PROGRAMME DE FORMATION

Analyse de données avec Python : pandas, NumPy et visualisation

Développement & Programmation

RÉFÉRENCE	VERSION	ÉDITION	DOMAINE
WF-0440	V2026-004	2 novembre 2026	Développement & Programmation

DURÉE	MODALITÉ	PROGRAMME	SANCTION
35 h	Distanciel asynchrone (e-learning)	9 modules 37 leçons	Attestation et certificat de réalisation

Objectifs pédagogiques

- Organiser un carnet Jupyter reproductible et calculer de façon vectorisée avec NumPy.
- Importer des données CSV, Excel et JSON dans pandas avec les bons types, puis sélectionner et filtrer.
- Contrôler la qualité d'un jeu de données et le nettoyer en documentant chaque choix (types, valeurs manquantes, doublons, texte).
- Traiter des dates et des séries temporelles, agréger, remodeler et combiner des tables sans perte ni doublon.
- Produire des statistiques descriptives et des graphiques lisibles et accessibles avec Matplotlib et pandas.
- Livrer une analyse reproductible et performante sous forme de rapport exporté, avec ses constats et ses limites.

L'essentiel de la formation

Public visé

Développeurs, analystes, contrôleurs de gestion, chargés d'études et techniciens qui savent écrire un petit programme Python et veulent importer, nettoyer, analyser et représenter des données tabulaires avec pandas, NumPy et Matplotlib, au-delà du tableur.

Prérequis

Écrire de petits programmes Python (variables, listes, dictionnaires, boucles, fonctions, lecture d'un fichier CSV) ; à défaut, les leçons de Python (WF-0110) sont réemployées en rappel.

Connaître les notions de moyenne, de médiane et de pourcentage.

Disposer d'un ordinateur avec une connexion internet et un navigateur récent ; aucune installation n'est nécessaire, les carnets tournent dans l'environnement d'exercice.

Modalités et déroulement

Formation à distance, en ligne, à rythme libre, organisée autour d'organisations fictives (Kestrel Mobilité, vélos électriques en libre-service avec application et API ; Imprimerie Brasselin, PME de 45 salariés sur deux sites ; Collectif Ombelle, studio de design et de vidéo ; Médiathèques de Valtorne, réseau intercommunal avec portail en ligne ; Les Vergers de Lirac, coopérative fruitière — organisations fictives, données fictives). Un test de positionnement ouvre le parcours. Chaque leçon associe une vidéo de démonstration, une fiche pratique à conserver et deux mises en pratique corrigées (code écrit dans des carnets de l'environnement d'exercice et vérifié par des tests automatiques, fichiers à produire, analyses d'erreurs, décisions de traitement et écrits de restitution) ; chaque module se clôt par un atelier de synthèse corrigé, une vidéo récapitulative et un quiz. Les manipulations se font dans un environnement d'exercice en ligne fourni par l'organisme (terminal, éditeur de code et services simulés, signalés comme tels) pour Python, pandas, NumPy et Matplotlib dans des carnets Jupyter : aucune installation, aucune licence ni aucun compte payant n'est nécessaire ; l'apprenant peut ensuite refaire les exercices sur son propre poste. Temps de formation réparti à environ 60 % de pratique, 30 % de démonstrations et d'apports, 10 % d'évaluation. Accompagnement par un formateur par messagerie (réponse sous 48 h ouvrées), correction individuelle des livrables sur demande.

Méthodes pédagogiques

- Positionnement en début de parcours
- Apports en vidéo : explications et démonstrations
- Exercices d'application et mises en pratique corrigées
- Production de livrables professionnels
- Atelier de synthèse en fin de module
- Quiz et QCM d'auto-évaluation
- Accompagnement et retours du formateur

Moyens pédagogiques et techniques

- Plateforme de formation en ligne, accès individuel à distance
- Vidéos pédagogiques
- Fiches pratiques à conserver
- Quiz et exercices interactifs en ligne
- Correction des productions sur grille de critères
- Messagerie avec le formateur (réponse sous 48 h ouvrées)

Modalités d'évaluation

Positionnement initial (non éliminatoire) ; correction automatique ou sur grille critériée de chaque exercice ; atelier de synthèse corrigé et quiz à la fin de chaque module ; évaluation finale interne : analyse complète d'un jeu de données fictif non vu en cours (prêts des Médiathèques de Valtorne : import de deux sources, contrôle qualité et nettoyage documentés, dates, agrégats et remodelage, jointure contrôlée, quatre graphiques lisibles, rapport exporté avec constats et limites), corrigée sur grille, et questionnaire de 20 questions. Seuil de réussite interne : 70 % des critères. Une attestation de fin de formation précise les compétences évaluées ; aucune certification, ni d'éditeur ni d'État, n'est délivrée ni préparée.

