



Aymen Khatrani

Élève-ingénieur Data | Python · SQL · ML · Pipelines · BI
Polytech Lille — ISIA 4^e année

Alternance Data (Analyst · Engineer · Scientist) — 1 an
— Septembre 2026

✉ aymen.khatrani@polytech-lille.net
☎ +33 7 68 87 05 39
📍 Lille (mobile Île-de-France)
🌐 [linkedin.com/in/aymen-khatrani](https://www.linkedin.com/in/aymen-khatrani)
🐙 github.com/aymen-khatrani

Résumé Professionnel

Élève-ingénieur en 4^e année à Polytech Lille en **Informatique, Statistique et Intelligence Artificielle**. Couvre l'ensemble de la chaîne de valeur data : **construction de pipelines** (ETL/ELT, SQL avancé, Python), **modélisation prédictive** (ML, scoring, séries temporelles) et **restitution d'insights** (BI, tableaux de bord, visualisation). Expériences en environnements bancaire, immobilier et recherche académique. Recherche une **alternance Data** d'un an à partir de **septembre 2026**.

Formation

2023 – 2027 Polytech Lille
Diplôme d'Ingénieur — Informatique, Statistique et Intelligence Artificielle

Statistiques inférentielles, ML supervisé/non supervisé, séries temporelles,

SQL/NoSQL, Big Data, BI, Data Viz, architectures distribuées, Scrum.

2021 – 2023 Lycée Jean Bart, Dunkerque
CPGE PSI — Mathématiques, Physique, Sciences de l'Ingénieur

Certifications

Google AI Essentials — Google
Claude Certified Architect — Anthropic
AWS Cloud Practitioner — En cours

Langues

Français : natif **Anglais** : C1, TOEIC 870 **Espagnol** : B1

Centres d'intérêt

Veille sur l'écosystème Data et IA (outils, recherche, architectures), Karting et Formule 1, Cinéma et documentaires

Expériences Professionnelles

Stagiaire R&D — Pipelines et modélisation de données neuroimagerie (en cours) *Mai – Août 2026 / MIMLAB, Boğaziçi University, Istanbul*

- Construction d'un **pipeline Python** de prétraitement, feature extraction et classification de signaux IRM/EEG multi-sites (3 laboratoires); benchmark de modèles ML/DL (accuracy, F1); co-rédaction d'un rapport scientifique en anglais.

Stagiaire Data Quality & Fiabilisation — CRM Immobilier *Juin 2025 / Cimmobilier*

- Réduit le taux d'erreur des données CRM de **22 % à 6 %** en quatre semaines via un pipeline Python/SQL de déduplication et standardisation sur ~1 500 fiches clients; documenté pour reprise autonome.

Stagiaire Analyse de Données *Juin – Juil. 2024 / Banque Populaire du Maroc, Agence Lille*

- Structuré et automatisé un pipeline d'agrégation (clients, transferts, crédits) réduisant le reporting mensuel d'environ **60 %**; livré un tableau de bord avec **5 KPIs actionnables** à la direction.

Projets

Scoring d'appétence crédit — Dataset Cofidis (5 344 clients) *R · scikit-learn · XGBoost · SMOTE · Logistic Reg · Random Forest · MLP*

- Cycle complet : EDA, feature engineering, gestion du déséquilibre (~1 :7) par SMOTE, comparaison de modèles (**AUC = 0,769, Gini = 0,537**); recommandation régression logistique pour la mise en production (interprétabilité réglementaire); restitution aux non-techniques.

Pipeline SQL avancé & contrôles qualité automatisés — Base e-commerce *PostgreSQL · PL/pgSQL · Git*

- CTEs récursives, window functions et **4 vues matérialisées** sur ~500k lignes de transactions (réduction reporting **60 %**); contrôles qualité automatisés via triggers PL/pgSQL (anomalies stock, doublons, dérivés CA).

Profilage statistique et classification — EA Sports FC (FIFA 2016–2024) *Python · pandas · scikit-learn · SHAP · Plotly · Jupyter*

- Profilage sur 9 saisons (~150 000 observations) : corrélations salaire/performance, détection d'outliers (z-score, IQR), classification multi-modèles et analyse **SHAP** de l'importance des variables; visualisations interactives Plotly.

Solidar Tool — Plateforme de prêt d'équipements entre particuliers *Java EE · JPA · EJB · PostgreSQL · Servlet/JSP*

- Conception d'une **architecture 3-tier Java EE** : EJB Session Beans (interfaces @Remote/@Local), persistance JPA/PostgreSQL, gestion transactionnelle EJB, filtrage géographique GPS (≤5 km) et système de crédits utilisateurs.