

MASOOD NAZARI

Senior Data Scientist | Clinical Data Engineer | AI & BI Specialist

📍 Southampton, UK | ☎ 07938 935499 | 📩 Michaelnazary@gmail.com

LinkedIn GitHub Portfolio

PERSONAL PROFILE

Senior Data Scientist & Analytics Engineer (MSc Distinction) with 4+ years of experience transforming complex datasets into strategic assets within Clinical Research, Healthcare, and Fintech. Currently leading high-level data engineering and automation initiatives at a major UK Clinical Trials Unit. Expert in deploying **Python/R pipelines, API-driven architectures, and Machine Learning models** to replace legacy manual processes. Proven ability to bridge the gap between clinical operations and technical data science, delivering measurable efficiency gains (**70%+ reduction** in manual workload).

TECHNICAL SKILLS

- **Languages:** Python, R, SQL (PostgreSQL/T-SQL), Bash, SAS.
- **BI & Visualisation:** Power BI, Tableau, R Shiny, Streamlit, Excel (Advanced).
- **ETL & Data Engineering:** API Integration, Data Cleaning, ETL pipelines, Google Earth Engine, Docker.
- **Machine Learning & AI:** Scikit-learn, XGBoost, Prophet, SHAP, LIME, Neural Networks (BERT, SpaCy).
- **Cloud & DevOps:** Git/GitHub, Azure, Google Cloud Platform (GCP).
- **Compliance:** GDPR, Metadata Governance, GCP Standards, FCA Regulatory Knowledge.

PROFESSIONAL EXPERIENCE

Lead Data Engineer / Clinical Data Analyst (Official title: Coordinator)

Southampton Clinical Trials Unit (SCTU) | Academic Research Centre

Sep 2024 – Present

Southampton, UK

Acting as the technical lead for data automation and reporting infrastructure across the Cancer Vaccine Launch Pad (CVLP) and oncology portfolio.

- **Automated Safety Monitoring System (API Development):** Engineered a custom Python-based API system to automate AE/SAE cross-checks. Eliminated **15 hours of weekly manual reconciliation**, reducing the data validation workload by ~40% for the trial team while ensuring 100% consistency in safety reporting.
- **Real-Time Clinical Dashboards:** Deployed interactive Streamlit dashboards to replace static Excel reporting. Cut reporting latency from weekly to real-time, saving Trial Managers an estimated **4–6 hours per week** in status tracking and reducing data query response times by **25%**.
- **Audit Trail Automation:** Developed a programmatic API solution to extract audit logs, reducing Quality Assurance (QA) retrieval time by **80%** and ensuring immediate inspection-readiness for regulatory audits.
- **System Architecture & Risk Mitigation:** Conducted technical gap analysis on the BC Platforms Referral Tool. Identified critical authentication risks and proposed architectural changes to prevent potential data access bottlenecks for 70+ external sites.
- **Core Data Management:** Led the data build lifecycle and targeted User Acceptance Testing (UAT), achieving **100% delivery** on assigned validation workstreams despite strictly limited resource allocation.

Data Scientist (Contract)

SC Solutions | Environmental Analytics & Consulting Firm

Jun 2024 – Sep 2024

Remote/Hybrid

- **Geospatial Pipeline Architecture:** Architected automated ETL pipelines using Google Earth Engine and Python APIs to ingest multi-terabyte satellite data (NASA GEDI, Sentinel-2). Reduced data preprocessing latency by **80%**, accelerating the model iteration cycle from days to hours.
- **ML Model Performance:** Developed and deployed the BioVision Analytics Hub (Streamlit/Plotly), utilising ensemble learning and automated hyperparameter tuning to reduce Root Mean Square Error (RMSE) by **15%** for real-time biomass prediction.
- **Commercial AI Implementation:** Scaled biomass prediction models across 10,000+ hectares, improving carbon stock estimation accuracy by **20%**. Integrated Explainable AI (SHAP, LIME) to provide transparent, audit-ready insights.

