# XLSTAT Analyses multivariées, classification et PLS (ACP, CAH, K-MEANS, AFD, PLS)

*Lien : https://innov-systems.com/formation/xlstat-analyses-multivariees-classification-et-pls-acp-cah-k-means-afd-pls*

DURÉE
**5 jours (35h)**

RÉFÉRENCE
**BSI570**

CATÉGORIE
**Outils Décisionnels : Essbase, Hyperion, Matlab, Jaspersoft, Talend, Qlik, Cognos,...**

Innov Systems

## OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Identifier la méthode d’analyse multivariée ou de classification la plus adaptée selon le contexte métier
- Comprendre les concepts mathématiques et statistiques associés aux méthodes multivariées
- Mettre en œuvre et interpréter les principales analyses : ACP, CAH, K-means, AFD, PLS
- Différencier les notions de contribution, cosinus carré et VIP
- Maîtriser les critères et indicateurs de qualité d’une analyse multivariée
- Optimiser le choix et l’utilisation des modèles de régression PLS et PLS-DA
- Développer une autonomie dans l’analyse et l’interprétation des sorties graphiques et statistiques
- Comparer et combiner différentes approches pour des problématiques complexes

## POUR QUI ?

- Analystes de données et statisticiens
- Ingénieurs qualité et responsables R&D
- Chercheurs et doctorants utilisant XLSTAT dans leurs études
- Professionnels en marketing, sciences sociales, santé, agroalimentaire, environnement
- Toute personne souhaitant approfondir l’analyse multivariée et la modélisation prédictive avec XLSTAT

Innov Systems

## Programme détaillé

### 1 / INTRODUCTION À XLSTAT ET AUX DONNÉES

- Présentation générale et interface utilisateur XLSTAT
- Activation, chargement et organisation des modules
- Gestion et préparation des données : codage, regroupement, transformation
- Complémentarité entre Excel et XLSTAT
- Bonnes pratiques de structuration des jeux de données

### 2 / PRINCIPES DES ANALYSES MULTIDIMENSIONNELLES

- Limites des approches statistiques classiques
- Champs d'application et objectifs : description, prédiction, exploration
- Concepts fondamentaux : distance, inertie, variance, axes factoriels, variables latentes
- Présentation de l'éventail des méthodes : ACP, AFC, AFD, CAH, K-means, PLS
- Introduction au data mining et au machine learning supervisé vs non supervisé

### 3 / ANALYSE EN COMPOSANTES PRINCIPALES (ACP) ET VARIANTES

- Objectifs et contexte d'application
- Choix du nombre de composantes principales (critères et validation)
- Interprétation des graphiques : cercle des corrélations, projection des individus
- Contribution et qualité de représentation ($cos^2$)
- Comparaison ACP normée vs non normée, ACP non paramétrique
- Variables illustratives et interprétation avancée
- Différences conceptuelles avec l'AFC
- Atelier pratique avec données réelles

### 4 / MÉTHODES DE CLASSIFICATION : CAH ET K-MEANS

Innov Systems

- CAH : choix des distances et des méthodes de liaison, lecture et interprétation du dendrogramme
- Détermination du nombre de classes et validation des résultats
- K-means : principe de l'algorithme, détermination des clusters optimaux
- Comparaison des performances CAH vs K-means
- Classification sur individus et sur variables
- Utilisation de l'ACP en complément de la classification
- Applications pratiques sur grand jeu de données
- Conseils d'interprétation et erreurs à éviter

## 5 / ANALYSE FACTORIELLE DISCRIMINANTE (AFD)

- Notions de classification et discrimination
- Étapes méthodologiques de l'AFD
- Comparaison avec l'ACP et apports complémentaires
- Interprétation des résultats : axes discriminants, graphiques, corrélations
- Évaluation de la qualité : lambda de Wilks, tests multivariés
- Matrice de confusion et courbe ROC
- Bonnes pratiques et pièges à éviter

## 6 / RÉGRESSION PARTIELLE PAR LES MOINDRES CARRÉS (PLS) ET PLS-DA

- Problématiques des données colinéaires et petits échantillons
- Présentation et comparaison avec PCR et régression Ridge
- Méthodologie générale de la PLS et de la PLS-DA
- Critères de sélection du nombre de composantes ($Q^2$, validation croisée)
- Interprétation des sorties : coefficients, VIP, cercle factoriel
- Indicateurs de qualité : $R^2$, $Q^2$, qualité d'ajustement et de prédiction
- Sélection et importance des variables
- Applications pratiques avec données de prédiction et classification

## 7 / ATELIERS DE SYNTHÈSE ET MISE EN PRATIQUE

Innov Systems

- Études de cas sur des jeux de données réels (marketing, industrie, santé…)
- Comparaison ACP, CAH, K-means et PLS pour une même problématique
- Élaboration d’une démarche d’analyse complète : préparation, choix de méthode, validation et interprétation
- Conseils méthodologiques et stratégie d’application en entreprise
- Discussion sur l’intégration des méthodes XLSTAT dans des projets de data science

## Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

## Prochaines dates programmées

| | |
|---|---|
| 23 au 27 Nov. 2026 | Distanciel |
| 15 au 19 Fév. 2027 | Présentiel - Casablanca |
| 12 au 16 Avr. 2027 | Présentiel - Casablanca |
| 07 au 11 Juin 2027 | Présentiel - Casablanca |
| 02 au 06 Août 2027 | Présentiel - Casablanca |

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Innov Systems

## Réservation & Renseignements

**Téléphone :** +212 522 247 210
**Email :** contact@innov-systems.com
**Web :** https//www.innov-systems.com


Innov Systems — Tous droits réservés