

Mis à jour le 18/03/2024

S'inscrire

Formation Modélisation Statistique

2 jours (14 heures)

Présentation

La **modélisation statistique** est un processus d'application de l'analyse statistique à un ensemble de données. Elle consiste à utiliser des modèles mathématiques et des hypothèses statistiques pour générer des données d'échantillon et faire des prédictions sur le monde réel.

Un **modèle statistique** est une collection de distributions de probabilité représentant tous les résultats possibles d'une expérience. Durant notre formation vous verrez les différents modèles statistiques tels que les modèles linéaires ou les modèles ANOVA.

L'application de la modélisation statistique aux données brutes permet aux spécialistes de la Data à aborder l'analyse des données de manière stratégique. Elle leur fournit des visualisations intuitives aidant à identifier les relations entre les variables et à prédire leur résultat.

Notre formation Modélisation statistique, l'essentiel vous présentera les bases de l'analyse statistique appliquée, des formules et des tests statistiques. Vous manipulerez les outils R et Excel. Vous saurez concevoir un rapport d'analyse basé sur les faits, exploiter des paramètres statistiques et valider la précision d'une estimation.

À l'issue de cette formation, vous serez capable de mettre en œuvre une analyse des données statistiques, descriptives et prévisionnelles.

Objectifs

- Maîtriser les bases de l'analyse statistique appliquée
- Appliquer des formules statistiques
- Utiliser des tests statistiques fondamentaux
- Savoir concevoir un rapport d'analyse basé sur les faits
- Comprendre une série de données en utilisant les paramètres statistiques
- Découvrir des outils R et Excel afin de mettre en œuvre des modèles étudiés
- Confirmer la précision d'une estimation avec des intervalles de confiance

- Être capable de prévoir les comportements suivants
- Savoir vérifier l'adéquation à un modèle

Public visé

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

Pré-requis

Connaissances générales en mathématiques, analyse statistique et Excel

Programme de notre formation Modélisation statistique, l'essentiel

Introduction à la modélisation statistique

- Qu'est-ce que la modélisation statistique ?
- Modélisation mathématique
- Définir la statistique descriptive
- Méthodes d'échantillonnage
- Inférence statistique

Méthodes paramétriques

- Régression linéaire simple/multiple
 - Modèle
 - Moindres carrés
 - Estimations
 - Intervalles de confiance
 - Sélection de variables
- Analyse de la variance
- Analyse de la covariance
- Comparaison avec Anova
- Comparaison avec la régression linéaire
- Modèle linéaire généralisé
 - Régression de Poisson
 - Modèle polytomique
- Modélisation Bayésienne
- Analyse de séries temporelles

Méthodes non paramétriques

- Régression spline
- Estimation par moyennes locales
- Estimation à noyau
- Régression polynomiale locale

Variables quantitatives polytomiques

- Variables à modalités ordonnées
- Variables à modalités non ordonnées
- Variables à choix emboîtés
- Formalisation et exemples
- Cas d'applications des variables

Variables quantitatives binaires

- Estimation et interprétation des paramètres
- Valider les résultats
- Exemples et formalisation
- Cas d'application

Nettoyage des données

- Repérer d'éventuels codes superflus
- Repérer des données manquantes en grande quantité
- Suppression de variables
- Analyses univarié
- Analyse bivarié

Intervalles de confiance

- Estimation et intervalles de confiance
- Lois statistiques d'intervalle de confiance
- Valider une précision d'une estimation

Paramètres statistiques

- Exploiter des paramètres statistiques
- Comprendre une série de données
- Aspects statistiques de modélisation
 - Validation du modèle
 - Tests de coefficients
 - Étude des résidus et des points influents
- Vérification d'adéquation à un modèle

Modélisation avec Python et Excel

- Mettre en œuvre des méthodes de régression
- Analyser la variance sous Python
- Traiter une question concrète à travers d'une approche de modélisation
- Mise en œuvre de modélisation avec Python

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.