

# DATA SCIENCE MENTORING : CAS PRATIQUE D'APPLICATION

5 jours

Fiche d'inscription  
sur [www.caritat.fr](http://www.caritat.fr)

E-learning

120<sup>Pts</sup> PPC  
Actuaire

## À qui s'adresse cet e-learning ?

À toute personne souhaitant analyser des données dans les différents domaines d'assurance : marketing, tarification, résiliation, sinistres, fraudes, innovation produits.

## Pour obtenir quoi ?

Démystifier la Data Science en réalisant un projet avec « R ».

## Comment ?

1h de mentoring par appel vidéo, 2h d'exercices par jour, assistance pour le code, réponses aux questions (téléphone, appel vidéo, mail). Les supports seront adressés par mail, et comprendront les explications sur les différentes techniques et fonctions, des exercices et des mises en pratique pour la réalisation du projet. Chaque jour avant 20h, les réponses aux questions doivent être envoyées. Vous pouvez poser des questions à tout moment ; la question sera traitée en moins de 2h de 9h à 22h.

Vous devez envoyer le projet finalisé dans un dossier comprenant : un fichier rmd, un fichier html, directement compilé à partir du fichier rmd. En fin de formation, vous recevrez la correction pour la réalisation du projet.

## Quels sont les pré-requis ?

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

! Chaque participant se munira d'un ordinateur portable avec webcam et micro, d'une bonne connexion internet, sur le navigateur Chrome.

## Qui anime cet e-learning ?

**Kezhan SHI,**  
Membre certifié de l'IA. Il est titulaire d'un master en actuariat de l'Université Paris Dauphine. Il a également fait l'École Centrale Paris. Il a travaillé chez Prim'Act, Direct Assurance et Aviva, avant de rejoindre Allianz en 2017, au titre de Data Scientist.



## L'e-learning en pratique...

### Quand et où ?

**5 au 9 mars 2018**

9 h 00 - 10 h 00

Site Internet : <https://appear.in/caritat>

### Combien ça coûte ?

500 € HT + TVA 20 %, soit 600 € TTC.

Les frais de participation couvrent les journées de formation et la documentation complète.

## Qu'allez-vous apprendre ?

La Data Science est une nouvelle matière qui s'appuie sur trois piliers : la capacité de calcul, l'accès à de volumineuses bases de données, et l'application à des modèles prédictifs très utiles en assurance.

Caritat a mis en place une formation à distance qui permet des calculs réels en utilisant des outils accessibles sur le net et des bases open data. La toile n'offre pas accès à des bases utilisables pour les problématiques assurantielles.

Caritat a donc appuyé ses exemples pratiques sur des données du prix de l'immobilier qui présente des similitudes certaines avec l'assurance dans la mesure où l'on doit déterminer le prix d'un produit dont on ne connaît pas le prix de revient.

**Jour 1 :** introduction à « R »

**Jour 2 :** webscrapping

**Jour 3 :** nettoyage de données

**Jour 4 :** visualisation des données

**Jour 5 :** machine learning

**Projet final :** analyse du prix immobilier à Paris

Vous approfondirez les éléments suivants :

- **Webscrapping**
  - Connaissance de la structure des pages html
  - Manipulation des codes html
  - Manipulation des expressions régulières
- **Nettoyage de données**
  - Manipulation des vecteurs et des data.frame
  - Fusion et sélection de données
- **Machine learning**
  - Apprentissage supervisé
  - Régression
  - Prédiction du prix immobilier
- **Visualisation des données**
  - Utilisation du package ggplot2
  - Types de visualisation
  - Graphiques interactifs
  - Formattage des tables
- **Visualisation géographique**
  - Manipulation des fichiers shapefile
  - Carte choroplète
  - Graphiques interactifs
- **Outil de production**
  - Rmarkdown
  - htmlwidgets
  - Tableaux de bord

01 44 51 04 00  
[info@caritat.fr](mailto:info@caritat.fr)

## Qu'en disent les stagiaires ?

Cet e-learning est une nouveauté du catalogue Caritat.