Introduction

Cette formation de trois jours offre une introduction à l'environnement d'**Exploration de Données** KNIME®. Aucune expérience préalable en programmation ou connaissance de KNIME® n'est requise. Les thèmes abordés au travers du cours sont la prise en main de l'outil KNIME®, l'exploration et la visualisation de données, ainsi que la programmation par diagrammes de nœuds et flux de données.

A qui s'adresse cette formation ?

Cette formation s'adresse aux scientifiques qui souhaitent acquérir une connaissance solide de l'environnement KNIME® et s'initier aux techniques de l'Exploration de Données.

Vous allez apprendre à

Maîtriser les bases de KNIME®, l'outil par excellence pour l'exploration de données dans de multiples domaines. Lire, fusionner, manipuler, convertir, mettre en valeur, visualiser et sauvegarder vos données scientifiques avec un outil ouvert, puissant et polyvalent.

Quels sont les prérequis ?

•Avoir une expérience préalable dans l'utilisation de logiciels scientifiques.

•Connaitre les bases du logiciel Excel® ou équivalent.

•Comprendre l'anglais informatique élémentaire (outil en langue anglaise).

Méthodes, modalités, délais d'accès et adaptation aux personnes en situation d'handicap

Cette formation est entièrement basée sur le logiciel pédagogique KNIME avec la mise en place de workflows comme méthode d'apprentissage. Une évaluation continue par QCM est effectuée sur les exercices proposés lors de la formation et des questionnaires de satisfaction à chaud et à froid sont proposés aux apprenants. La formation est proposée uniquement en présentiel en intra-entreprise. Adaptation du temps et mode de formation aux personnes en situation d'handicap en lien avec l'entreprise. Délai d'accès maximum de deux mois entre la commande et le début de la formation.

Programme complet

Premier jour

Introduction

- Comment installer l'environnement d'exploration de données KNIME®.
- L'exploration et l'analyse de données par diagrammes de nœuds dans KNIME®.
- Familiarisation avec la zone de travail de KNIME® et ses composants.

Mon premier diagramme de données

- Création, sauvegarde et suppression d'un diagramme.
- Manipulation d'un nœud : création, configuration, exécution et appellation d'un nœud.
- Description d'un nœud et visualisation des données associées.
- Structure de données dans KNIME®: organisation des données dans chaque nœud sous forme de tableur avec lignes et colonnes.
- Comment filtrer des colonnes dans le tableur d'un nœud.
- Comment filtrer des lignes dans le tableur d'un nœud.
- **Mise en application.**

Ma première exploration de données

- Introduction à l'exploration de données.
- Remplacer des valeurs dans une colonne.
- Scinder une colonne avec des données de type texte.
- Types de données dans KNIME® et conversion de types.
- Visualisation de données : surbrillance, nuage de points, matrice de points, tableur interactif, graphe linéaire, graphe par intervalles, histogramme, coordonnées parallèles.
- **Mise en application.**

Deuxième jour

Mise en pratique sur des données expérimentales des connaissances KNIME acquises lors du premier jour de formation

- **Lecture** et écriture de données hétérogènes expérimentales à partir de plusieurs fichiers au format CSV et/ou Excel et à l'aide des nœuds -File Reader-, -Excel Reader-, -CSV Writer- et -Excel Writer-.
- **Normalisation** des noms de colonnes des données expérimentales à l'aide des nœuds -Column Rename- et –Column Rename (regex)-.
- **Filtrage** des données expérimentales à l'aide des nœuds –Row Splitter-, -Column Splitter-, -Rule Engine- et –Rule Base Splitter-.
- **Transformation** des données expérimentales à l'aide des nœuds -Cell Splitter by Position-, -Cell Splitter by [Délimiter]-, -String Manipulation- et -String to Date & Time-.
- **Visualisation et mise en valeur** des données transformées à l'aide des nœuds -Color Manager-, -Scatter Plot-, -Scatter Matrix-, -Histogram-, -Pie Chart-, -Radar Plot-, -Parallel Coordinates- et -Interactive Table-.

Suite de l'exploration de données et découverte de nouvelles fonctionnalités

- Transposition de tableaux.
- Groupement et dégroupement de données.
- Pivotement et dépivotement de données.
- Concaténation et Jonction de tableaux.
- **Mise en application.**

Troisième jour

Mise en pratique sur des données expérimentales des connaissances KNIME acquises lors du deuxième jour de formation

- **Groupement** de données transformées afin de calculer des résultats statistiques sur la nature des données.
- **Transposition**, **pivotement** et **dépivotement** de données dans le but de transformer l'organisation des données pour une meilleure compréhension et lecture des résultats statistiques.
- **Jonction** des résultats statistiques sur différentes données pour analyse et comparaison.

Une première analyse de données

- Introduction à l'analyse de données.
- Séparation et mélange de données : séparation par ligne, partitionnement, brassage et concaténation.
- Transformation de colonnes : Gérer les valeurs manquantes, normalisation de données.
- Mon premier modèle de données : apprentissage et prédiction des données par la mise en pratique d'un modèle de classification.
- Sauvegarde et lecture d'un modèle de données à partir d'un fichier.
- **Mise en application.**

Mise en pratique sur des données expérimentales des connaissances KNIME acquises lors du troisième jour de formation

- Mise en place d'un premier modèle de classification sur les données transformées préalablement avec évaluation par des estimateurs statistiques et visualisation des résultats.