Formation Réaliser ses analyses statistiques avec R

Objectifs : Apprendre à utiliser le logiciel R pour analyser des données. Mettre en oeuvre dans R les méthodes de statistique descriptive, décisionnelle, analyse de la variance, régression linéaire et analyse de données multidimensionnelles.

Compétences visées : - Importer, manipuler et analyser des données dans R
- Effectuer des analyses statistiques univariées
- Établir avec R des intervalles de confiance autour de paramètres comme moyenne, proportion et variance
- Réaliser les principaux tests d'hypothèses paramétriques et non paramétriques dans le logiciel
- Étudier numériquement et graphiquement la liaison entre 2 variables quantitatives, qualitatives ou de chaque nature
- Mettre en œuvre des analyses de la variance à un ou plusieurs facteurs pour la comparaison de moyennes de population
- Élaborer un modèle établissant une relation linéaire entre une variable à expliquer et une (régression simple) ou plusieurs (régression multiple) variables explicatives.
- Découvrir les principales méthodes d'analyse multidimensionnelle des données à l'aide du package FactoMineR

Durée : 3 jours (21 heures)

Public : Toute personne souhaitant analyser des données avec R

Pré-requis : Pour suivre ce stage dans de bonnes conditions, il est recommandé d'avoir suivi en amont la formation R – Prise en main, analyses statistiques et graphiques

Méthode pédagogique : Pédagogie active mêlant exposés, exercices et applications pratiques dans le logiciel R.

Tarifs :
- Présentiel : 1650 € HT
- Distanciel : 1500 € HT
(-10% pour 2 inscrits, -20% dès 3 inscrits)

Option(s) :
- Forfait déjeuners : 75 € HT

Nos prochaines sessions

Bordeaux
du 23 au 25 juin 2027

Distance

du 10 au 12 mars 2027
du 16 au 18 juin 2027
du 3 au 5 novembre 2027

Lyon
du 30 mars au 1 avr. 2027
du 29 sept. au 1 oct. 2027

Paris
du 9 au 11 décembre 2026
du 10 au 12 mars 2027
du 2 au 4 juin 2027
du 1 au 3 septembre 2027
du 8 au 10 décembre 2027

Toulouse
du 10 au 12 mai 2027
du 13 au 15 octobre 2027

Programme :

- Statistiques descriptives

- Gestion d'un jeu de données - dataframe
  Importation, caractérisation, sélection, sous-ensembles
- Premières analyses d'un jeu de données
  Premières vérifications, valeurs manquantes, recodage
- Résumé d'une variable quantitative – numeric
  Indicateurs numériques, représentations graphiques
- Résumé d'une variable qualitative – factor
  Indicateurs numériques, représentations graphiques

- Intervalle de confiance

- Le raisonnement à partir d'un échantillon
  Généralités, échantillonnage, estimation d'un paramètre
- Intervalle de confiance d'une moyenne
- Intervalle de confiance d'une proportion
- Intervalle de confiance d'une variance

- Tests d’hypothèses

- Qu’est-ce qu’un test d’hypothèse ?
  Généralités, règle de décision, risques d’erreur, puissance
- Les tests de conformité ou de comparaison à une norme
  Conformité d’une moyenne, d’une proportion
- Les tests de comparaison de deux populations
  Comparaison de deux moyennes, de deux proportions
- Test d’ajustement à une loi de probabilité normale
  Le test de Shapiro-Wilk
- Introduction aux tests non paramétriques
  Test de Wilcoxon, Mann et Whitney, Kruskal Wallis, Friedman

- Liaisons entre deux variables

- Liaison entre deux variables quantitatives
  Nuage de points, intensité de la liaison, significativité
- Liaison entre deux variables qualitatives
  Tableau de contingence, intensité et significativité du lien de dépendance : test du khi2
- Liaison entre une variable qualitative et quantitative
  Comparaison de plusieurs populations, le rapport de corrélation
- Liaisons entre plusieurs variables
  Approches graphiques : matrice de nuages de points, treillis
  Caractériser des sous-populations par plusieurs variables

- L’analyse de la variance – Anova

- Analyse de la variance à un facteur
  Variabilité inter, intra, totale – Rapport de corrélation - Le test de Fisher
- Comparaisons multiples de moyennes
  Analyses post hoc, la procédure de Tukey
- Analyse de la variance à deux facteurs et interaction
- Extensions de l’Anova
  Modèle à effet fixe ou aléatoire, modèle hiérarchisé

- Régression linéaire simple et multiple

- De la corrélation à la régression
  L’intérêt d’un modèle - Variables à expliquer, explicatives, erreur

- La régression linéaire simple
  Ajustement par la méthode des moindres carrés - Tests et validation du modèle
- La régression linéaire multiple
- Choix d'un modèle de régression
  Pourquoi sélectionner un sous-ensemble de variables explicatives ?
  Les différentes approches et critères de sélection d'un modèle

- Analyse de données multidimensionnelles

- Un panorama des méthodes multidimensionnelles
  Analyses factorielles, classification – Le package FactoMineR
- ACP : Analyse en Composantes principales
- AFC : Analyse Factorielle des Correspondances
- ACM : Analyse des correspondances Multiples
- CAH : Classification Ascendante Hiérarchique

Modalités d'évaluation : Un formulaire d'auto-évaluation proposé en amont de la formation nous permettra d'évaluer votre niveau et de recueillir vos attentes. Ce même formulaire soumis en aval de la formation fournira une appréciation de votre progression.
Des exercices pratiques seront proposés à la fin de chaque séquence pédagogique pour l'évaluation des acquis.
En fin de formation, vous serez amené(e) à renseigner un questionnaire d'évaluation à chaud.
Une attestation de formation vous sera adressée à l'issue de la session.
Trois mois après votre formation, vous recevrez par email un formulaire d'évaluation à froid sur l'utilisation des acquis de la formation.

Qualification du formateur : Les formations DATA VALUE sont animées par des intervenants experts, alliant expérience opérationnelle et compétences pédagogiques reconnues. Ils disposent de plus de 10 ans d'expérience professionnelle et assurent une veille constante sur les pratiques, outils et normes de leur domaine afin de garantir des contenus actualisés et adaptés aux besoins des stagiaires.

Délais d'accès et modalités de formation : Les inscriptions peuvent être effectuées jusqu'à 48 heures avant le début de la formation.
Vous pourrez trouver des informations plus détaillées concernant les modalités de nos formations à cette adresse : https://www.datavalue.fr/modalites-acces-formation.

Accessibilité : Vous souhaitez suivre notre formation Réaliser ses analyses statistiques avec R et êtes en situation de handicap ? Merci de nous contacter afin que nous puissions envisager les adaptations nécessaires et vous garantir de bonnes conditions d'apprentissage

Date de dernière modification : 30 octobre 2026