

## Formation en STATISTIQUES

### « Logiciel R pour analyses statistiques avancées »

<p><b>Dates et durée</b> Du 13 au 15 novembre 2019  (3 jours)</p> <p><b>Lieu</b> Nantes</p> <p><b>Public</b> Toute personne souhaitant se perfectionner sur R dans un contexte de mise en œuvre d'analyses statistiques de données</p> <p><b>Pré-requis</b> &gt; Connaissances de base sur R (importation données, création et exécution scripts simples, création et manipulation vecteurs et data frames ...) &gt; Bonnes connaissances sur analyses statistiques telles que : analyses descriptives, tests d'hypothèses (Student ...), ANOVA, régression linéaire, analyses multifactorielles, classification</p> <p><b>Intervenant</b> Formateur de la Sté Arkesys</p>	<p><b>OBJECTIFS :</b> Se perfectionner dans l'utilisation du logiciel R pour la mise en œuvre d'analyses statistiques de données</p> <ul style="list-style-type: none"> <li>▪ Manipuler efficacement R (facteurs, data frames)</li> <li>▪ Filtrer un tableau de données</li> <li>▪ Créer des tableaux de synthèses (de moyennes, d'écart-types, par groupe)</li> <li>▪ Lire et écrire des scripts avancés d'analyse de données</li> <li>▪ Mettre en œuvre et extraire les résultats des méthodes statistiques suivantes :             <ul style="list-style-type: none"> <li>- Analyse descriptive de données</li> <li>- Tests paramétriques et non paramétriques classiques (Student, Wilcoxon)</li> <li>- Anova à 1 facteur et ANOVA multi-facteurs</li> <li>- Tests post-Hoc de Tukey, de Bonferroni</li> <li>- Régression linéaire</li> <li>- Analyses factorielles multivariées (ACP, AFC), et classification (CAH)</li> </ul> </li> </ul> <p><b>PROGRAMME :</b></p> <ul style="list-style-type: none"> <li>➤ <b>Revoir les fondamentaux de R :</b> Fonctionnement général et documentation en ligne ; Organisation d'un projet R ; Importation de données Excel ; Installation et chargement d'un package ; Trucs et astuces</li> <li>➤ <b>Structurer et extraire de l'information d'un tableau de données :</b> Création et manipulation efficace de variables qualitatives (facteurs) et quantitatives ; Ajout ou suppression d'une variable ; Empilement/dépilement d'un tableau ; Filtres (à partir d'indices des éléments ou par le biais de conditions)</li> <li>➤ <b>Synthétiser l'information d'un tableau de données :</b> Commande tapply ; Commande aggregate ; Création de tableaux de synthèse/croisés</li> <li>➤ <b>Mettre en œuvre des analyses statistiques de base sous R :</b> Analyse descriptive de données ; Intervalles de confiance ; Tests d'hypothèses paramétriques et non paramétriques (Test de Student pour comparaison de moyennes, Tests de Wilcoxon et Mann-Whitney pour comparaison de médianes, Test de normalité Shapiro-Wilk et QQ-plot)</li> <li>➤ <b>Mettre en œuvre une ANOVA sous R :</b> Commandes aov() et lm() ; ANOVA à un facteur (tableau d'analyse de variance, vérification des conditions du modèle sur la normalité et l'homogénéité des dispersions (test de Bartlett et Levene), représentation graphique des résultats, Test post-Hoc pour comparaisons multiples avec méthode de Tukey et correction Bonferroni, Tests non paramétriques Kruskal-Wallis) ; ANOVA multi-facteurs (effets simples et croisés, graphique des interactions, somme des carrés de type I et III)</li> <li>➤ <b>Mettre en œuvre une régression linéaire sous R :</b> Ecriture du modèle avec la commande lm() ; Commandes summary() et anova() ; Régression linéaire simple et multiple ; Tableau d'analyse de variance et estimation des coefficients ; Vérification des conditions (normalité ou homogénéité des résidus) ; Diagnostic des multi colinéarités</li> <li>➤ <b>Mettre en œuvre une analyse factorielle multivariée et une classification sous R :</b> Principaux package ; Analyse en Composantes Principales (ACP) avec diagramme des valeurs propres, cercle factoriel, graphique des individus, contributions, cosinus<sup>2</sup> ; Analyse Factorielle des Correspondances (AFC) avec graphique des modalités ; Classification Ascendante Hiérarchique (CAH) avec commande hcpc(), dendrogramme, récupération des classes</li> </ul> <p><b>METHODE</b> Explications théoriques suivies de pratiques guidées puis mises en autonomie (1 ordi par stagiaire avec R). Les supports et exercices pourront être récupérés sur clé USB par les stagiaires</p>
--	--

**Date limite d'inscription : 4 octobre 2019 sur Sirene :** <https://www.sirene.inserm.fr/>

(onglet « Agent formation », menu « Demander une formation », domaine « TS4 – Statistiques)

**CONTACT :** Michèle HAYS - Responsable Formation - INSERM DR Grand Ouest

63, quai Magellan – CS 32116 – 44021 NANTES Cedex 1 – Tél : 02 40 35 86 80 – Mail : [michele.hays@inserm.fr](mailto:michele.hays@inserm.fr)