

PROGRAMME DE FORMATION · VERSION 1 · 13 SEPTEMBRE 2026

Data Fullstack : de Python et SQL au produit data

Un parcours pratique de 140 heures pour préparer les données, analyser, évaluer un modèle et concevoir un produit data reproductible.

20 jours équivalents · 140 heures · Hybride · Sur devis

Public et prérequis

Analystes, professionnels en évolution vers la data et développeurs souhaitant relier analyse et livraison.

Aisance avec les fichiers et tableurs, proportions et logique. Bases Python/SQL à consolider au positionnement.

Poste adapté requis pour les ateliers Docker, API et BI.

Objectifs pédagogiques

- Définir des indicateurs et contrôler la qualité des données.
- Écrire des traitements Python et des requêtes SQL vérifiables.
- Explorer une distribution et expliquer l'incertitude.
- Évaluer un modèle sans fuite de données et le comparer à une référence.
- Concevoir un pipeline relançable, un contrat API et un tableau BI.
- Soutenir un dossier reproductible et organiser son transfert.

Votre projet fil rouge

Solari Services, entreprise fictive de maintenance : 120 interventions synthétiques pour définir les indicateurs, contrôler les données et soutenir un produit data.

Organisation

20 journées équivalentes de 7 heures, à répartir selon la cohorte. Durées incluant lecture, pratique, retours et évaluation ; rythme à confirmer au positionnement.

Positionnement et vérification du matériel avant inscription. Lire le programme, préparer son objectif et vérifier les modalités avec le centre.

PROGRESSION · MODULES 1 À 2

1. Cadrer une question et définir un indicateur

Jour 1 · 420 minutes, pratique et retours compris

- Besoin métier
- Dictionnaire de données
- Unité d'analyse

Atelier

Rédigez une fiche de décision pour répartir une équipe de maintenance entre trois sites fictifs. Définissez trois indicateurs, leur grain et leurs limites. Construisez un dictionnaire de dix champs et expliquez comment traiter les interventions encore ouvertes.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

2. Contrôler et préparer un tableau source

Jour 2 · 420 minutes, pratique et retours compris

- Types et formats
- Doublons
- Contrôles de réconciliation

Atelier

Créez un petit fichier contenant un doublon, une date ambiguë, un code avec zéro initial et une valeur manquante. Préparez une feuille de contrôles séparée. Expliquez chaque transformation et vérifiez que le total retenu se réconcilie avec les lignes rejetées.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

PROGRESSION · MODULES 3 À 4

3. Utiliser les bases de Python

Jour 3 · 420 minutes, pratique et retours compris

- Variables et types
- Conditions et boucles
- Collections

Atelier

Dans le notebook fourni, calculez le nombre de dossiers clos, la durée moyenne disponible et la somme des montants en centimes. Traitez séparément une liste vide, une durée manquante et une chaîne non convertible.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

4. Structurer du code réutilisable

Jour 4 · 420 minutes, pratique et retours compris

- Fonctions et contrats
- Objets simples
- Tests de comportement

Atelier

Implémentez ce contrat dans une fonction sans accès au réseau ni au fichier. Ajoutez des tests pour les limites et un appel utilisant le résultat. Proposez ensuite une petite classe de configuration seulement si elle simplifie réellement les paramètres.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

PROGRESSION · MODULES 5 À 6

5. Nettoyer et joindre avec pandas

Jour 5 · 420 minutes, pratique et retours compris

- Lecture typée
- Valeurs manquantes
- Cardinalité des jointures

Atelier

Utilisez les fichiers synthétiques du notebook pour profiler les données, séparer les anomalies et joindre un référentiel de sites. Introduisez volontairement une clé dupliquée dans le référentiel et démontrez que votre contrôle bloque le calcul.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

6. Explorer sans tromper avec un graphique

Jour 6 · 420 minutes, pratique et retours compris

- Distribution et dispersion
- Comparaisons
- Communication des limites

Atelier

Créez trois figures sur le jeu Solari : distribution des durées, taux de respect par site avec effectifs et volume par semaine. Ajoutez sous chaque figure une observation, une limite et une question à vérifier.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

