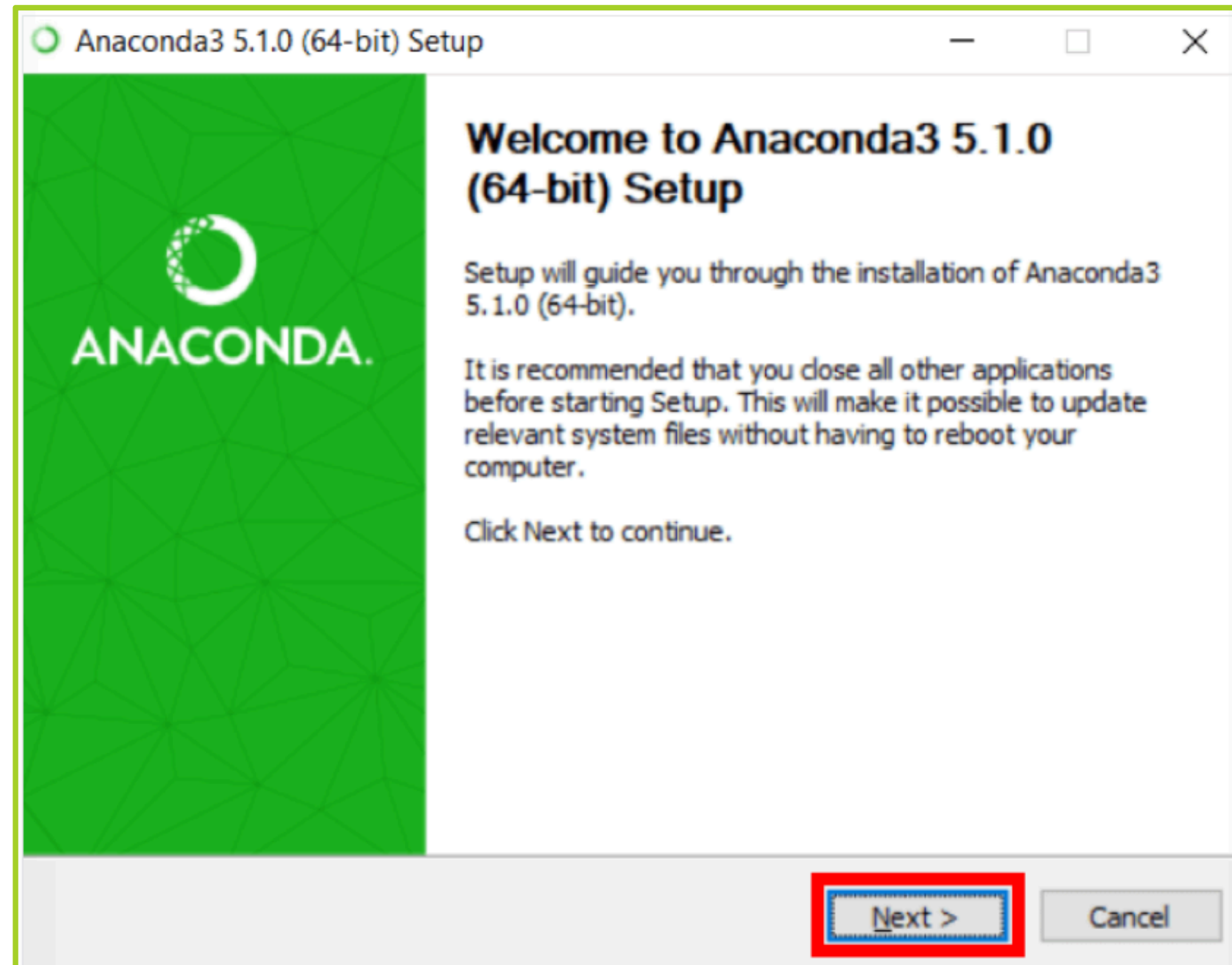
**2. Localisez votre téléchargement et double-cliquez dessus.**



# Environnement: **Installation Anaconda**

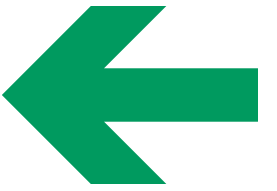


Lorsque l'écran ci-dessous apparaît, cliquez sur Suivant.

