

FORMATION STATISTIQUE ET ANALYSE DE DONNEES

REF. : STA101

DESCRIPTION

Cette formation statistique et analyse de données de 3 jours vous permet de découvrir et maîtriser les bases de la statistique et de l'analyse de données.

Vous avez besoin de comprendre les analyses statistiques de bases et interpréter les résultats, cette formation répond à ce besoin. Il s'agit d'une formation pratique qui s'appuie sur des exemples concrets afin d'illustrer les logiques de la statistique et de l'analyse de données.

Cette formation se base sur différents outils dont la maîtrise n'est pas nécessaire (Excel, R, XLSTAT...).

DUREE

3 jours

PUBLIC ET PREREQUIS

Aucune connaissance préalable

MOYENS PEDAGOGIQUES

Alternance d'exposés et d'applications pratiques avec des exercices sur des données.

OBJECTIFS

- Connaître les concepts de base en statistique
- Maîtriser les principes du traitement des données
- Savoir quelle méthode utiliser en fonction de votre problématique
- Savoir lire et interpréter les résultats obtenus par les différents outils

PROGRAMME DETAILLE

Jour 1 :

- Matin :
 - Introduction à la statistique et à l'analyse de données
 - Les principes de bases : échantillonnage, population vs. échantillon...
 - La préparation des données (transformation, traitement des données manquantes, filtrage...)
 - Cas du traitement d'une enquête
- Après midi :
 - Les statistiques descriptives pour décrire vos données (moyenne, variance, médiane, mode...)
 - Les intervalles de confiance et les tests statistiques
 - Les tests de comparaison de moyennes
 - Les tests paramétriques et non paramétriques
 - Cas pratiques : applications sur des données réelles

STAT4DECISION

37-39 avenue Ledru-Rollin – 75012 Paris – France

Tel. 01.72.25.40.82 – E-mail : info@stat4decision.com – www.stat4decision.com

SIRET : 810 489 856 00015 – Numéro de déclaration d'activité : 11 75 53522 75

stat4decision est validé dans le DATA-DOCK

Jour 2 :

- Matin :
 - Principes de l'analyse de données multivariée
 - L'analyse en composantes principales
 - Les autres méthodes d'analyse factorielles (AFC, ACM, analyse discriminantes...)
- Après midi :
 - Présentation des méthodes de clustering (apprentissage non supervisé)
 - La classification ascendante hiérarchique
 - Les k-means

Jour 3 :

- Matin :
 - Les principes de la modélisation
 - Le modèle linéaire : régression, analyse de la variance
 - La régression logistique
- Après midi
 - Vers le machine learning et le data mining
 - Les principes et les différences avec la statistique classique
 - Quelques exemples : Les arbres, les forêts...