

Mis à jour le 03/05/2024

S'inscrire

# Formation Polars dataframe

3 jours (21 heures)

## Présentation

Notre formation Polars dataframe vous permettra de tirer le maximum de Polars, la librairie Open source pour le traitement des données. Si les pipelines de données rencontrent souvent des problèmes lors du traitement de [Big Data](#) avec Python, Polars permet d'ouvrir la voie au traitement de grands ensembles de données grâce à ses structures très performantes.

Notre programme vous permettra d'installer Polars sur votre système et à le configurer selon vos besoins précis. Vous apprendrez à travailler avec des fichiers CSV et Parquet ainsi que les techniques de manipulation de données dans Polars (filtrage, sélection, tri, etc.).

Notre formation couvre également l'utilisation du Lazy mode ainsi que l'interopérabilité avec la conversion vers et depuis Pandas & Numpy. Vous apprendrez également la lecture de fichiers CVS et JSON.

Comme toutes nos formations, elle se déroulera sur la dernière version de l'outil : [Polars 0.20.23](#)

## Objectifs

- Manipuler de grands ensembles de données avec Polars
- Analyser ses ensembles de données avec les expressions
- Utiliser le Lazy mode

## Public visé

- **Data scientists**
- Data analysts

## Pré-requis

- Maitrise de Python

# PROGRAMME DE NOTRE FORMATION POLARS DATAFRAME

## INTRODUCTION À POLARS

- Qu'est-ce que Polars ?
- Différences avec Pandas
- Utiliser Polars pour l'analyse de données
- Aperçu de la performance et de la syntaxe expressive de Polars
- Structure des DataFrames et des Series dans Polars
- Présentation du mode Lazy Evaluation

## INSTALLATION ET CONFIGURATION DE L'ENVIRONNEMENT

- Installer Polars sur différents systèmes d'exploitation
- Configuration de l'environnement de développement
- Vérification de l'installation et résolution des problèmes courants
- Préparation de l'espace de travail pour les exercices pratiques

## STRUCTURES DE DONNÉES DE BASE DANS POLARS

- Explication détaillée des Series
- Comprendre les DataFrames et leurs avantages
- Création et manipulation basique des Series et DataFrames
- Conversion entre Polars et Pandas/Numpy pour une interopérabilité

## MANIPULATION DE DONNÉES AVEC POLARS

- Importation et exportation de données avec des fichiers CSV et Parquet
- Techniques de sélection, de filtrage et de tri des données
- Ajout et transformation de colonnes dans un DataFrame
- Gestion des valeurs manquantes et des types de données

## LAZY MODE

- Introduction au Lazy mode et ses avantages
- Construction de requêtes
- Optimisation des performances avec l'évaluation paresseuse
- Exemples pratiques d'utilisation

## STATISTIQUES ET AGRÉGATIONS

- Utilisation des fonctions statistiques de base
- Techniques de comptage et de résumé des valeurs
- Group By et agrégations, y compris avec des expressions complexes
- Agrégation en mode Lazy et optimisations associées

## OPÉRATIONS SUR LES GROUPES ET JOINTURES

- Pivot, melt et autres transformations de groupes
- Concaténation et jointures (inner, left, fast-track) de DataFrames
- Filtrage avancé en utilisant les conditions et jointures

## TRAVAILLER AVEC LES DONNÉES TEMPORELLES

- Introduction aux types de données temporelles et gestion des fuseaux horaires
- Manipulation et analyse des données temporelles
- Group by dynamique pour l'analyse de séries temporelles

## LECTURE ET ÉCRITURE DE FICHIERS DE DONNÉES

- Lecture et écriture de divers formats de fichiers (CSV, Excel, JSON, Parquet)
- Techniques de traitement des jeux de données plus grands que la mémoire
- Lecture de données depuis des bases de données

## CONCLUSION ET MEILLEURES PRATIQUES

- Récapitulation des concepts clés de la formation
- Discussion sur les raisons de choisir Polars par rapport à d'autres outils
- Échange sur les cas d'usage avancés et l'intégration avec d'autres bibliothèques
- Conseils pour continuer à apprendre et à s'améliorer après la formation

## 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.