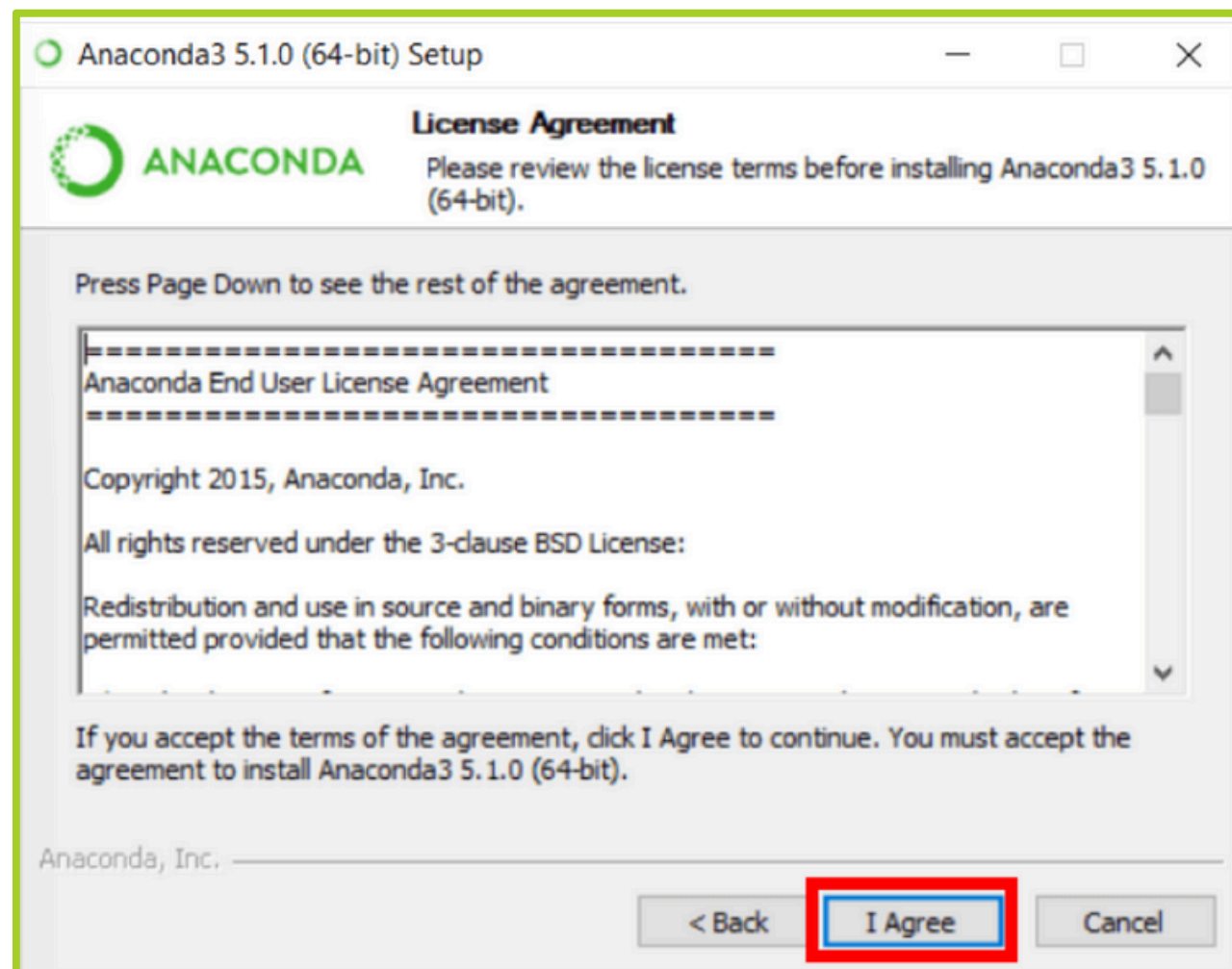




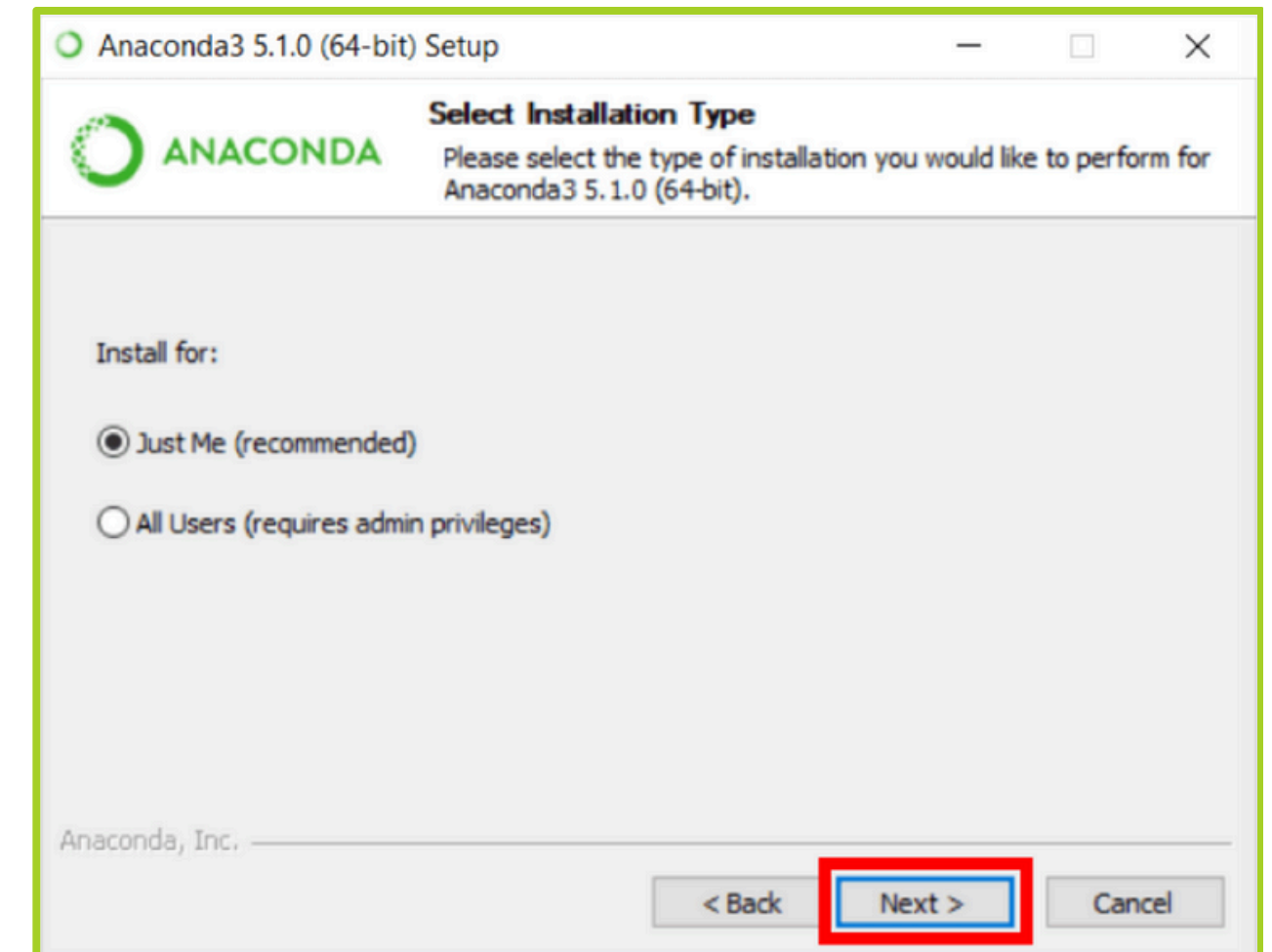
# Environnement: Installation Anaconda



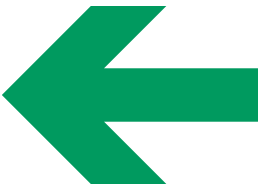
3. Lisez le contrat de licence et cliquez sur J'accepte.



