

# Traitement statistique de données agronomiques

Principes et mise en oeuvre avec R



## A propos

Ingénieur, chercheur ou technicien supérieur au sein d'un service de R&D ou d'une unité de recherche dans les domaines de l'agronomie ou de l'agroalimentaire, vous êtes confronté à l'exploitation de données expérimentales, voire à la mise en place de dispositifs expérimentaux.

Modulaire et progressif, combinant formation à distance et présentiel, **ce stage vous apportera à la fois les outils et les éléments méthodologiques permettant de manipuler, explorer, présenter et exploiter statistiquement vos jeux de données.**



## Objectifs de la formation

- Découvrir le langage R et l'environnement RStudio. Manipuler des données avec R. Rédiger des scripts
- Connaître les principales approches univariées et bivariées de statistique descriptive ; savoir les mettre en oeuvre à l'aide de R et RStudio
- Savoir choisir et mettre en oeuvre les principales méthodes de statistique inférentielle en expérimentation agronomique ; savoir choisir et mettre en oeuvre les dispositifs expérimentaux classiques en agronomie ; savoir en traiter les données et en présenter les résultats
- Savoir choisir et mettre en oeuvre les principales méthodes multivariées pour l'exploration de données statistiques

## Public cible

Techniciens ou ingénieurs en agriculture et environnement

## Equipe pédagogique

- > Jean-Pierre Da Costa  
Professeur, Bordeaux Sciences Agro
- > Lionel Bombrun  
Maître de Conférences, Bordeaux Sciences Agro

# Contenu de la formation

Les parties A, B et C, sont dissociables, mais la partie A est un prérequis des parties B et C.

## ● **Partie A : Auto-formation à R et aux statistiques descriptives** (14h, à son rythme)

### **Jour 1 – Découverte de R et RStudio**

#### **I. Prendre en main le langage R et l'environnement de développement RStudio**

- Télécharger, installer, découvrir R et RStudio
- Editer, manipuler un script R
- Manipuler des variables
- Manipuler des fonctions, utiliser l'aide en ligne
- Installer des Packages

#### **II. Manipuler un jeu de données avec R**

- Créer et importer un fichier CSV sur R
- Visualiser un jeu de données individus/variables
- Extraire un sous ensemble
- Modifier, sauvegarder un jeu de données.

### **Jour 2 – Statistique descriptive uni et bivariee**

#### **III. Mettre en oeuvre des outils de statistique descriptive univariée**

- Connaître, identifier, modifier les types de variables
- Présenter une variable qualitative sous forme de tableau ou de graphique (diagrammes en barres, en secteurs)
- Présenter une variable quantitative sous forme de résumés numériques (paramètres de tendance centrale et de dispersion, statistiques de rang, mise en classe) ou graphiques (histogrammes, box-plots)

#### **IV. Mettre en oeuvre des outils de statistique descriptive bivariee**

- Croiser deux variables qualitatives : produire et normaliser un tableau de contingence, le représenter graphiquement à l'aide de diagrammes en barres.
- Croiser une variable qualitative et une variable quantitative à l'aide de résumés numériques et graphiques (e.g. box-plots)
- Croiser deux variables quantitatives à l'aide de résumés numériques (covariance et coefficients de corrélation, paramétriques et non paramétriques) et graphiques (nuages de points)

## ● **Partie B – Statistique et expérimentation** (2 jours, en présentiel)

### **Jour 3 – Principes d'échantillonnage, d'estimation et de tests hypothèses**

#### **V. Echantillonnage et estimation**

- Notions de moyenne, d'écart-type, de variance ; lois usuelles (normale, binomiale...)
- Estimation ponctuelle d'une moyenne, d'une proportion, d'une variance
- Estimation par intervalle de confiance d'une moyenne, d'une proportion, d'une variance

### **Jour 4 – ANOVA et dispositifs en expérimentation agricole**

#### **VII. Analyse de Variance (ANOVA)**

- Mettre en oeuvre et analyser une ANOVA à 1 facteur et ses tests post-hoc
- Savoir mettre en oeuvre et interpréter une ANOVA à 2 facteurs avec et sans interaction
- Savoir vérifier les conditions d'applications, appliquer une transformation ou choisir une alternative non paramétrique

#### **VIII. Dispositifs expérimentaux**

- Connaître les dispositifs expérimentaux usuels, savoir les mettre en oeuvre et en traiter les résultats

## ● **Partie C – Analyse de données multivariées** (1 jour, en présentiel)

### **Jour 5 – Méthodes multivariées pour l'analyse de variables quantitatives**

#### **IX. Analyse en composantes principales (ACP)**

- Connaître les objectifs et le principe d'une ACP, savoir la mettre en oeuvre et en interpréter les résultats

#### **X. Classification non supervisée (K-means et classification ascendante hiérarchique)**

- Connaître les principales méthodes de classification non supervisée. Savoir choisir, mettre en oeuvre ces méthodes et en interpréter les résultats



# Contenu de la formation

...suite

## Option

dans le cadre d'une formation intra-entreprise, sur mesure

- **Partie D – Support à distance post formation**

A l'issue de la formation, une assistance est apportée aux stagiaires afin de faciliter l'appropriation et la mise en pratique des outils et méthodes abordées dans leur contexte professionnel. L'équipe de formateurs se rendra disponible pour répondre aux sollicitations des stagiaires. L'assistance sera apportée, de façon asynchrone, sur un forum dédié mis en place sur la plate-forme pédagogique de l'établissement. Les supports de formations et autres supports pédagogiques seront mis à disposition des stagiaires sur cette même plate-forme.



## Contact et inscriptions

**Inscriptions et financements**  
formco@agro-bordeaux.fr  
05 57 35 07 67

## Informations pratiques

### Durée

5 jours (35h), 2 j en distanciel (14h à son rythme) + 3 j en présentiel (21h)

### Modalités et Lieu

Hybride

### Pré-requis

Notions élémentaires en statistiques et traitement de données

### Modalités d'évaluation

Test de positionnement, évaluation QCM de l'atteinte des objectifs, questionnaire de satisfaction

### Modalités pédagogiques

**Partie distancielle (asynchrone) :**

- utilisation de la plateforme pédagogique Moodle
- assistance asynchrone de l'équipe pédagogique sur un forum

**B+C : Parties présentiels (synchrone) :**

- alternance d'apports théoriques et de mise en pratique sur logiciel R et RStudio

Dans les deux cas, auto-évaluations sous formes de QCM (sur Moodle).

### Nombre de participants par session

de 6 à 15

### Dates et Tarifs

- > Consultez notre site internet, rubrique **Formation Continue**
- > **Modalités de financement :** plan de développement des compétences, OPCO, France Travail, financement personnel



Suivez-nous sur notre page **LinkedIn**  
«Executive - Bordeaux Sciences Agro»

[www.agro-bordeaux.fr](http://www.agro-bordeaux.fr)

