Analyse statistique avancée avec R

Réalisation d'analyses statistiques spécifiques

Présentiel ou en classe à distance

3 jours (21 h)

Prix inter : 2.190,00 € HT
Forfait intra : 7.650,00 € HT

Réf.: BI104

L'analyse statistique des données est une des compétences requises pour mettre en oeuvre des projets Big Data. Selon la nature des données manipulées et le type d'analyses souhaitées, il est parfois nécessaire de recourir à des techniques d'analyses avancées. Aussi, tout Data Scientist doit-il aujourd'hui maitriser l'implémentation d'outils statistiques sous R pour réaliser des analyses non ponctuelles et inférentielles paramétriques ou de données uni ou multivariées pour tous les domaines dans le service, l'industrie ou encore la R&D. Cette formation avancée permettra aux participants de maitriser R et ainsi d'être à même d'analyser tous types de données dans leurs projets Big Data.

A qui s'adresse cette formation ?

- Ingénieurs
- Analystes
- Data Analysts
- Toute personne intéressée par l'analyse statistique avec R

- Avoir suivi la formation "Les fondamentaux de l'analyse statistique avec R" (BI103) ou connaissances équivalentes
- Avoir une première expérience de la programmation
- Disposer de solides connaissances théoriques en statistiques
- Disposez-vous des connaissances nécessaires pour suivre cette formation ? Testez-vous !

Programme

Pendant

1 - Introduction

- Générer et représenter des variables aléatoires
- Puissance d'un test Z à 1 échantillon en bilatéral
- Taille d'échantillon (effectif) d'un test Z en bilatéral
- Puissance d'un test t à 1 échantillon en bilatéral

2 - Travailler avec des échantillons

- Taille d'échantillon (effectif) d'un test t à 1 échantillon en bilatéral
- Taille d'échantillon (effectif) d'un test p (proportion) à 1 échantillon en bilatéral
- Taille d'échantillon (effectif) d'un test p (proportion) à 2 échantillons en bilatéral

3 - Réaliser des tests d'ajustement

- Test d'ajustement d'Anderson-Darling (ie Agostino-Stephens)
- Test d'ajustement de Shapiro-Wilk

4 - Estimation et intervalles de confiance

- Intervalle de confiance de la moyenne (test Z à un 1 échantillon)
- Intervalle de confiance de la moyenne (test t à un 1 échantillon)

5 - Analyses statistiques avancées

- Test t-Student bilatéral d'un échantillon
- Test t-Student pour données appariées
- Test t-Student homoscédastique bilatéral d'égalité de la moyenne
- Test t-Student hétéroscédastique bilatéral d'égalité de la moyenne (test de Welch)
- Test de Poisson à un échantillon unilatéral/bilatéral
- Comparaison de proportions sur une même population (test binomial exact)
- Intervalle de confiance de la proportion
- Comparaison de proportions sur 2 échantillons indépendants
- Test de Fisher d'égalité des variances
- Test de Levene d'égalité de deux variances
- Robustesse de tests statistiques
- Transformations de Box-Cox
- Transformations de Johnson

6 - Analyse de la variance et de la covariance

- ANOVA à un facteur fixe (ANOVA-1 canonique) désempilé
- ANOVA à deux facteurs fixes (ANOVA-2 canonique) sans répétitions avec ou sans interactions
- ANOVA à deux facteurs fixes (ANOVA-2 canonique) avec répétitions avec ou sans interactions
- Comparaisons multiples du test de Student avec correction de Bonferroni
- Test de (l'étendue) de Tukey HSD
- Test de Levene et Bartlett d'égalité des variances d'une ANOVA canonique
- ANOVA Imbriquée/Hiérarchique complète
- ANOVA Carré Latin
- ANCOVA (Analyse de la Covariance)
- MANOVA
- ACP (Analyses en Composantes Principales) paramétrique
- Analyse factorielle exploratoire (AFE)
- AFE avec méthode ACP sans rotation
- AFE avec méthode ACP et rotation VariMax



Les objectifs de la formation

- Connaître les outils et méthodes d'analyses statistiques avancées en environnement R
- Être capable de programmer des analyses avec R
- Savoir utiliser les packages de R pour mettre en oeuvre des modélisations statistiques (régression, ACP...)
- Comprendre comment réaliser des analyses prédictives à l'aide d'équations de régression



Evaluation

Pendant la formation, le formateur évalue la progression pédagogique des participants via des QCM, des mises en situation et des travaux pratiques.
Les participants passent un test de positionnement avant et après la formation pour valider leurs compétences acquises.



Les points forts de la formation

- Une formation très pratique : 90% du temps de la formation est dédié à la mise en pratique pour une meilleure assimilation de notions de base
- Cette formation est basée sur des exercices principalement proposés par le formateur et tirés de l'ouvrage qui sert de support pour la formation
- Des consultants expérimentés partagent leur savoir-faire avec les participants.
- 83% des participants à cette formation se sont déclarés satisfaits ou très satisfaits au cours des 12 derniers mois.

Evaluation

Pendant la formation, le formateur évalue la progression pédagogique des participants via des QCM, des mises en situation et des travaux pratiques. Les participants passent un test de positionnement avant et après la formation pour valider leurs compétences acquises.

Dates et villes 2027 - Référence BI104

Dernières places disponibles Session garantie

A distance

du 22 févr. au 24 févr.	du 20 sept. au 22 sept.
du 24 mai au 26 mai	du 15 nov. au 17 nov.

Paris

du 22 févr. au 24 févr.	du 20 sept. au 22 sept.
du 24 mai au 26 mai	du 15 nov. au 17 nov.