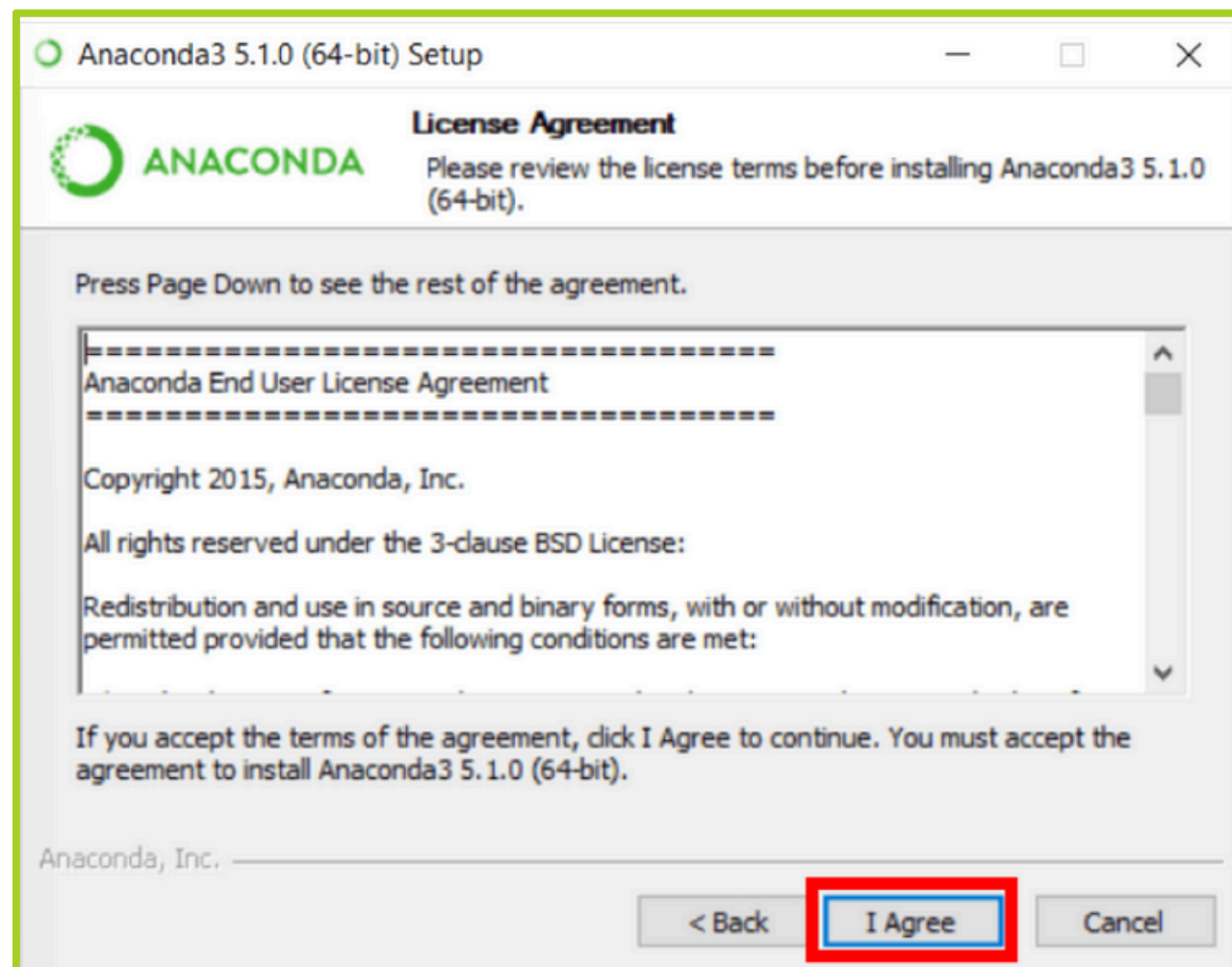
4. Cliquez sur Suivant.