Sanction

Une attestation de fin de formation et un certificat de réalisation sont remis au participant sous réserve de la réalisation du parcours.

Cette formation n'est pas présentée comme préparant à une certification enregistrée : elle ne comprend ni inscription ni présentation à des épreuves de certification, et n'ouvre droit à aucune délivrance de certification ni au financement par le CPF.

Modalités et délais d'accès

La demande d'inscription peut être effectuée par téléphone, par courrier électronique ou au moyen du formulaire en ligne.

Un entretien ou un questionnaire de positionnement permet de vérifier l'adéquation de la formation avec les besoins, le niveau initial et les prérequis du candidat.

L'inscription devient définitive après validation du dossier administratif, acceptation des documents contractuels et, le cas échéant, confirmation du financement.

Les accès à la plateforme sont généralement activés sous 48 à 72 heures ouvrées après validation complète du dossier.

Les formations à distance sont proposées en entrées et sorties permanentes, sous réserve des délais ou dates imposés par le financeur ou le certificateur.

Accessibilité aux personnes en situation de handicap

Vidéos sous-titrées et transcrites, fiches pratiques, énoncés et corrigés en version texte lisible par les lecteurs d'écran, code et commandes fournis en texte copiable (jamais en image), contrastes et tailles de police réglables dans l'environnement d'exercice, navigation au clavier, rythme libre et durée allongeable pour les exercices chronométrés. Référent handicap joignable à handicap@will-formation.com pour étudier toute adaptation (durée, supports, outils, modalités d'évaluation).

Tarif de la formation

Prix : **3 060 €**

Financement : prise en charge possible par les organismes financeurs, selon les critères d'éligibilité applicables.

Qualiopi processus certifié

La certification qualité a été délivrée au titre de la catégorie d'action suivante : ACTIONS DE FORMATION. Certificat QUA25070009. WILL FORMATION — Société par actions simplifiée à associé unique au capital de 30 000 € — NDA 11757059675 (cet enregistrement ne vaut pas agrément de l'État) — SIRET 893 713 180 00014 — 01 85 09 25 99 — info@will-formation.com — www.will-formation.com

Programme détaillé

9 modules, 37 leçons — durée totale 35 h.

01 Carnets Jupyter et rappels Python

1. Positionnement — Python et lecture de données
2. Rappels Python pour l'analyse de données
3. Organiser un carnet Jupyter reproductible
4. Structurer un projet d'analyse
5. Atelier de synthèse — Carnets Jupyter et rappels Python

02 Calcul numérique avec NumPy

1. Tableaux NumPy : création, types et forme
2. Vectorisation, masques et diffusion
3. Agrégats, tri et aléatoire reproductible
4. Atelier de synthèse — Calcul numérique avec NumPy

03 Importer et sélectionner avec pandas

1. Series et DataFrame : index, colonnes et types
2. Importer CSV, Excel et JSON
3. Sélectionner et filtrer : loc, iloc et query
4. Atelier de synthèse — Importer et sélectionner avec pandas

04 Contrôler et nettoyer les données

1. Types et conversions
2. Valeurs manquantes et doublons
3. Nettoyer le texte et les catégories
4. Atelier de synthèse — Contrôler et nettoyer les données

05 Dates et séries temporelles

1. Convertir et décomposer des dates
2. Rééchantillonner et regrouper dans le temps
3. Fenêtres glissantes et comparaisons de périodes
4. Atelier de synthèse — Dates et séries temporelles

06 Agréger, remodeler et combiner

1. Regrouper et agréger avec groupby
2. Remodeler : pivot_table, melt et crosstab
3. Combiner des tables : merge et concat
4. Atelier de synthèse — Agréger, remodeler et combiner

07 Statistiques descriptives et visualisation

1. Statistiques descriptives et corrélation
2. Graphiques avec Matplotlib et pandas
3. Graphiques lisibles et accessibles
4. Atelier de synthèse — Statistiques descriptives et visualisation

08 Performances et chaînes de traitement

1. Mesurer et accélérer un traitement
2. Traiter de gros fichiers
3. Écrire une chaîne de traitement testée
4. Atelier de synthèse — Performances et chaînes de traitement

09 Restituer une analyse reproductible

1. Cadrer une analyse à partir d'une question
2. Produire un rapport exporté et rejouable
3. Relire et critiquer une analyse
4. Atelier de synthèse — Restituer une analyse reproductible