

ROODELINE GERMAIN

Data Engineer | Consultante Data | Certifiée Dataiku

Dataiku	Python	SQL	PySpark	Power BI	GenAI
---------	--------	-----	---------	----------	-------

3 ans d'expérience • Paris, France • Mobile France entière
roodeline.germain@gmail.com • 07 49 22 25 51

À PROPOS

Data Engineer spécialisée Dataiku, je transforme des besoins métiers complexes en pipelines de données robustes, industrialisés et auditables. Certifiée Dataiku Developer, Dataiku ML & MLOps Practitioner et Dataiku Generative AI Practitioner, j'interviens de bout en bout : du cadrage des besoins métiers à la mise en production des solutions. À l'interface entre les équipes métiers, data et IT, j'anime des ateliers et accompagne la montée en compétence, avec une approche orientée valeur métier et performance.

COMPÉTENCES

Data Dataiku (Recipes, scénarios, API, Deployer, bundle, monitoring) SQL (requêtes avancées, optimisation) Python (Pandas, API, automatisation) PySpark (pipelines à grande échelle) SAS IA Générative (RAG, agents autonomes, prompt engineering)	Data Viz Power BI (DAX) Tableau Qlik Sense Excel Cloud & données AWS (S3, Athena) Hadoop / Hive, PostgreSQL, Snowflake, Databricks, BigQuery
Data Quality Contrôles qualité, règles, documentation, fiabilisation Outils & méthodes Git, JIRA, SCRUM, Agile, Data Governance	MLOps Industrialisation, déploiements, CI/CD, mise en production

PARCOURS PROFESSIONNEL

Data Engineer – Dataiku – Alten

Mai 2026 – Aujourd'hui | Mission secteur bancaire

Contexte

Mission au sein d'un grand groupe bancaire, centrée sur l'industrialisation à grande échelle de pipelines de données en environnement Hadoop/Hive et sur l'orchestration multi-projets.

Missions

- Industrialisation et mise en production de pipelines de données à grande échelle (PySpark, Dataiku DSS).
- Conception et déploiement de scénarios de lancement inter-projets et gestion des dépendances entre chaînes de traitement.
- Gestion des mises en production : bundles Dataiku, Deployer, coordination avec les équipes infrastructure.
- Diagnostic d'incidents de production et analyse de cause racine sur des chaînes de traitement complexes (pipelines de scoring, croisements de référentiels).
- Conception d'un dispositif de monitoring et d'alerting : détection des échecs et non-démarrages de scénarios, contrôles de volumétrie et de qualité des données.
- Contrôle de cohérence des données entre systèmes et réconciliation de volumétrie multi-sources.
- Développement de ciblages marketing data-driven à partir de données comportementales digitales.

Environnement technique

Dataiku DSS, PySpark, Hadoop/Hive, Deployer & bundles, SQL

Data Engineer – Dataiku – Syner-IA

Juin 2025 – Février 2026 | Paris • 9 mois

Contexte

Mission Data Engineer experte Dataiku, menée en parallèle sur un projet client (Supply Chain) et un projet interne (Challenge Agentic AI). Rôle transversal entre métiers, équipes data, IT et éditeur Dataiku : industrialiser les pipelines, améliorer la performance, mettre en place une gouvernance Dataiku durable et concevoir un workflow IA agentique complet.

Industrialisation Dataiku pour la Supply Chain (projet client)

- Ingestion multi-sources (ERP, CSV, API) et construction d'un flow Dataiku complet : préparation, jointures, règles de gestion.
- Mise en place de contrôles qualité, de mécanismes de traçabilité et d'un monitoring des jobs avec alertes en cas d'échec.
- Optimisation SQL et automatisation via scénarios (rafraîchissements, alertes, logs) pour réduire les temps de traitement et garantir une vision quasi temps réel.
- Création de dashboards Dataiku pour le pilotage temps réel des équipes Supply Chain.

Optimisation, monitoring & gouvernance Dataiku

- Optimisation des recettes pour réduire les temps d'exécution et mise en place d'un monitoring complet (KPI, performance, santé des pipelines).
- Contribution à la création d'un centre de services Dataiku (gouvernance, bonnes pratiques) et documentation technique/fonctionnelle pour l'auditabilité.
- Support aux utilisateurs, animation d'ateliers et montée en compétence des équipes data & métiers.

Conception du workflow industrialisé Agentic AI (projet interne)

- Lead technique sur le Challenge Agentic AI : création d'une roadmap, répartition des tâches et construction d'agents IA autonomes pour orchestrer les traitements.
- Mise en place de la chaîne NLP (sentiment analysis, classification, détection d'irritants) et automatisation du traitement des verbatims.
- Conception du dashboard Voix Client, avec contrôles de cohérence et logs pour assurer la traçabilité - workflow industrialisation-ready dès la phase d'innovation.