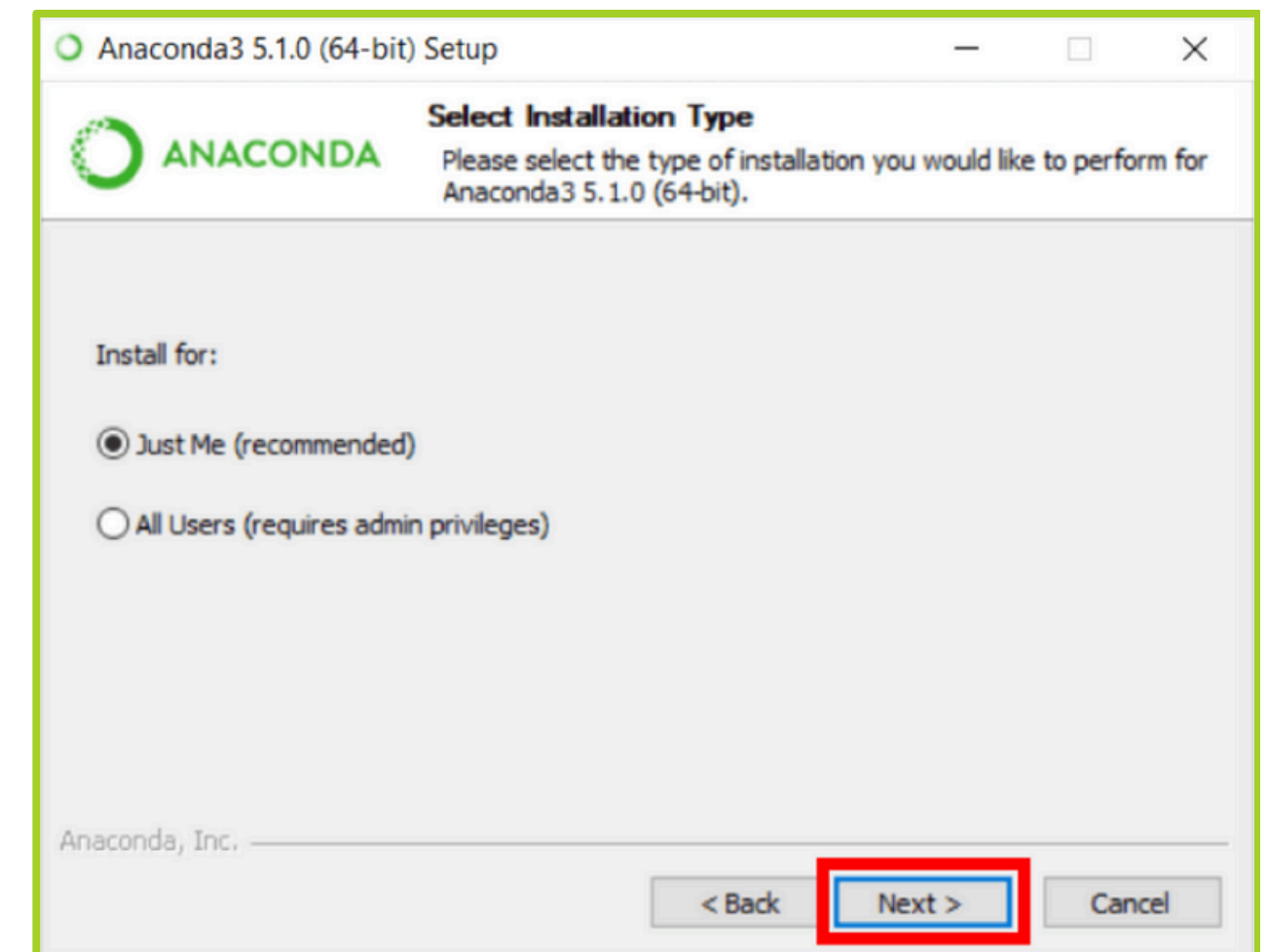
# Environnement: Installation Anaconda



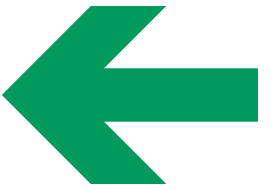
3. Lisez le contrat de licence et cliquez sur J'accepte.



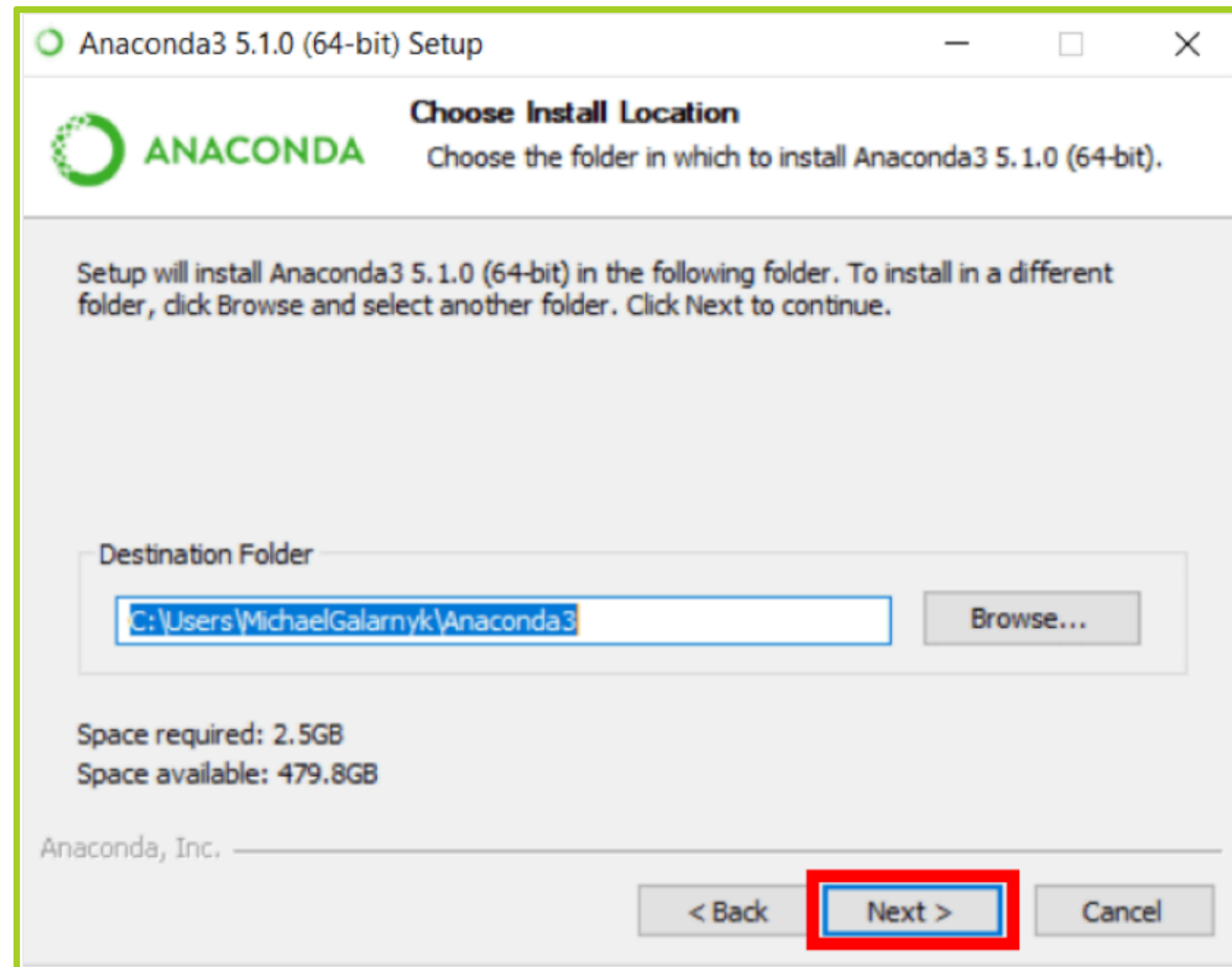
4. Cliquez sur Suivant.



