



PROGRAMME SUR PYTHON

1. Au travers d'IPython (version pour l'analyse interactive de Python)

- Présentation de l'environnement python
- Configuration de l'espace de travail sur python
- Présentation du navigateur ANACONDA
- Présentation des différents IDEs (**IPython, Jupiter notebook, spider, etc**)

2. Les bases de Python

- Organisation d'un programme
- Types de données simples et complexes, List

3. Compréhension

- Structures de contrôle (if, while, for, etc)
- Gestion des fichiers
- Principaux modules et fonctions

4. Préparation des données avec pandas

- Traitement des données manquantes et aberrantes
- Lecture et écriture des données depuis et vers différentes sources (fichiers textes, Excel, SPSS, STATA, CSV bases de données, etc.)
- Filtrage, sélection, transformation, calcul, agrégation, jointure, sorties simples

5. La visualisation des données avec matplotlib & seaborn



- Revue des différents tableaux statistiques (univarié, bivarié)
- Revue des différents types de graphiques (histogrammes, diagrammes, camembert, etc)

6. Apprentissage et analyse statistique avec sci-kit learn & statsmodels

- Gestion des ensembles d'apprentissage et de test (Tests de khi deux, test student, test de Pearson, etc.)
- Évaluation des modèles (régressions linéaires et non linéaires)

7. Projet de formation (rapport statistique)

Durée de la formation : 30 jours

TOUS - STATISTICIEN