



# Data Science pour l'agriculture

Du 23 au 27 septembre 2024  
Formation à distance

## Objectifs

- Acquérir les bases permettant de manipuler les principales méthodes de data science pour des objectifs de prédiction
- Appréhender la pratique réelle de ces méthodes au travers d'exemples et de travaux pratiques avec le logiciel R

## Public

- Ingénieurs et chercheurs travaillant dans les domaines de l'agriculture et de l'environnement.

## Pré-requis

- Connaissances de base du logiciel R
- Maîtriser les méthodes de base de régression
- Avoir un compte google pour utilisation de google colab

L'inscription à la formation est validée par un questionnaire de positionnement.

## Méthode

### Moyens pédagogiques

- Exposés, échanges et quizz
- Travaux pratiques personnels

### Équipements nécessaires

Disposer d'un équipement pour participer à une classe virtuelle (ordinateur avec connexion internet et web camera). Durant la formation, il est recommandé de s'isoler.

### Moyens d'évaluation

- Questionnaire de positionnement
- Évaluation des compétences acquises lors de la formation par des quizz et des exercices pratiques individuels
- Évaluation de la formation par un questionnaire de satisfaction

### Moyen de suivi et formalisation à l'issue de la formation

- Émargement par ½ journée
- Certificat de réalisation transmis par courriel à l'issue de la formation

### Ressources pédagogiques

- PDF de la formation fourni en début de session
- Ouvrage collectif « Data science pour l'agriculture et l'environnement – Méthodes et applications avec R et Python » Editions Ellipses

# Contenu de la formation

## Classe virtuelle du 23/09 de 9h à 12h10

- 9h-10h : Présentation de la formation, tour de table et introduction à la datascience
  - 10h10-12h10 : Méthodes de régression (pénalisée, PLS, GAM, sélection)
- + Travail individuel « **Méthodes de régression** » à réaliser en asynchrone et à rendre avant le soir aux formateurs (durée estimée 1h)

## Classe virtuelle du 24/09 de 9h à 12h10

- 9h-10h : Retour sur le travail individuel « Méthodes de régression »
  - 10h10-12h10 : Arbres et random forest
- + Travail individuel « **Arbres et random forest** » à réaliser en asynchrone et à rendre avant le soir aux formateurs (durée estimée 1h)

## Classe virtuelle du 25/09 de 9h à 11h40

- 9h-10h : Retour sur le travail individuel « Arbres et random forest »
  - 10h10-11h40 : Méthodes d'évaluation (qualité de prédiction)
- + Travail individuel « **Méthodes d'évaluation** » à réaliser en asynchrone et à rendre avant le soir aux formateurs (durée estimée 1h)  
(merci de veiller à profiter de la longue pause avant la reprise à 14h pour réaliser ce travail individuel)

## Classe virtuelle du 25/09 de 14h à 16h10

- 14h-15h30 : Data mining
  - 15h40-16h10 : Introduction Google Colab et Python - partie 1
- + Travail individuel « Google Colab et Python » à réaliser en asynchrone et à rendre avant le soir aux formateurs (durée estimée 1h)

## Classe virtuelle du 26/09 de 9h à 12h10

- 9h-9h30 : Retour sur les travaux individuels « Méthodes d'évaluation » et « google colab et python »
  - 9h30-10h30 : Introduction Google Colab et Python – partie 2
  - 10h40-12h10 : Réseau de neurones et extension multicouche
- + Travail individuel « **Réseau de neurones et extension multicouche** » à réaliser en asynchrone et à rendre avant le soir aux formateurs (durée estimée 1h30)

## Classe virtuelle du 27/09 de 9h à 12h40

- 9h-10h : Retour sur le travail individuel « Réseau de neurones et extension multicouche »
- 10h10-12h10 : Réseau de neurones et convolutions
- 12h10-12h40 : Conclusion et clôture de la formation

## Informations complémentaires

**Durée :** 25h sur 6 demi-journées (dont 2 le mercredi)

**Tarif :** 1250 € HT

**Tarif réduit :**

950 € HT pour les partenaires RMT Science des données et Modélisation pour l'agriculture et l'agroalimentaire ([www.modelia.org](http://www.modelia.org)) ou pour les doctorants financés ou labellisés par l'institut de convergence DigitAg – Agriculture Numérique ([www.hdigitag.fr](http://www.hdigitag.fr)).

**Intervenants :**

- **François Brun**, ingénieur, animateur du RMT Modelia, intervenant sur les méthodes d'évaluation (Acta)
- **David Makowski**, directeur de recherche, animateur du RMT Modelia, intervenant sur l'introduction et les méthodes de régression (INRAE)
- **Maxime Legris**, ingénieur, intervenant sur les arbres et forêts aléatoires (Idele)
- **Florent Duyme**, ingénieur, intervenant sur les méthodes de régression (ARVALIS)
- **Mohammed El Jabri**, ingénieur, intervenant sur les réseaux de neurones (Idele)
- **Alexandre Termier**, professeur, intervenant sur le data mining (IRISA (université de Rennes-INRIA))

**Modalités et délais d'accès :** inscription sur [acta.asso.fr/formations/](http://acta.asso.fr/formations/)

*Vous pouvez vous inscrire jusqu'à 3 semaines avant la formation. Au-delà, merci de nous contacter.*

**Accessibilité aux personnes en situation de handicap :** nous contacter

**CONTACT**



**Pédagogique**

**François Brun**

Tél. : 06 25 78 29 94

Email : [francois.brun@acta.asso.fr](mailto:francois.brun@acta.asso.fr)

**Administratif**

**Volimata Camara**

Tél : 01 81 72 17 07

Email : [volimata.camara@acta.asso.fr](mailto:volimata.camara@acta.asso.fr)