PROGRESSION · MODULES 7 À 8

7. Raisonner avec un échantillon et un test A/B

Jour 7 · 420 minutes, pratique et retours compris

- Échantillonnage
- Incertitude
- Plan d'expérience

Atelier

Rédigez un protocole pour tester un rappel de maintenance. Précisez unité de tirage, critère principal, durée, exclusions et risque de contamination. Calculez l'écart observé puis expliquez ce que l'échantillon permet et ne permet pas d'affirmer.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

8. Interroger et agréger avec SQL

Jour 8 · 420 minutes, pratique et retours compris

- SELECT et filtres
- GROUP BY
- JOIN et NULL

Atelier

Dans SQLite via le notebook, créez une table d'interventions, calculez les volumes par site et les moyennes sur dossiers clos. Ajoutez un site sans intervention, puis comparez jointures gauche et interne et expliquez les écarts.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

PROGRESSION · MODULES 9 À 10

9. Analyser une chronologie avec les fenêtres SQL

Jour 9 · 420 minutes, pratique et retours compris

- PARTITION BY
- Ordre déterministe
- Cumuls et décalages

Atelier

Écrivez une requête calculant rang, cumul et variation par site. Testez deux dates identiques et un jour absent. Décrivez la différence entre une fenêtre de trois lignes et une fenêtre de trois jours calendaires.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

10. Concevoir un modèle décisionnel

Jour 10 · 420 minutes, pratique et retours compris

- Faits et dimensions
- Schéma en étoile
- Historisation

Atelier

Dessinez une étoile avec une table de faits Intervention et trois dimensions. Donnez les clés, le grain et trois mesures. Décrivez le traitement d'un site inconnu et d'un changement de région en cours d'année.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

PROGRESSION · MODULES 11 À 12

11. Choisir un stockage documentaire et organiser un lac

Jour 11 · 420 minutes, pratique et retours compris

- Documents JSON
- Contrats de schéma
- Traçabilité des fichiers

Atelier

Définissez deux documents JSON valides et un invalide, puis un contrat de contrôle. Proposez l'arborescence d'un lac pédagogique et un manifeste contenant source fictive, date, nombre de lignes et empreinte des fichiers.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

12. Collecter une API de manière robuste

Jour 12 · 420 minutes, pratique et retours compris

- HTTP et statuts
- Pagination
- Timeouts et reprise

Atelier

Avec une API simulée dans le notebook, traitez pagination, page répétée et erreur temporaire. Définissez les erreurs à réessayer et celles à signaler immédiatement. Rédigez un journal qui permet le diagnostic sans exposer un secret.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

PROGRESSION · MODULES 13 À 14

13. Construire un pipeline relançable

Jour 13 · 420 minutes, pratique et retours compris

- Extraction et transformation
- Idempotence
- Qualité et transactions

Atelier

Implémentez dans SQLite un import de lot relançable. Exécutez-le deux fois, injectez un identifiant invalide et simulez une erreur avant validation finale. Comparez le nombre de lignes, les montants et le journal de traitement.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

14. Versionner et contrôler en équipe

Jour 14 · 420 minutes, pratique et retours compris

- Git et revue
- Tests en intégration continue
- Secrets

Atelier

Dans un dépôt d'exercice local, préparez une branche, un changement de transformation et un test indépendant. Rédigez une description de revue expliquant le problème, le résultat et la validation. Concevez un workflow qui n'a besoin d'aucun secret pour les tests de données fictives.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

PROGRESSION · MODULES 15 À 16

15. Reproduire un environnement avec des conteneurs

Jour 15 · 420 minutes, pratique et retours compris

- Image et conteneur
- Ports et volumes
- Configuration

Atelier

Sur un poste de travail autorisé, décrivez une image Python minimale, son fichier de dépendances et son contrôle de santé. Identifiez les fichiers à exclure du contexte de construction. Expliquez où résident les données et comment les restaurer.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

16. Évaluer un modèle supervisé sans fuite de données

Jour 16 · 420 minutes, pratique et retours compris

- Cible et référence simple
- Séparation des jeux
- Prétraitement et métriques

Atelier

Le notebook sépare une série synthétique par date et entraîne une régression linéaire simple. Comparez son erreur à la moyenne d'entraînement, puis expliquez quelles variables auraient été indisponibles au moment de prédire. Proposez une extension scikit-learn sur un poste configuré.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

