

IFT870 – SÉANCE 3

ATELIER

Nadia Tahiri, Ph. D.
Professeure adjointe
Université de Sherbrooke

Nadia.Tahiri@USherbrooke.ca



PYTHON ET BIBLIOTHÈQUES



Python 3 : intègre des fonctionnalités générales et de nombreuses bibliothèques spécifiques pour l'analyse de données

Installer Python (<https://www.python.org/>) et
pip (<https://pip.pypa.io/en/stable/>)

scikit-learn : implémentation de nombreux algorithmes de forage de données supervisé et non-supervisé, avec une documentation bien fournie (http://scikit-learn.org/stable_).

Pour l'installer voici un exemple de ligne de commande : `pip install scikit-learn`

Jupyter Notebook : permet l'interaction directe avec du code via un navigateur (<https://jupyter.org/>)

```
pip install ipython
pip install jupyterlab
pip install notebook
jupyter notebook
```

PYTHON ET BIBLIOTHÈQUES



numpy et scipy : pour des calculs scientifiques en Python: calcul matricielle, opérations d'algèbre linéaire, génération de nombres aléatoires, fonctions d'optimisation, distributions statistiques, et autres

```
pip install numpy scipy
```

pandas : bibliothèque pour la manipulation de données

```
pip install pandas
```

pillow : bibliothèque pour la manipulation de données d'imagerie

```
pip install pillow
```

matplotlib : bibliothèque pour la visualisation de données

```
pip install matplotlib
```

seaborn : bibliothèque pour la visualisation de données statistiques

```
pip install seaborn
```

LA SEMAINE PROCHAINE

INTRODUCTION

Algorithmes usuels de forage de données.

EXPLORATION DES DONNÉES

Types de données et d'attributs, descriptions statistiques, mesures de similarité, dissimilarité, et visualisation.