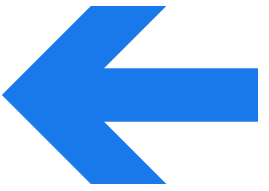
# Environnement: Installation Anaconda



5. Notez votre emplacement d'installation, puis cliquez sur Suivant.

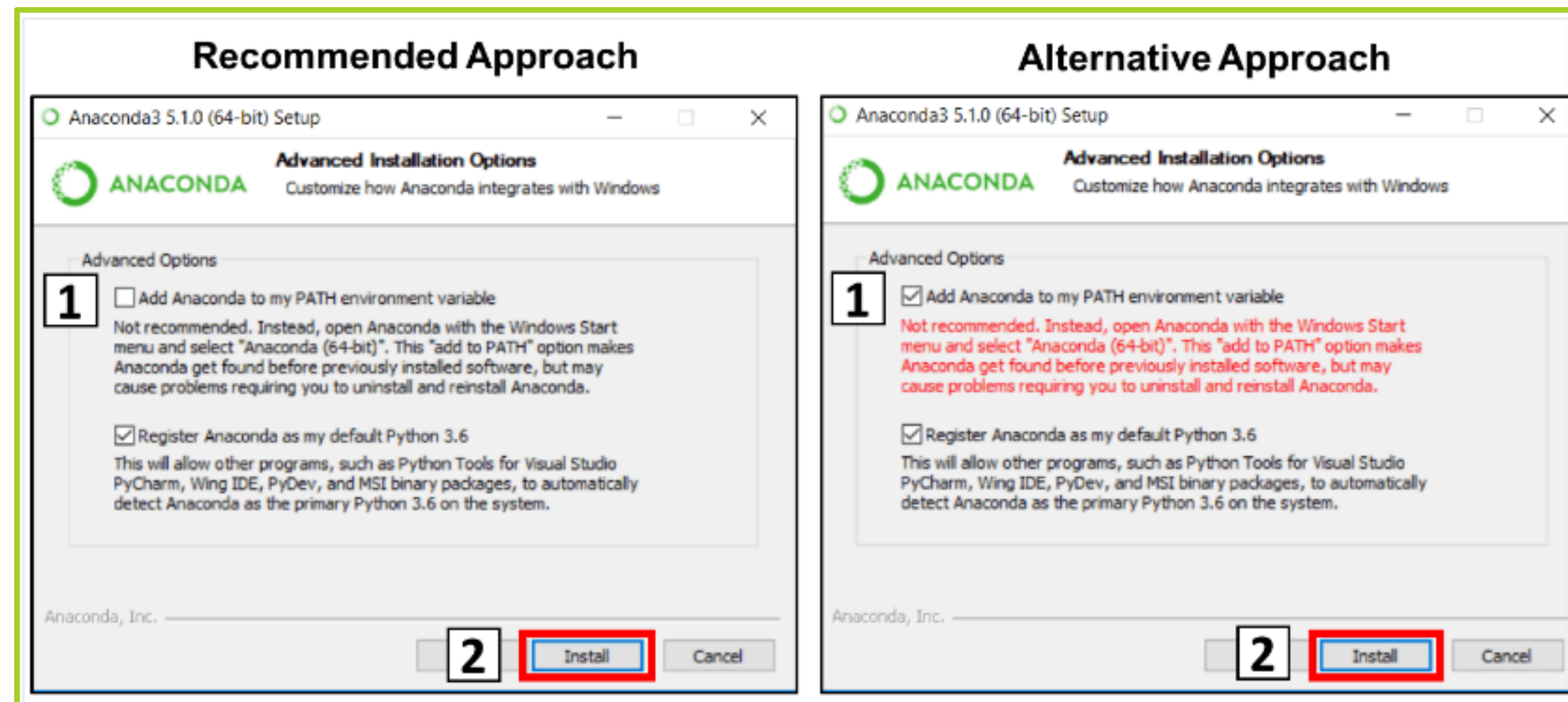


# Environnement: Installation Anaconda

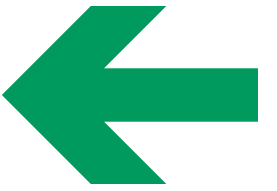


6. Il s'agit d'une partie importante du processus d'installation. L'approche recommandée consiste à ne pas cocher la case pour ajouter Anaconda à votre chemin. Cela signifie que vous devrez utiliser Anaconda Navigator ou l'invite de commande Anaconda (située dans le menu Démarrer sous « Anaconda »). lorsque vous souhaitez utiliser Anaconda (vous pouvez toujours ajouter Anaconda à votre PATH plus tard si vous ne cochez pas la case).

