

Formation Modélisation statistique, L'essentiel

Présentation

Ce cours présente l'essentiel de la modélisation statistique. Il vous permettra de comprendre son rôle dans le monde de l'analyse décisionnelle, du Big Data et du Data Mining ainsi que les mécanismes qui permettent de transformer et d'affiner des données pour en tirer des informations métiers utiles.

Durée : 14,00 heures (2 jours)

Tarif INTRA : Nous consulter

Objectifs de la formation

- Connaître les fondamentaux de l'analyse statistique appliquée,
- Maîtriser l'utilisation des tests et formules statistiques de base,
- Savoir concevoir des rapports analytiques factuels,
- Explorer les paramètres statistiques pour comprendre les séries de données,
- Explorer des outils comme R et Excel pour implémenter des modèles étudiés,
- Valider l'exactitude de l'estimation, à l'aide d'intervalles de confiance,
- Prédire les comportements futurs,
- Savoir comment vérifier l'ajustement du modèle.

Prérequis

- Connaissances générales en mathématiques, analyse statistique et Excel.

Public

- Utilisateurs et gestionnaires métiers de bases de données,
- Data Scientist,
- Ingénieurs,
- Data Analysts,
- Toute personne intéressée par l'analyse statistique appliquée.

Programme de la formation

Généralités sur les modèles

1/5

Docaposte Institute

ACI 3A3-303 - 45-47 Boulevard Paul Vaillant Couturier, 94200 IVRY SUR SEINE - Email :
hello.institute@docaposte.fr

SIRET : 813 058 542 00032 - Code APE : 8559A - N TVA Intra : FR 58 813 058 542

Déclaration d'activité enregistrée sous le numéro 11941117994 auprès du préfet de région d'Ile-de-France



- Une représentation du réel
- Buts et utilités d'un modèle
- Domaines d'applications
- Modèles mathématiques
- Modélisation stochastique

Domaines d'utilisation

- Statistiques officielles
- Presse, médias
- Banques – assurances
- Sciences de la vie
- Environnement (foresterie, pêche ...)
- Santé
- Sciences humaines
- Entreprises & industrie
- Finance
- Recherche fondamentale et appliquée

Techniques de modélisation

- Analyse de données
- Probabilités / statistiques
- Théorie de l'information
- Intelligence artificielle
- Base de données
- Datavisualisation

Modèles préliminaires

- Nettoyage des données
- Exploration des données
- Transformation / regroupement de variables

Méthodes statistiques non paramétriques

- Méthodes d'estimation
- Tests non paramétriques
- Test d'égalité de distributions Kolmogorov-Smirnov
- Statistique semi-paramétrique

Modèles paramétriques classiques

- Modèle linéaire (gaussien) de base
- Modèle linéaire généralisé
- Modèles linéaires généralisés
- Modèles mixtes

Paramètre de position et de dispersion

- Mode, valeur modale, valeur la plus probable
- Moyenne d'une population (ou d'un échantillon)
- Médiane, partager une série numérique
- Etendue, différence entre valeurs extrêmes.
- Utiliser les quantiles
- Ecart-Type, calculer la dispersion d'un ensemble de données.
- Calcul de la variance et de la covariance

Etude de cas : Calcul de paramètres de position et de dispersion sur différents échantillonnages et comparaisons des résultats.

Tests et intervalle de confiance

- Lois statistiques et intervalle de confiance.
- Tests statistiques courants (Test de Student, Analyse de variances, Khi^2).
- Valider la précision d'une estimation.
- Amplitude de l'intervalle.

Etude de cas : Exercices sur le logiciel R.

Organisation

Formateur

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes.

Moyens pédagogiques et techniques

- Apports des connaissances communes.
- Mises en situation sur le thème de la formation et des cas concrets.
- Méthodologie d'apprentissage attractive, interactive et participative.

- Equilibre théorie / pratique : 60 % / 40 %.
- Supports de cours fournis au format papier et/ou numérique.
- Ressources documentaires en ligne et références mises à disposition par le formateur.
- Pour les formations en présentiel dans les locaux mis à disposition, les apprenants sont accueillis dans une salle de cours équipée d'un réseau Wi-Fi, d'un tableau blanc ou paperboard. Un ordinateur avec les logiciels appropriés est mis à disposition (le cas échéant).

Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation

En amont de la formation

- Recueil des besoins des apprenants afin de disposer des informations essentielles au bon déroulé de la formation (profil, niveau, attentes particulières...).
- Auto-positionnement des apprenants afin de mesurer le niveau de départ.

Tout au long de la formation

- Évaluation continue des acquis avec des questions orales, des exercices, des QCM, des cas pratiques ou mises en situation...

A la fin de la formation

- Auto-positionnement des apprenants afin de mesurer l'acquisition des compétences.
- Evaluation par le formateur des compétences acquises par les apprenants.
- Questionnaire de satisfaction à chaud afin de recueillir la satisfaction des apprenants à l'issue de la formation.
- Questionnaire de satisfaction à froid afin d'évaluer les apports ancrés de la formation et leurs mises en application au quotidien.

NB : dans le cadre d'une Action collective, chaque stagiaire bénéficiaire sera contacté par un prestataire choisi par l'Opco Atlas afin d'évaluer « à chaud » la qualité de la formation suivie.

Accessibilité

Nos formations peuvent être adaptées à certaines conditions de handicap. Nous contacter pour toute information et demande spécifique.

