

Data Scientist junior

SOSSA DANIEL

PROFIL

Data Scientist Junior avec une expérience pratique en analyse de données, Machine Learning et ingénierie des données. Projets end-to-end incluant l'analyse de plus de 35 millions d'enregistrements et la conception de pipelines temps réel.

Passionné par la fintech, j'exploite la donnée pour analyser les comportements, automatiser les traitements et éclairer la décision.

CONTACT

Tél : +225 05 03 85 89 97

Email : daniel.sossa@epitech.eu | contact@danielsossa.dev

Adresse : Abidjan, Côte d'Ivoire

LinkedIn : daniel-sossa

GitHub : danielxdata

Portfolio : danielsossa.dev

COMPÉTENCES TECHNIQUES

Langages : Python, SQL

Data & Analytics : Pandas, NumPy, EDA, Feature Engineering, Power BI

Machine Learning : Scikit-learn, XGBoost, Régression, Classification, Ensemble Learning

Deep Learning & NLP : TensorFlow, Keras, CNN, NLTK, TF-IDF

Data Engineering : Apache Kafka, ETL, SQLAlchemy

Bases de données : PostgreSQL, MySQL, MongoDB

Outils & DevOps : Git, Docker, Linux, Azure, Jupyter

LANGUES

Français : courant (langue maternelle)

Anglais : intermédiaire B1 : lecture technique, communication professionnelle

APTITUDES PROFESSIONNELLES

- Rigueur analytique
- Esprit d'initiative et autonomie
- Curiosité technologique
- Orientation résultats
- Travail en équipe
- Adaptabilité et leadership

EXPÉRIENCE PROFESSIONNELLE

Développeur Web Full-Stack & QA

Ebenyx Technologies (Stage) • *Septembre à Novembre 2025, 3 mois* • *Le Plateau, Abidjan*

Contribution au développement d'une plateforme web de gestion des zones industrielles (SIGETI) couvrant la gestion des lots industriels, des conventions de location et le suivi des activités opérationnelles.

- Développement de fonctionnalités front-end et back-end avec React.js, Express.js et l'écosystème Node.js dans un environnement GitFlow.
- Participation à la conception de dashboards par rôles et de modules statistiques facilitant le suivi des indicateurs métier.
- Automatisation des tests fonctionnels avec Selenium et Playwright, permettant une détection plus rapide des anomalies et une meilleure stabilité des livraisons.
- Contribution aux pipelines CI/CD et à l'automatisation des processus de déploiement.

PROJETS RÉALISÉS

Real-time Sentiment Analysis Pipeline *Juin 2026*

Contexte : Architecture ML orientée production avec streaming de données en temps réel.

Objectif : Classifier automatiquement les sentiments de commentaires Reddit en flux continu.

Technologies : Python, Apache Kafka, PostgreSQL, SQLAlchemy, XGBoost, Docker, Docker Compose.

Résultat : Pipeline end-to-end fonctionnel (ingestion Kafka, classification, stockage, visualisation Power BI) ; modèle XGBoost atteignant 83,5 % d'accuracy et un F1-score de 0,84 sur un jeu de test de 10 000 commentaires.

GitHub:https://github.com/danielxdata/Real-time-Sentiment-Analysis-Pipeline

E-commerce Data Analysis : EDA & Segmentation Client *Mai 2026*

Contexte : Analyse de 5 jeux de données transactionnelles d'un e-commerce comprenant plus de 3 millions de commandes et 32 millions de produits commandés.

Objectif : Identifier des insights exploitables sur le comportement d'achat des clients et les tendances de consommation.

Technologies : Python, Pandas, NumPy, Matplotlib, Seaborn, Power BI.

Résultat : Analyse exploratoire de plus de 35 millions d'enregistrements ; segmentation comportementale des clients et identification des produits, départements et périodes les plus performants pour orienter les décisions métier.

GitHub : https://github.com/danielxdata/ecommerce-sales-analysis

FORMATIONS & CERTIFICATIONS

Bourse DataCamp × iSheero : ML Engineer Roadmap

Formation continue certifiante (30+ cours) • Mai 2026 à présent

Coding Academy : Data Science & Intelligence Artificielle

Epitech Côte d'Ivoire, Abidjan • Janvier à Juillet 2026

BTS Informatique : Développeur d'Applications

GEIGE, Abidjan • 2023 à 2025