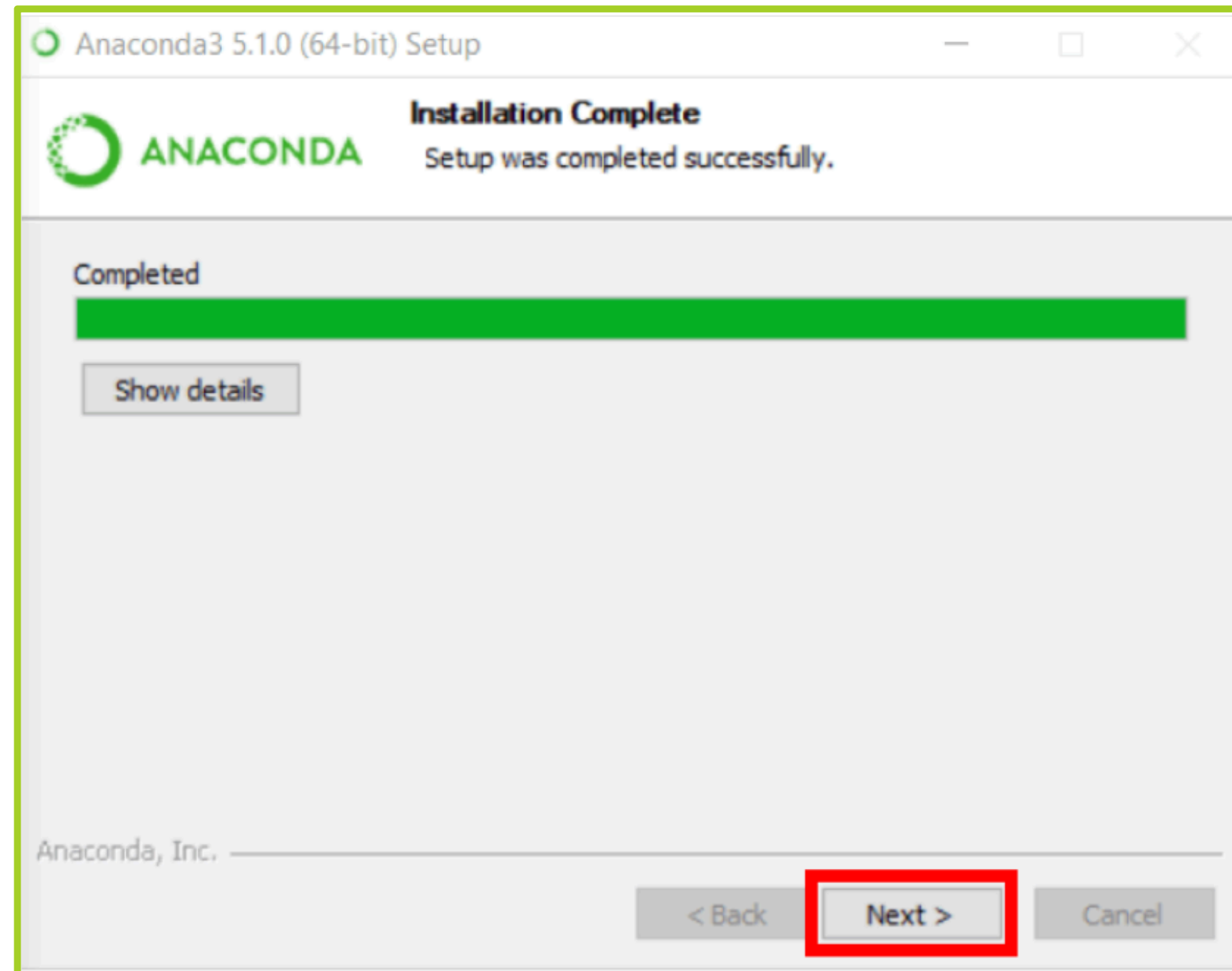
Si vous souhaitez pouvoir utiliser Anaconda dans votre invite de commande (ou git bash, cmd, powershell, etc.), veuillez utiliser l'approche alternative et cocher la case.



