

Durée de la formation ?

2 jours – soit 14 heures.

À qui s'adresse cette formation ?

À toute personne susceptible d'intervenir dans l'élaboration et la mise en œuvre d'une stratégie Big Data, et ce dans les domaines : produits, marketing, distribution et tarification.

Ainsi qu'à toute personne désireuse de se familiariser avec ce sujet en devenir.

Pour obtenir quoi ?

Une compréhension globale de la problématique Big Data et des outils disponibles à ce jour et nécessaires à son traitement.

Quels objectifs pédagogiques ?

Répondre aux problématiques d'utilisation de R en entreprise.

Maîtriser R pour la manipulation des données et le machine learning.

Présenter les résultats des modèles prédictifs.

Développer les capacités à interpréter le fonctionnement des modèles.

Quelles méthodes mobilisées ?

Des exemples pratiques animeront chaque étape du processus de traitement du Big Data, et compléteront la définition et les perspectives de ce nouveau métier.

Quels sont les prérequis ?

Aucun en particulier. L'intérêt et l'utilisation des outils numériques sont cependant fortement recommandés.

Quelles modalités d'évaluation ?

Une évaluation des acquis des objectifs sera réalisée durant la formation.

Qui anime cette formation ?

Kezhan SHI,

Il est diplômé de l'École Centrale Paris et titulaire d'un master en actuariat de l'Université Paris Dauphine. Il a travaillé chez Axa Global Direct et Allianz, avant de rejoindre Diot Siaci en 2022, au titre de Responsable du Data Lab.



La formation en pratique...

Quand et où ?

9 et 10 décembre 2021

9 h 00 - 12 h 30 et 14 h 00 - 17 h 30
Caritat, Paris 8°

Combien ça coûte ?

2 100 € HT + TVA 20 %, soit 2 520 € TTC.

Les frais de participation couvrent les deux journées de formation, la documentation complète, les déjeuners et les pauses café.

Qu'allez-vous apprendre ?

Introduction à la programmation avec « R »

- Installation de R et de RStudio
- Présentation de RStudio

Collecte des données web avec « R »

- Import des données
- Bases de la structure des pages web
- Scraping des données html avec « R »
- Cas d'application : analyse des avis et notations des produits
- Import et analyse des données de twitter
- Construction d'un outil de veille pour les produits ou les services

Nettoyage de données html et textes

- Text mining
- Manipulation des corpus avec « R »
- Exploitation des document term matrix (DTM)
- Association des mots
- Visualisation des nuages de mots

Machine learning avec « R »

- Analyse prédictive des mails indésirables
- Pratique de différents algorithmes avec « R »
- Application d'un arbre de classification
- Application de SVM (Support Vector Machine)
- Application de la régression logistique
- Agrégation de modèles
- Comparaison des différents algorithmes

Performance des algorithmes

- Matrice de confusion et taux d'erreur
- Courbe ROC et AUC
- Coefficient de Gini

Conclusions

- Intérêt de l'analyse des données textes
- Intérêt du processus d'analyse prédictive

01 44 51 04 00
info@caritat.fr

! Chaque participant se munira d'un ordinateur portable pour les travaux pratiques.

Qu'en disent les stagiaires ?

« Le programme est très intéressant et le formateur est passionné. »

GL, Responsable de l'actuariat – ADDING

« Le programme est en cohérence avec mes attentes. L'exposé a été très clair. »

MK, Actuaire – AXA

« Formateur à l'écoute et pédagogue. »

NP, Chargé d'études statistiques – MACIF