# JACOB MACKEY

*Data Scientist & Financial Economist*

UK-based | Open to UAE relocation | jacobmackey01@gmail.com | linkedin.com/in/jacob-mackey-27a834349 | github.com/jacobmackey01

## PROFILE

Data scientist and financial economist who uses Python, SQL and statistical modelling to turn imperfect data into defensible commercial decisions. Builds reproducible analysis, compares models with credible baselines and makes weak or null results visible before recommending action. Targeting data science, analytics, risk and quantitative research opportunities in the UAE.

## SELECTED PROJECTS

**German Wind Forecasting & Market-Signal Research** | Python, XGBoost

Evaluated an XGBoost residual calibrator over **8,615 out-of-sample hours** and 12 expanding folds. Its **0.34% pooled MAE gain** was not statistically distinguishable from zero, so I froze a stricter day-ahead design instead of claiming a deployable edge.

**Retail Product Funnel, Activation & Cohorts** | SQL, DuckDB

Modelled an ordered journey across **2,756,101 events**. Measured 2.35% view-to-basket, 30.22% basket-to-transaction and 0.71% end-to-end conversion; recommended deeper measurement before intervention rather than unsupported causal marketing claims.

**Credit-Card Fraud Detection & Alert Thresholds** | Python, classification

Built chronological model selection and validation around the real **284,807-row** public dataset. The held-out logistic result returned **82.61% precision, 76.00% recall and 0.8075 PR-AUC**; the final threshold remains an operational capacity and customer-friction decision.

**RAG for European Energy-Market Research** | Python, ChromaDB

Built checksummed document ingestion, page-aware retrieval, constrained generation, refusal behaviour and citation-source checks over public ENTSO-E and ACER reports. Positioned the tool for evidence discovery with human claim-level review, not autonomous truth.

## APPLIED WORK

**RMS for Girls Sports Centre - Customer Service Associate (part-time), 2025-present**

Self-initiated and management-approved pricing and utilisation analysis for the 2026/27 academic year. Built Python cleaning/QA and an Excel decision model for rate-card options. Recommendations were scenario-based and not yet implemented; no realised uplift is claimed.

## CORE SKILLS

**Programming & data**
Python, Pandas, NumPy, SQL, DuckDB, Excel, REST APIs

**Modelling & statistics**
XGBoost, scikit-learn, classification, time-series forecasting, walk-forward validation, IV/2SLS, Newey-West inference

**Analytics**
Funnels, activation, cohorts, data QA, scenario analysis, stakeholder reporting

**Applied AI**
OpenAI Responses API, structured outputs, embeddings, ChromaDB, evidence retention, evaluation cases

## EDUCATION

**MSc Financial Economics (Merit)**
Cardiff University
Completed December 2024

**BSc (Econ) Economics & Finance**
Cardiff University
2023

## RESEARCH

**MSc dissertation**
Analysed 5,280 firm-month observations across 44 U.S. companies (2014-2023). Diagnosed endogeneity in the S&P; 500 control and re-estimated with IV/2SLS, while documenting proxy, survivorship and model-fit limits.

## WORKING STANDARD

Start with the decision.
Audit the data.
Establish a baseline.
Validate as time unfolds.
Make uncertainty usable.