

Programme de Formation

Analyse de données (Big data)

Organisation

Durée : 21 heures

Mode d'organisation : Présentiel

Contenu pédagogique



Public visé

- Responsable Recherche & Développement
- Responsable de laboratoire d'analyses
- Responsable de l'amélioration continue dans l'entreprise
- Agent méthode & industrialisation
- Chargé(e) d'amélioration continue / excellence opérationnelle
- Technicien/ne en Recherche & Développement



Objectifs pédagogiques

L'objectif de la formation est d'apprendre à analyser un nombre important de données à l'aide des principales méthodes d'analyse statistique multivariée.

L'apprenant apprend à faire le bon choix de la méthode selon le type de données et l'objectif défini (décrire, classer, modéliser).

L'apprenant interprète et prend des décisions sur de nombreux cas pratiques, en s'aidant de solutions informatiques et de l'intelligence artificielle (IA).



Description

Maîtriser les principales méthodes d'analyse statistique de données multivariée afin d'explorer, modéliser et interpréter des (grandes) bases de données complexes pour l'aide à la décision et l'optimisation dans des processus de reporting (tableau de bord), de recherche (data science), de performance industrielle, ainsi que de services (données économiques). Découvrir comment les outils d'intelligence artificielle (IA) permettent de gagner en performance et rapidité d'analyse.

Reconnaissance de formation : Attestation SMACA (accordée sous condition d'une présence minimale de 60 %).



Contenu

- Introduction au Big Data
- Les premiers pas en Big Data : Décrire et valider la qualité des données
- Les différentes méthodes d'apprentissage supervisé
 - Régression linéaire $Y = f(X)$
 - Régression logistique binaire $(0 ; 1) = f(X)$, et classification supervisée
 - Analyse Discriminante Linéaire (LDA)
- Validation de modèle : apprentissage supervisé
- Les différentes méthodes d'apprentissage non supervisé

- Classification non supervisée : Clustering K-Means
- Réduction de dimension : Analyse en composantes principales
- Les réseaux de neurones : la puissance de l'approximation
- La régression PLS : la mixité d'apprentissage

Tout au long des 3 journées de formation, exploitation numérique et graphique de plusieurs études (base de données) de cas pratiques avec l'aide de différentes solutions logicielles et en s'aidant de l'intelligence artificielle (IA).

★ **Prérequis**

Formation SMACA :

- Statistiques de base
- ou formation équivalente,
ou expérience professionnelle



Modalités pédagogiques

Atelier :

- Postes de contrôle & instruments de mesure connectés en simulation de situation professionnelle de mesure & d'analyse de données
- Travail sur des études de cas basées sur des données réelles

École : Salle de formation théorique & média interactif pour l'animation ; utilisation de logiciel statistique ainsi que de l'intelligence artificielle (IA)

Livrables pour l'apprenant :

1 accès individuel sur la plateforme digitale permettant à l'apprenant en amont des journées en présentiel d'accéder :

- Au test d'auto-évaluation des connaissances (quiz de quelques questions)
- A la documentation de formation téléchargeable, regroupant les thèmes abordés durant la ou les journées en présentiel

Cet accès individuel permet aussi au terme des journées en présentiel, d'accéder de nouveau :

- Au test final de validation des acquis de la formation
- Au pack de fichiers téléchargeables en lien avec les notions théoriques & pratiques développées durant la ou les journées en présentiel