# Environnement: **Installation Anaconda**

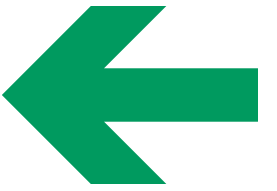


**7. Cliquez sur Suivant.**

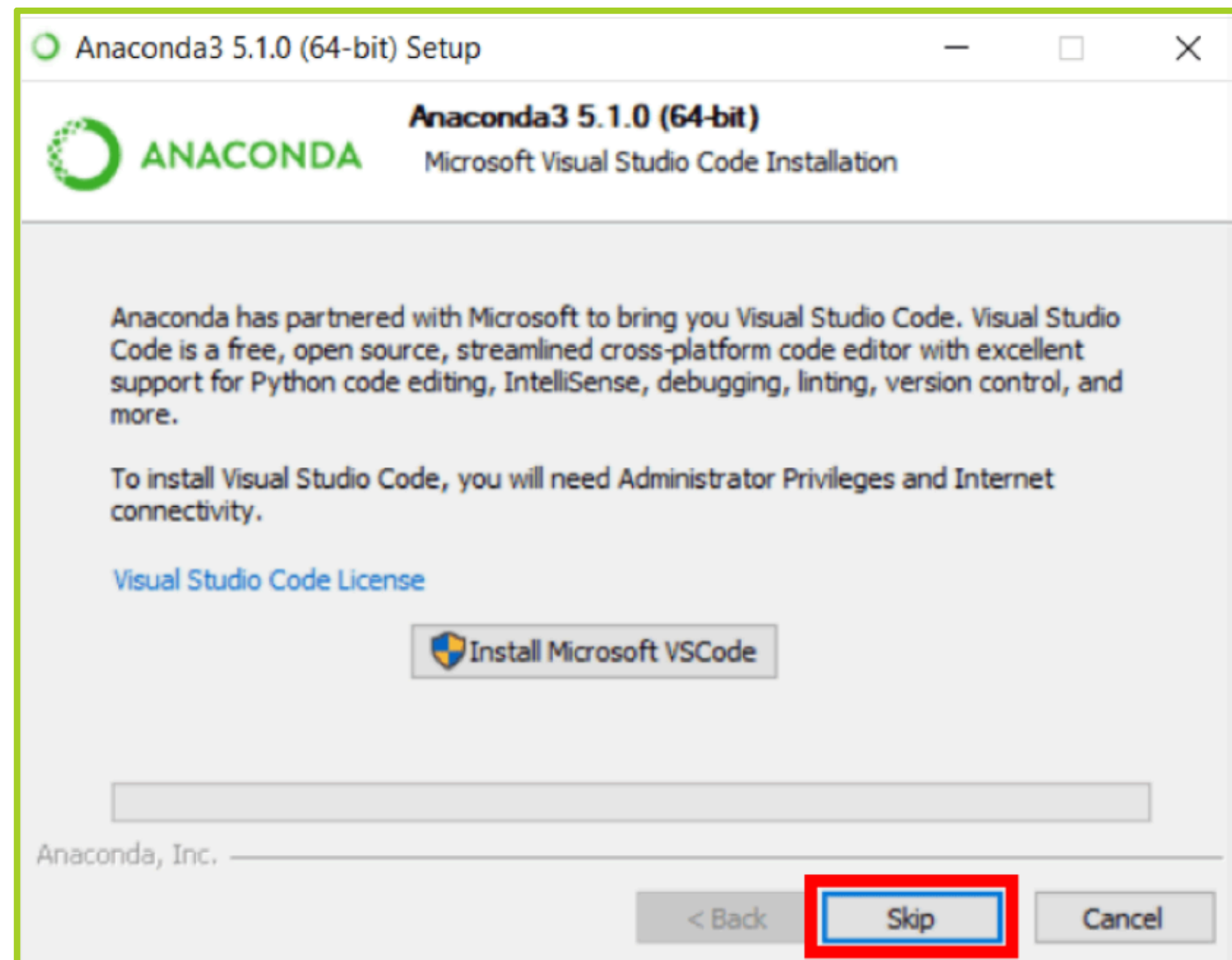




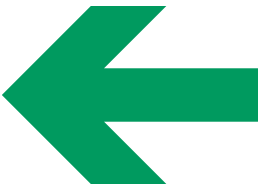
# Environnement: **Installation Anaconda**



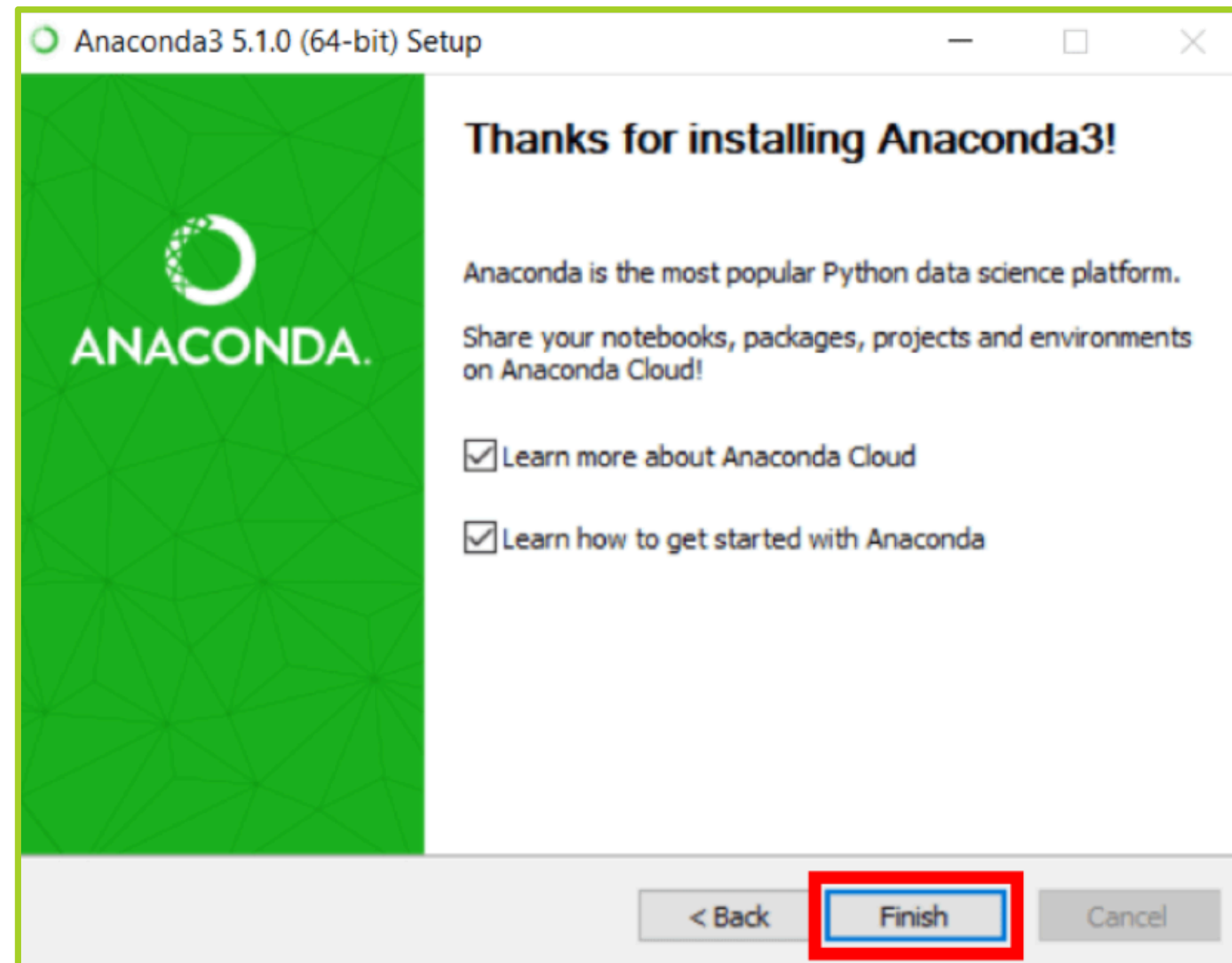
**8. Vous pouvez installer Microsoft VSCode si vous le souhaitez, mais cela est facultatif.**