Senior Data Analyst & Accountant

Parsian E-Commerce Co. | Fintech & E-Commerce Platform

Oct 2019 – Sep 2023

Tehran

- **Fraud Detection System:** Engineered Machine Learning models for transaction anomaly detection (e.g., geospatial mismatches), improving fraud capture rates by **26%** and significantly reducing financial liability.
- **Financial Workflow Automation:** Designed SQL and Python-based reconciliation scripts to replace manual ledgers. Reduced manual handling time by **43%**, cutting the month-end closing cycle by **2 days**.
- **Strategic BI Dashboards:** Built executive-level KPI dashboards (Power BI/Tableau) to identify revenue bottlenecks. Increased decision-making efficiency by ~**27%** by providing management with real-time visibility into cost drivers.
- **Team Leadership:** Mentored 5 junior analysts on SQL and Python, achieving a 95% project completion rate.

Sales Accountant

Jun 2018 – Sep 2019

- **Data Integrity & SQL:** Leveraged SQL (PostgreSQL) to query large transactional datasets, identifying financial discrepancies and improving supplier payment accuracy by **23%**.
- **Reporting Automation:** Developed automated queries for tax and compliance reporting, replacing manual Excel checks and reducing reporting error rates for regulatory audits.

KEY PORTFOLIO PROJECTS

InsurePrice: Actuarial Risk & Pricing Engine | Personal Project

2025

- **Tech Stack:** Python, CatBoost, SHAP, FastAPI, Streamlit.
- **The Build:** Architected an **FCA-compliant** insurance data platform, integrating an ensemble pricing model (CatBoost/Neural Networks) with automated pipelines for real-time risk reporting.
- **The Impact:** Delivered a **0.62+ AUC** risk model optimised across 10,000+ policies. Cut operational data processing time by **94%** and identified **£2.4M** in potential fraud exposure.

Fraud Detection Dashboard API | Personal Project

2025

- **Tech Stack:** Python, FastAPI, Docker, Network Analysis, Explainable AI.
- **The Build:** Engineered a production-ready API backend processing **5M+ transactions** with **sub-second response times**, enabling real-time identification of fraud rings and money mule accounts.
- **The Impact:** Demonstrated **3,700% ROI potential** through comprehensive model monitoring, automated drift detection, and 24-hour predictive risk scoring.

Bank Marketing Analytics Platform | Personal Project

2025

- **Tech Stack:** Python, React (Frontend), Flask (Backend), Streamlit, Machine Learning.
- **The Build:** Developed a full-stack predictive analytics platform featuring 6 comparative ML algorithms and K-Means clustering to segment **41,000+ customer records**.
- **The Impact:** Achieved **90.5% model accuracy** and simulated a **15–20% conversion rate improvement** by optimising contact frequency. Designed to reduce customer churn and maximise ROI.

NIHR Research Intelligence Dashboard | University of Southampton

2025

- **Tech Stack:** Python, Streamlit, automated ETL pipelines, Plotly.
- **The Build:** Architected a comprehensive BI platform (4,500+ lines of code) to consolidate a **£171.6M research portfolio** across 11 programmes and 314 constituencies.
- **The Impact:** Delivered parliament-ready visualisations and 70% faster reporting cycles. Verified a **122:1 ROI** potential, projecting **£116K–£172K** in annual savings.

EDUCATION

MSc Data & Decision Analytics (Distinction)

Sep 2023 – Sep 2024

University of Southampton, UK

- **Honours:** Awarded the Presidential Scholarship for Academic Excellence & Alignment with UN SDGs.
- **Key Projects:** FTSE 100 Predictive Analysis (ARIMA models, **18% improved accuracy**); S&P 500 Classification (**76% accuracy**).

BSc Materials Engineering (Distinction)

Sep 2010 – Sep 2014

Semnan University

- Graduated in the **Top 5%**. Strong foundation in mathematical modelling, computational analysis, and process optimisation.