MOHAMED AYMAN

Senior Data Scientist | Customer Intelligence, Machine Learning & GenAI on Microsoft Fabric

Cairo, Egypt | +20 106 006 0292 | mohammed.ayman.stat@gmail.com
linkedin.com/in/mohamed-ghareb | github.com/ayman-stat | mohamedaymanstat.vercel.app

SUMMARY

Senior data scientist with 5+ years in customer and decision intelligence for fitness, healthcare-diagnostics and telecom businesses. Built a 260-program retention engine that now generates ~SAR 20M a month at Leejam Sports, and replaced manual campaign targeting at IDH with ML segmentation that quadrupled monthly campaign impact. Proves impact with A/B tests against holdout groups, ships models as governed data products on Microsoft Fabric, and grounds LLM assistants in governed KPI definitions.

EXPERIENCE

Leejam Sports Company (fitness-centre chain, Saudi Arabia) — Remote from Egypt
Senior Data Scientist — Apr 2025 – Present

- Built a retention engine of 260 personalised-offer programs across the member lifecycle (retention before expiry, growth for active members, reactivation of members lapsed 1–11 and 12+ months, and lead-to-member conversion), now generating ~SAR 20M (~USD 5.3M) a month.
- Built Member 360, a Microsoft Fabric gold-layer table of 80+ engineered behavioural and commercial features that is the single governed input to churn scoring, RFM and clustering segmentation, next-best-offer scoring, personalised messaging and target/control assignment.
- Developed the member churn model and the target/control design that judges every retention and offer campaign on incremental lift against holdout groups (A/B), not raw conversion.
- Owned the Amazon Bedrock Knowledge Base that grounds a Claude-based analytics assistant in Microsoft Teams (schemas, KPI definitions, business rules, guardrails, numeric validation), improving the consistency of its KPI answers.
- Cut semantic-model maintenance time 40% with a galaxy-schema Fabric model unifying 5+ source systems into governed leadership KPIs.
- Applied NLP sentiment and intent classification to CX tickets and survey text to tie experience drivers to specific centres; lead delivery with executives and mentor junior analysts.

IDH – Integrated Diagnostics Holdings (Al Borg, Al Mokhtabar; 600+ lab branches) — Smart Village, Egypt
Data Scientist Supervisor — May 2024 – May 2026 (full-time; consultant from Apr 2025)

- Grew monthly CVM campaign impact from ~EGP 10M / 15K visits to ~EGP 40M / 45K visits (~USD 0.8M) in 3 months, lifting the annual incremental-revenue run-rate above EGP 300M (~USD 6M), by replacing rule-based list selection with ML customer segmentation.
- Built new customer models on visit behaviour, disease and test profiles, and payment behaviour to send each segment a customised offer, and A/B-tested every offer against control groups before scaling it.
- Had 15 of the first 20 recommended sites accepted with a geospatial site-selection model (branch performance, customer behaviour, competitor and healthcare-facility density, population, GDP) that replaced manual feasibility studies.
- Added ~EGP 3M (~USD 60K) a month in incremental revenue with a House Call propensity model targeting branch customers likely to convert to home visits.
- Halved reporting time with Oracle SQL ETL (stored procedures, scheduled jobs) behind automated Power BI dashboards for sales, operations, marketing and executives; supervised junior data scientists.

Credit Agricole Egypt — New Cairo, Egypt
Data Analytics Officer — Jul 2023 – Apr 2024

- Raised reporting efficiency 40% and cut manual errors by automating recurring retail-portfolio reporting in SQL, SAS and R and migrating SAS workflows to R Markdown.
- Built Power BI dashboards for portfolio performance, KPIs and policy impact used by retail leadership.
- Segmented the retail lending portfolio by product, program, sales source and dealer for monthly management reporting and policy-impact analysis.

Huawei Technologies — New Cairo, Egypt
Data Analyst — Jan 2022 – Jun 2023

- Cut executive reporting from ~4 h to under 1 min per cycle (30+ h saved a month) by rebuilding it in Power Query and Power BI.
- Owned dealer and operator incentive calculations (sell-out / sell-through) and set regional sales targets aligned to stock.
- Built competitor pricing, promotion and market-share dashboards for the commercial team.

Upwork & private clients Remote

Freelance Data Scientist & Power BI Consultant Oct 2019 – Present

- Top Rated with 100% Job Success across 58 jobs: predictive modelling, customer segmentation, Power BI dashboards, hypothesis testing and survey analytics for business and academic clients in Python, R and SPSS.
- Built the Python data and AI backend of an AI interior-design MVP (May–Jul 2026): a 5-vendor product-catalogue pipeline with data-quality checks and PostgreSQL loads, and a FastAPI middleware for room analysis and image generation.

SELECTED PROJECTS

E-Commerce Return Prediction M.Sc. research project, 2025

Framed returns as a classification problem on ~100K Olist orders with 25+ engineered features; benchmarked logistic regression, random forest, XGBoost and a neural network with class weighting and cross-validated tuning. Random forest reached ROC AUC 0.826; SHAP ranked delivery delay and freight cost as top drivers.

Customer Churn Intelligence App github.com/ayman-stat/customer-churn-intelligence-streamlit | live demo

Streamlit and scikit-learn app on synthetic data that compares models on ROC AUC and average precision, explains drivers with permutation importance and ranks customers under a retention budget; model card and pytest tests.

Real-Time E-Commerce Data Pipeline github.com/ayman-stat/Data-Engineering--DEPI

Streams Kafka events through PySpark Structured Streaming into a Bronze/Silver/Gold Parquet lakehouse and MongoDB on Docker Compose, plus a daily-order forecasting layer.

SKILLS

Machine learning: Python (pandas, NumPy, scikit-learn, XGBoost, SHAP), churn and propensity models, segmentation (RFM, clustering), next-best-offer, geospatial scoring, forecasting, NLP

Experimentation: A/B testing, target/control design, uplift and incremental-impact measurement, hypothesis testing, R

Data & BI: Microsoft Fabric (lakehouse, semantic models), Power BI (DAX, Power Query), SQL (Oracle, SQL Server, PostgreSQL), ETL, Kafka, PySpark, MongoDB

Generative AI: Amazon Bedrock (Knowledge Bases, guardrails), LLM grounding, Boto3, FastAPI, Streamlit

EDUCATION

M.Sc. Data Science (in progress), Faculty of Graduate Studies for Statistical Research, Cairo University Expected 2027

Coursework: advanced machine learning, pattern recognition, Bayesian inference, deep neural networks.

B.Sc. Statistics, Faculty of Economics and Political Science, Cairo University 2021

CERTIFICATIONS & TRAINING

Completed

- Python for Everybody Specialization – University of Michigan / Coursera (2026)
- Data Engineering Scholarship, AI & Data Science track – DEPI Egypt (2026)
- Microsoft Power BI Engineer track – DEPI Egypt (2024)
- Business Intelligence Developer, 369 hours – ITI (2023)
- Data Science Diploma – Machinfy (2022)
- IBM Data Analyst Professional Certificate – Coursera (2019); Data Analyst Nanodegree – Udacity (2019)

In progress

- Machine Learning Diploma and Deep Learning Diploma – CSkilled (Dr. Mostafa Saad)
- AWS Generative AI and AI Agents with Amazon Bedrock Professional Certificate – Coursera