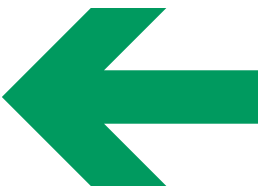
# Environnement: **Installation Anaconda**



**9. Cliquez sur Terminer.**



# Environnement: Installation Anaconda



## Comment ajouter Anaconda au chemin (facultatif)

Il s'agit d'une étape facultative. Elle concerne le cas où vous n'avez pas coché la case à l'étape 6 et souhaitez maintenant ajouter Anaconda à votre chemin. L'avantage de cette opération est que vous pourrez utiliser Anaconda dans votre invite de commande, Git Bash, cmdr, etc.

### 1. Ouvrez une invite de commande.

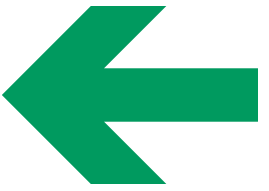
A screenshot of a Windows Command Prompt window. The title bar reads "Command Prompt". The window content shows the text: "Microsoft Windows [Version 10.0.16299.431]" followed by "(c) 2017 Microsoft Corporation. All rights reserved." and the command prompt "C:\Users\MichaelGalarnyk>". The window has standard Windows window controls (minimize, maximize, close) in the top right corner.

```
Command Prompt
Microsoft Windows [Version 10.0.16299.431]
(c) 2017 Microsoft Corporation. All rights reserved.

C:\Users\MichaelGalarnyk>
```



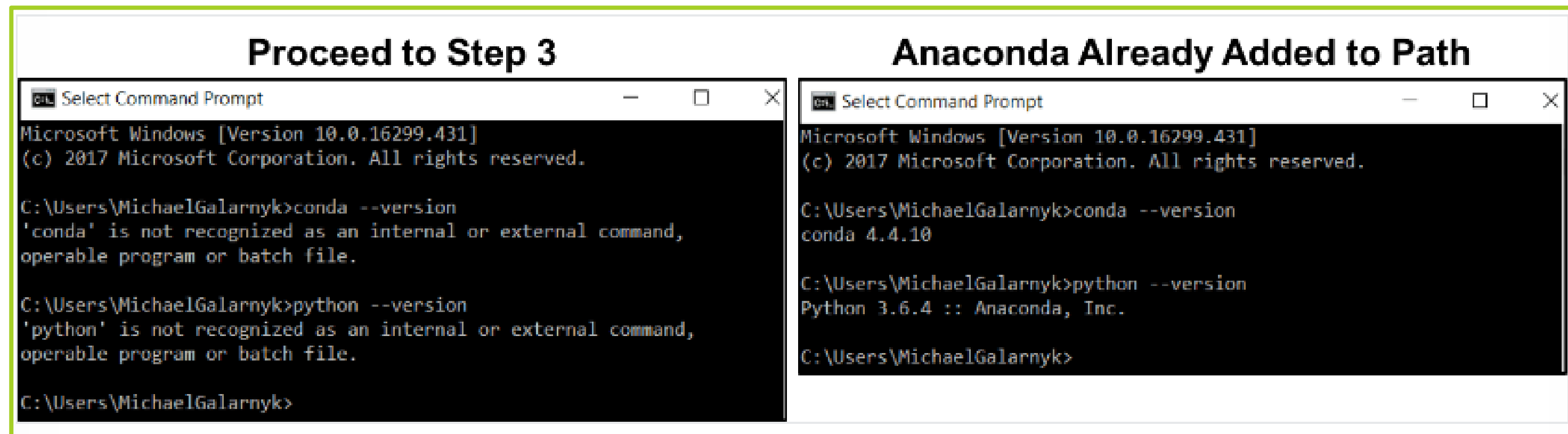
# Environnement: Installation Anaconda



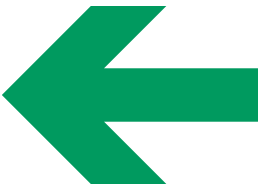
2. Vérifiez si Anaconda est déjà ajouté à votre chemin. Entrez les commandes ci-dessous dans votre invite de commande. Cela permet de vérifier si Anaconda est déjà ajouté à votre chemin. Si vous obtenez une erreur de commande non reconnue comme dans le côté gauche de l'image ci-dessous, passez à l'étape 3. Si vous obtenez un résultat similaire au côté droit de l'image ci-dessous, vous avez déjà ajouté Anaconda à votre chemin.

**conda --version**

**python --version**



# Environnement: Installation Anaconda



3. Si vous ne savez pas où se trouvent votre conda et/ou python, ouvrez une invite Anaconda et saisissez les commandes suivantes. Cela vous indique où se trouvent conda et python sur votre ordinateur.

where conda

where python

```
Anaconda Prompt

(base) C:\Users\MichaelGalarnyk>where conda
C:\Users\MichaelGalarnyk\Anaconda3\Library\bin\conda.bat
C:\Users\MichaelGalarnyk\Anaconda3\Scripts\conda.exe

(base) C:\Users\MichaelGalarnyk>where python
C:\Users\MichaelGalarnyk\Anaconda3\python.exe

(base) C:\Users\MichaelGalarnyk>
```

# **Atelier Pratique**

(Jupyter Notebook)

# Thank You

