

## ASM Fouille de données



Niveau d'étude  
BAC +4



Composante  
École Nationale  
Supérieure des  
Ingénieurs en  
Arts Chimiques



Volume horaire  
22h

### En bref

- **Code:** LP1A525V
- **Ouvert aux étudiants en échange:** Oui

## Présentation

### Objectifs

À l'issue du cours, vous serez capable de :

#### Cerner les enjeux de l'analyse de données

- Identifier les défis spécifiques aux données industrielles (qualité, volume, complexité).
- Comprendre le rôle de l'analyse dans la prise de décision (optimisation, diagnostic, prédition).

#### Structurer et préparer les données

- Nettoyer, transformer et organiser des jeux de données multidimensionnelles.
- Gérer les données manquantes et les erreurs avec une approche méthodique.

#### Sélectionner et appliquer des méthodes adaptées

- Choisir entre techniques d'exploration (ACP), de segmentation (K-means) ou d'apprentissage supervisé (régression) selon le problème.
- Justifier le choix des outils statistiques (R/Python, visualisation) pour répondre à une question métier.

#### Conduire une analyse professionnelle

- Mettre en œuvre des analyses sur des cas réels (ex : clustering de défauts, régression pour l'optimisation des coûts).
- Valider les résultats à l'aide d'indicateurs pertinents (silhouette score, R<sup>2</sup>, p-valeurs).

#### Interpréter et communiquer les résultats

- Tirer des conclusions actionnables pour l'industrie (ex : seuils d'alerte, paramètres critiques).
- Présenter les résultats sous forme visuelle (graphiques, dashboards) adaptée aux parties prenantes.

#### **Évaluer les limites et alternatives**

- Critiquer les biais ou incertitudes des modèles utilisés.
- Proposer des améliorations (ex : collecte ciblée, méthodes hybrides) pour des analyses futures.

#### **Approche pédagogique :**

- Études de cas issues du génie industriel (maintenance, logistique, qualité).
- Travaux pratiques sous **R/python** avec jeux de données réels.
- Focus sur la **r rigueur scientifique et l'applicabilité terrain.**

---

## Description

Ce cours pratique initie les futurs ingénieurs industriels aux **méthodes fondamentales d'analyse de données** appliquées à la prise de décision. À travers des études de cas concrets, les étudiants découvrent :

- **L'acquisition et préparation des données** (capteurs, traçabilité, gestion des données manquantes)
- **Les techniques d'exploration** (ACP, analyse de variance) et de **segmentation** (K-means, dendrogrammes)
- **L'apprentissage supervisé** (régression, classification) avec validation rigoureuse des modèles

---

## Pré-requis obligatoires

- Mathématiques
- Statistique
- Manipulation de la donnée
- Connaissance basique des langages R/python