PROGRESSION · MODULES 17 À 18

17. Comparer des modèles et explorer des groupes

Jour 17 · 420 minutes, pratique et retours compris

- Validation croisée
- Arbres et régularisation
- Clustering

Atelier

Concevez un tableau comparant référence, régression et arbre selon erreur, complexité et coût d'usage. Dans le notebook, affectez des sites synthétiques à des centres imposés puis examinez l'effet d'un changement d'échelle. Proposez des critères de stabilité des groupes.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

18. Exposer un résultat dans un produit data

Jour 18 · 420 minutes, pratique et retours compris

- Contrat d'API
- Validation des entrées
- Interface et exploitation

Atelier

Définissez le contrat d'un endpoint local d'estimation : entrée, bornes, sortie et erreurs. Testez la fonction pure dans le notebook. Utilisez ensuite le kit FastAPI fourni sur un poste configuré, puis rédigez une maquette d'interface indiquant limites et date du modèle.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

PROGRESSION · MODULES 19 À 20

19. Livrer un tableau de bord BI gouverné

Jour 19 · 420 minutes, pratique et retours compris

- Modèle sémantique
- Mesures et filtres
- Accès et actualisation

Atelier

Dessinez une page Solari avec trois indicateurs et deux filtres. Spécifiez les mesures en numérateurs et dénominateurs. Sur un poste Power BI configuré, importez les données synthétiques, vérifiez les totaux et préparez une matrice de tests de visibilité par rôle.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

20. Soutenir et transmettre un projet reproductible

Jour 20 · 420 minutes, pratique et retours compris

- Dossier de preuves
- Recette et transfert
- Limites et amélioration

Atelier

Livrez le dossier Solari : question métier, source synthétique, pipeline contrôlé, analyses SQL et graphiques, modèle comparé à une référence, contrat d'API, maquette BI et procédure de reprise. Préparez dix minutes de démonstration et dix minutes de discussion avec le formateur.

Livrable et vérification

Trace de calcul ou de décision, résultat argumenté et limites, à intégrer au dossier fil rouge. Expliquer la méthode, vérifier un cas limite et justifier les hypothèses devant le formateur.

Évaluation et accompagnement

Quiz formatifs originaux et dossier final avec soutenance. Évaluation humaine selon la grille communiquée ; aucune certification externe garantie.

Évaluation finale : 60 minutes indicatives, incluses dans le parcours.

- Cohérence entre le besoin et la méthode.
- Justesse des calculs et qualité des preuves.
- Traitement des limites et des erreurs.
- Clarté du dossier et capacité à défendre une décision.

Méthodes et ressources

Leçons structurées, exemples résolus, vingt ateliers, quarante questions originales et projet accompagné. Retours et soutenance avec un formateur selon le calendrier convenu.

Supports Beyond rédigés et intégrés dans l'espace pédagogique privé. Relecture et validation par le formateur avant attribution aux apprenants. Aucun contenu ni diplôme tiers n'est inclus par simple équivalence de thème. Le laboratoire navigateur couvre Python/pandas/SQL et les calculs fournis. Docker, FastAPI, Streamlit et Power BI s'utilisent sur un poste ou environnement adapté, à confirmer avec le centre.

Modalités à confirmer avec le centre

Positionnement, calendrier, tarif, accès aux supports et accompagnement convenus avant inscription. Pour un besoin d'accessibilité, indiquez le besoin pratique sans diagnostic médical. Contact : 09 54 70 23 80 · contact@beyondexpertise.eu.

Programme à valider par le formateur avant animation. Aucun titre national, certification éditeur ou financement n'est garanti par ce document.

Références officielles

- [Python — tutoriel officiel](#)
- [pandas — jointures et combinaison de données](#)
- [PostgreSQL — fonctions de fenêtre](#)
- [scikit-learn — erreurs fréquentes et fuites de données](#)
- [Microsoft — modèle en étoile Power BI](#)
- [Docker — conteneurs](#)
- [FastAPI — validation des entrées structurées](#)
- [Requests — délais et réponses HTTP](#)
- [GitHub Actions — gestion des secrets](#)