

MASOOD NAZARI

Data Scientist | Clinical Data Engineer | AI Specialist

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

PERSONAL PROFILE

Data Scientist & Analytics Engineer (MSc Distinction) with 5+ years of experience transforming complex datasets into strategic assets. Currently leading high-level data engineering initiatives at a major UK Clinical Trials Unit. Expert in deploying **Python/R pipelines**, **API-driven architectures**, and **Machine Learning models**. Bridging the gap between clinical operations and technical data science to deliver measurable efficiency gains (reducing manual workloads by **70%**+).

TECHNICAL SKILLS

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

PROFESSIONAL EXPERIENCE

Clinical Data Analyst (Data Engineering Lead)

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).

- **Automated Safety Monitoring (API):** Engineered a custom Python-based API system to automate AE/SAE cross-checks. Successfully eliminated **15 hours of weekly manual reconciliation**, reducing data validation workload by ~40% while ensuring 100% reporting consistency.
- **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** per trial in status tracking.
- **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.
- **Delivery:** Led the data build lifecycle and targeted User Acceptance Testing (UAT), achieving 100% delivery on assigned validation workstreams under strict regulatory deadlines.

Data Scientist (Contract)

SC Solutions | Environmental Analytics

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. Reduced data preprocessing latency by **80%**, accelerating model iteration from days to hours.
- **ML Model Performance:** Developed the BioVision Analytics Hub (Streamlit/Plotly), utilising ensemble learning 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) for audit-ready insights.

Senior Data Analyst & Accountant

Parsian E-Commerce Co. | Fintech

Oct 2019 – Sep 2023
Tehran

- **Fraud Detection System:** Engineered Machine Learning models for transaction anomaly detection, improving fraud capture rates by **26%** and 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:** Built executive-level KPI dashboards to identify revenue bottlenecks. Increased decision-making efficiency by ~27% by providing real-time visibility into cost drivers.

Sales Accountant

MaximTaxi | Private Transport

Jun 2018 – Sep 2019
Tehran

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

KEY PORTFOLIO PROJECTS

InsurePrice: Actuarial Risk Engine | Personal Project

2024

- **Tech Stack:** Python, CatBoost, SHAP, FastAPI, Streamlit.

- **Impact:** Architected an FCA-compliant platform with ensemble pricing models. Delivered a **0.62+ AUC** risk model optimised across 10,000+ policies, demonstrating robust separation of high-risk profiles.

Fraud Detection Dashboard API | Personal Project

2024

- **Tech Stack:** Python, FastAPI, Docker, Network Analysis.
➤ **Impact:** Engineered a production-ready API backend processing transactions with **sub-second response times**, utilising Network Analysis to identify complex fraud rings in real-time.

Bank Marketing Analytics Platform | Personal Project

2024

- **Tech Stack:** Python, React (Frontend), Flask (Backend), Streamlit, Machine Learning.
➤ **Impact:** Developed a full-stack predictive analytics platform using K-Means clustering to segment **41,000+ records**. Achieved **90.5% model accuracy** and optimised contact frequency strategies.

NIHR Research Intelligence Dashboard | University of Southampton

2024

- **Tech Stack:** Python, Streamlit, automated ETL pipelines, Plotly.
➤ **Impact:** Architected a BI platform to consolidate a **£171.6M research portfolio**. Delivered parliament-ready visualisations and 70% faster reporting cycles.

EDUCATION

MSc Data & Decision Analytics (Distinction)

Sep 2023 – Sep 2024

University of Southampton, UK

BSc Materials Engineering (Distinction)

Sep 2010 – Sep 2014

Semnan University