MOSTAFA SAAD

Data Analyst | BI, SQL, Python & Machine Learning | Power BI, DAX & Excel

Al Jizah / Cairo, Egypt | +20 101 009 8727 (Phone / WhatsApp) | mostafasaad0091@gmail.com
linkedin.com/in/mostafa-saad-171086191 | github.com/mostafasaad91 | mostafasaad.substack.com | x.com/TacTical_41
Portfolio: https://mostafasaad91.github.io/?track=data

PROFESSIONAL SUMMARY

Data Analyst & Analytics Engineer with 6+ years of experience turning complex institutional, operational, and sales datasets into live interactive BI dashboards and zero-leakage ML systems. Proven track record across SQL Server (T-SQL), SQLite Star Schema, Python (Pandas, Scikit-Learn, LightGBM, XGBoost, SHAP), Power BI (DAX, PBIP), and Excel (Power Query, Power Pivot, What-If modeling). Experienced in leading 7+ archival and research database projects, supervising data teams, and facilitating advanced Excel workshops.

FLAGSHIP PORTFOLIO PROJECTS (GITHUB)

Corporate FP&A; — Executive P&L;, Cash Flow & Margin Intelligence (Live 6-Page Power BI + SQL Server)

github.com/mostafasaad91/corporate-fpa-financial-intelligence | SQL Server Star Schema · Power BI PBIP (34 DAX · What-If) · SEC EDGAR

- Engineered a SQL Server Star Schema (CorporateFPnA, 7 pbi tables + 2 T-SQL views) on 21,451 official audited SEC EDGAR XBRL facts across 12 global giants (FY2015–FY2024, $3.01T FY24 revenue, pytest 4/4 GAAP tie-out) powering a live 6-page Power BI PBIP/TMDL suite (54 visuals, 34 DAX, P&L; Matrix, DuPont ROE, +55% AI CapEx audit & What-If Margin Simulator).

UK Retail — CLV Prediction, Survival & Market Basket Engine (Live 4-Page App)

Live App: mostafasaad91.github.io/uk-retail-clv-survival-engine | Multi-Sheet Excel · SQLite Star Schema · Stacking + SHAP · 8-Sheet Excel

- Engineered a 1,067,371-row multi-sheet Excel -> zero-fan-out SQLite Star Schema pipeline (£17.07M GMV, 36,594 invoices, 5,852 customers, pytest CI), built a strict out-of-time Two-Stage CLV engine (Stacking Ensemble 0.7979 Test AUC; Top 20% VIPs = 73.22% of 90d revenue; TreeSHAP), Kaplan-Meier survival (89.4% vs 25.2% M12), Isolation Forest anomaly audit (67.3% of canceled value in 2% accounts), 20D TruncatedSVD basket (31.6x lift), LightGBM weekly forecast (13.78% MAPE), and an 8-sheet live-formula Excel workbook.

Olist Supply Chain — SLA Risk, GMV Forecasting & Live 4-Page BI Platform (99,441 Orders)

Live App: mostafasaad91.github.io/olist-supply-chain-sla-forecasting | SQLite Star Schema · 6-Sheet Excel · LightGBM + SHAP · CI

- Built a zero-fan-out SQLite Star Schema across 8 official tables (R$15.29M GMV, pytest CI), proved 5.88x late-delivery bad-review risk (p<1e-300), trained Out-of-Time LightGBM (+45.4% CS ROI) & Ridge GMV forecast (3.50% MAPE vs 3.86% Naive), and deployed a Live 4-Page Web Dashboard + 6-sheet formula-driven Excel What-If Simulator.

AdventureWorks — Enterprise SQL, RFM & Power BI Analytics & Live 6-Page Power BI ($110M OLTP)

github.com/mostafasaad91/adventure-works-sales-analytics | SQL Server · Python pyodbc K-Means · Live Power BI (17 DAX)

- Analyzed a 72-table SQL Server 2022 OLTP database (121K sales lines, $110M revenue) via point-in-time cost views, Star Schema, Python pyodbc K-Means RFM write-back (k=4), cohort retention (~345d), market basket (lift > 90x), and live 6-page Power BI (17 DAX measures).

Telco Churn Retention ROI & FinTech Credit Risk Scorecard (2 CI-Tested ML Repos)

github.com/mostafasaad91/telco-churn-retention-roi | github.com/mostafasaad91/fintech-credit-risk-scorecard

- Telco Churn (7,043 clients): Calibrated XGBoost (0.848 Test AUC, Brier 0.135) + ROI threshold (0.37, +27.2% profit) & SHAP. Credit Risk (30,000 clients): SQL QA (399 codes) + WoE/IV + LightGBM scorecard (KS 0.429, Gini 0.564) + 300–850 scoring & NT$29.6M Basel EL.

CORE TECHNICAL SKILLS

BI & Live Apps	Power BI (PBIP, DAX, Star Schema), Excel (Power Query, Power Pivot, What-If), Interactive Web BI, Tableau
SQL & Databases	SQL Server (T-SQL, CTEs, Window Functions, Views), SQLite, Star Schema Modeling & Automated QA Tests
Python & ML	Python, Pandas, NumPy, Scikit-Learn, LightGBM, XGBoost, SHAP, Calibration, K-Means, Time-Series, pyodbc

Data Governance	Data Cleaning & Validation, Metadata Standards, Coding Systems, GitHub Actions CI, Team Lead & Workshops

PROFESSIONAL EXPERIENCE

Digital Archivist & Cultural Data Analyst / Metadata Archivist — NADIM Foundation

June 2024 – Present | Cairo, Egypt

- Manage cultural heritage datasets, metadata standardization, cross-sheet ID reconciliation, and automated QA checks across multi-source institutional databases.
- Deliver clean, structured reporting tables and analytical summaries supporting conservation and digital encyclopedia teams.

Freelance Data Analyst — Independent Consulting

October 2022 – Present | Remote

- Deliver end-to-end client analytics from raw operational data to interactive Power BI dashboards, DAX models, Power Query ETL pipelines, and recurring KPI reports.
- Use SQL and Python for automated data extraction, validation, segmentation, and statistical analysis.

Senior Data Analyst & Archive Officer (prev. Junior Analyst) — Daftar Ahwal Data Research Institute

September 2017 – September 2022 (5 yrs 1 mo) | Cairo, Egypt

- Led and managed 7+ archival and social-research dataset projects covering collection, cleaning, statistical analysis, graphing, and reporting in Excel and Power BI.
- Designed dataset methodologies and data coding systems to generate accurate indicators and forecasts, while supervising data teams and teaching advanced Excel workshops.

Content Writer — AlManassa

October 2017 – October 2018 | Cairo, Egypt

- Authored analytical articles and critical reviews, building concise executive storytelling skills in Arabic and English.

CERTIFICATIONS & EDUCATION

- **Football Analysis & Data Analytics Training:** FP Academy · PFSA Scouting · The FA · England Football (EE Playmaker)
- **Bachelor's Degree, Business Administration & Management** — Sadat University (2010 – 2014)
- **Languages:** Arabic (Native), English (Professional Analytical & Technical Writing)