



# Penthao

Mise à jour : Fév. 2024

**Durée** : 3 jours - 21 heures

## OBJECTIFS PÉDAGOGIQUES

- Implémenter une solution ETL avec PDI.
- Découvrir le traitement et l'alimentation des données, la sécurisation et l'optimisation des flux.

## PRÉREQUIS

- Connaissance du SQL.

## PARTICIPANTS

- Administrateurs systèmes/BDD, Chefs de projets, Développeurs

## MOYENS PÉDAGOGIQUES

- Réflexion de groupe et apports théoriques du formateur
- Travail d'échange avec les participants sous forme de réunion-discussion
- Utilisation de cas concrets issus de l'expérience professionnelle
- Validation des acquis par des questionnaires, des tests d'évaluation, des mises en situation et des jeux pédagogiques
- Remise d'un support de cours

## MODALITÉS D'ÉVALUATION

- Feuille de présence signée en demi-journée
- Evaluation des acquis tout au long de la formation
- Questionnaire de satisfaction
- Attestation de stage à chaque apprenant
- Positionnement préalable oral ou écrit
- Evaluation formative tout au long de la formation
- Evaluation sommative faite par le formateur ou à l'aide des certifications disponibles

## MOYENS TECHNIQUES EN PRÉSENTIEL

- Accueil des stagiaires dans une salle dédiée à la formation, équipée d'ordinateurs, d'un vidéo projecteur d'un tableau blanc et de paperboard

## MOYENS TECHNIQUES DES CLASSES À DISTANCE

- A l'aide d'un logiciel comme Teams, Zoom etc... un micro et éventuellement une caméra pour l'apprenant,
- suivez une formation en temps réel et entièrement à distance. Lors de la classe en ligne, les apprenants interagissent et communiquent entre eux et avec le formateur
- Les formations en distanciel sont organisées en Inter-Entreprises comme en Intra-Entreprise
- L'accès à l'environnement d'apprentissage (support de cours, labs) ainsi qu'aux preuves de suivi et d'assiduité (émargement, évaluation) est assuré
- Les participants reçoivent une invitation avec un lien de connexion
- Pour toute question avant et pendant le parcours, une assistance technique et pédagogique est à disposition auprès de notre équipe par téléphone au 03 25 80 08 64 ou par mail à [secretariat@feep-entreprises.fr](mailto:secretariat@feep-entreprises.fr)

## ORGANISATION

- Les cours ont lieu de 9h à 12h30 et de 13h30 à 17h

## PROFIL FORMATEUR

- Nos formateurs sont des experts dans leurs domaines d'intervention
- Leur expérience de terrain et leurs qualités pédagogiques constituent un gage de qualité

## ACCESSIBILITÉ

- Les personnes atteintes de handicap souhaitant suivre cette formation sont invitées à nous contacter directement, afin d'étudier ensemble les possibilités de suivre la formation. Notre organisme peut vous offrir des possibilités d'adaptation et/ou de compensations spécifiques si elles sont nécessaires à l'amélioration de vos apprentissages

# Programme de formation

## Introduction et découverte de l'environnement

- Pourquoi Pentaho Data Integration (PDI) ?
- Historique et présentation
- Architecture, vue d'ensemble des composants.
- Installation et Configuration

## Découverte des flux

- Les transformations
- Les jobs
- Text file Output et CSV Input
- Exécuter un job.
- Exécuter une transformation.

## Traitement des données

- Qualifier les données à traiter : Sort, Unique et Filter Rows...
- Gérer les chaînes de caractères : String cut
- Effectuer des calculs : Calculator
- Joindre des données : Join Rows, Merge Join et Rows...).

## Les métadonnées

- Métadonnées : vue d'ensemble et création.

## Travailler avec les bases de données

- Bases de données supportées.
- Récupérer une table d'une métadonnée.
- Interaction avec la base de données.

## Développement de transformations avancées

- Propriétés d'un job et transformation.
- Utiliser des variables : Set / Get Variables.
- Copy / Get Rows to Result.
- Alimenter des données via des services Web
- Monitoring (" Write to log " et mails).

## Gestion des logs et des erreurs

- Les types d'erreurs.
- Gérer les erreurs (dans un job, une transformation).
- Paramètres de gestion d'erreurs.
- Le composant " Data Validation ".
- Les niveaux de logs (via la console, sauvegardés en base...).

## Fonctionnalités avancées, automatisation et documentation

- Débogage (prévisualisation et ajout de points d'arrêt).
- Parallélisation des traitements.
- Export des travaux.
- Kitchen : l'exécuteur de jobs (fichiers, repository, tâches planifiées).
- Documenter ses jobs.