Résultats

- Vision temps réel retrouvée sur la production Supply Chain et réduction significative des temps d'exécution des pipelines.
- Mise en production fiable via Deployer (bundles, versioning) et automatisation avancée via API Dataiku.
- Forte adoption de Dataiku par les équipes métiers et data ; workflow IA autonome livré dans les délais, leadership reconnu sur le projet Agentic AI.

« Un pipeline n'a de valeur que s'il est stable, traçable et exploitable au quotidien. »

Environnement technique

Dataiku DSS, Python, SQL, API Dataiku, AWS, RAG, Git, JIRA, NLP, MLOps, Agile

Data Analyst Dataiku – Specialist Risk & Assurance - MSD / Merck (laboratoire pharmaceutique)

Septembre 2023 – Avril 2025 | Puteaux • 20 mois

Contexte

Seule data analyst au sein d'une Direction Finance fortement réglementée (SOX), en charge de concevoir et industrialiser des pipelines Dataiku pour détecter les anomalies, automatiser des calculs comptables complexes et fournir des reportings financiers fiables - en lien direct avec la Direction Finance, le contrôle interne, la comptabilité et des consultants externes (EY).

Détection d'anomalies & contrôles internes (SOX)

- Conception d'un pipeline Dataiku complet (ingestion, nettoyage, règles de contrôle, scoring d'anomalies) sur les notes de frais et factures, automatisé via scénarios.
- Ajout de contrôles qualité et structuration d'une traçabilité complète (logs, historisation, documentation, vidéos de formation).
- Restitution des analyses auprès de la Direction Finance et du contrôle interne, en collaboration étroite avec le directeur financier.

Automatisation des calculs comptables complexes

- Développement des calculs comptables (dotations, reprises, indemnités, reclassements) via Python & Dataiku, avec gestion des erreurs pour garantir la stabilité du flux.
- Construction et industrialisation d'un pipeline générant des fichiers comptables prêts à intégrer, réduisant les interventions manuelles.
- Documentation complète pour audit interne et conformité EY.

Pilotage financier & reporting automatisé

- Conception d'un modèle de données structuré pour le suivi Budget vs Réalisé.
- Développement de dashboards Power BI (DAX, KPI dynamiques, alertes d'écart).

- Connexion aux flux Dataiku pour automatiser la mise à jour des indicateurs.
- Animation d'ateliers utilisateurs et accompagnement à l'adoption.

Résultats

- Réduction des délais de traitement de 90 % grâce à l'automatisation et l'industrialisation Dataiku.
- Processus comptables fiabilisés, auditables et validés par EY ; détection de risques auparavant invisibles dans les notes de frais.
- Forte adoption de Dataiku par les équipes Finance, Comptabilité et Contrôle interne.

Environnement technique

Dataiku DSS, Databricks, Python, SQL, Power BI, AWS, JIRA, ETL/ELT, Agile

Chargée d'études - ESCDA

Avril 2022 – Septembre 2022 | Boulogne-Billancourt • 6 mois

Contexte

Mission au sein d'un cabinet spécialisé en expérience client : préparer et exécuter des études CX, analyser les données issues des enquêtes et bases clients, et accompagner les équipes dans la compréhension des résultats, en autonomie sur toute la chaîne (collecte, analyse, synthèse, restitution).

Missions

- Préparation de l'étude CX : définition des critères d'évaluation, organisation des missions clients mystères et protocoles de contrôle qualité.
- Exploitation des bases de données clients : extraction, nettoyage, analyse quantitative (statistiques descriptives) et qualitative (verbatim, tendances).
- Veille concurrentielle et benchmark CX ; production des rapports d'étude et restitution auprès des équipes internes.

Résultats

- Amélioration de la compréhension des comportements clients et recommandations adoptées pour optimiser l'expérience client.
- Qualité des analyses reconnue par les managers.

Environnement technique

Excel, PowerPoint, Word, Reporting

CERTIFICATIONS

- Dataiku Generative AI Practitioner - Juillet 2025
- Dataiku MLOps Practitioner - Juillet 2025
- Dataiku ML Practitioner - Juillet 2025
- Dataiku Developer - Février 2025
- IBM - Artificial Intelligence Fundamentals - Avril 2025
- Dataiku Advanced Designer - Juillet 2024
- Dataiku Core Designer - Juin 2024

FORMATION

- Master Data Analyst – Business Analytics - IAE Paris-Est | 2022–2024. Mémoire M2 : prévision des dépenses et optimisation des budgets - approches économétriques et Machine Learning (Dataiku, Python, SQL, R).
- Licence Économie et Gestion – parcours Économie - Université Gustave Eiffel | 2019–2022.

LANGUES

Français *langue maternelle*

Anglais *avancé (C1)*