Hicham Alaoui

Data Scientist

+212 762 389 821 | halaoui@insea.ac.ma | linkedin.com/in/hicham-alaoui-08ba35206 | github.com/hicham-alaoui0

Summary

Data Scientist who owns the analytics value chain end-to-end: framing ambiguous problems with stakeholders, building robust data pipelines and models (ML/NLP when relevant), and shipping production-ready controls (dashboards, traceability, monitoring) in high-stakes environments. Strong at translating complex systems into simple workflows, KPIs, and clear visual narratives for adoption.

Professional Experience

Société Générale ATS — **Casablanca**

Data Science Intern -> EQD Trading Analyst (Data Science & Analytics Products) — *Mar 2025 – Present*

- Owned the **end-to-end analytics workflow** supporting **index structuring/management**: methodology research, data preparation, constituent/weight computation for rebalances, and stakeholder-ready outputs for sign-off.
- Built and maintained **Python/SQL pipelines** ingesting and reconciling **market + reference data** (incl. lifecycle events), producing reproducible datasets with **feature engineering** (returns, volatility, liquidity, factor-style signals).
- Developed **scoring/ranking frameworks** combining **constraint logic** with **predictive modeling (ML/statistical)**; ran **backtests + sensitivity analyses** to assess robustness (performance, risk, exposures, turnover) and document trade-offs.
- Produced **diagnostics and decision visuals** (run comparisons, change attribution, exceptions/mismatch analysis) and translated results into clear stakeholder narratives.
- **Key initiative:** delivered a **production-control + observability layer** around rebalancing runs (governance, KPI tracking, automated reconciliation) to reduce rework and strengthen operational reliability.
- **PFE (Final-Year Project):** real-time pre-trade limits recalibration (12M+ trades): EVT-POT model (94% precision at 1.3% false-positive rate) + MLOps, enabling **EUR 1.8M RWA** release and **80% faster processing**.

gaea21 (NGO) — **Geneva**

Data Scientist Intern (Climate & NLP) — *Jun 2024 – Sep 2024*

- Built CO_2 emissions forecasting models (*XGBoost, LSTM*) and translated results into stakeholder-ready insights and clear visualizations.
- Developed an NLP-assisted content workflow in French supporting sustainability articles: data preparation, generation pipeline, and quality checks; delivered concise slide syntheses.

Haut-Commissariat au Plan (HCP) — **Rabat**

Data Analyst Intern — *Jul 2023 – Aug 2023*

- Produced a short reporting deliverable (data preparation + Power BI) and supported indicator computation for internal analysis.

Education

INSEA – Rabat — **2022 – 2025**

State Engineer in Applied Economics, Statistics & Big Data

Selected Projects

AI Equity Selection & Portfolio Allocation System (Decision Workflow Design) — **Personal**

- Designed an AI-driven decision workflow from research -> scoring -> weighting -> execution-ready outputs, emphasizing transparency and confidence.
- Integrated **LLM-assisted research synthesis** (RAG-ready design) to summarize issuer/news context and generate **explainable rationales** aligned with signals and portfolio rules.
- Built a framework for signal aggregation (composites + confidence) and rebalancing rules; architecture designed for API integration with external data providers and execution connectivity.
- Stack: *Python, portfolio logic, LLM integration, API-driven architecture*.

Risk-aware Sports Prediction Assistant (Decision Support System) — **Personal**

- Designed a decision-support system for predictions and risk-aware recommendations with confidence framing and structured outputs.
- Added **LLM-based reasoning/templated explanations** grounded by a structured match object and **data-quality gating** to reduce hallucinations in automated workflows.
- Target stack: *FastAPI, MLflow, TimescaleDB/Postgres, Kafka, LLM agent tooling*.

Skills

Programming	Python, SQL, R	**Viz & Apps**	Streamlit, Power BI, Tableau; dashboards & monitoring
Data	pandas, NumPy; data quality & controls; feature engineering	**Delivery**	stakeholder framing, slide synthesis, documentation, user enablement
ML/NLP	scikit-learn, XGBoost; NLP pipelines (screening / text analytics)	**Tools**	Git; reproducible workflows (automation/CI exposure)
Analytics	time-series (ARIMA/ARDL); anomaly/mismatch detection		

Languages

Arabic | French (C2) | English (C1)