

# Osama Barakat

Data Scientist | Quantitative Analyst

Email: [Osamabarakat021@gmail.com](mailto:Osamabarakat021@gmail.com) || Call Number: +201014198841 Or +201554545332

---

## Summary

**Data Scientist & Quantitative Analyst** with 2 years of hands-on technical experience in data science, machine learning, and advanced analytics, complemented by 5 years in business-oriented roles across Marketing, Sales, and Finance. Adept at bridging the gap between data-driven insights and strategic business decisions, with a strong ability to translate complex analytical findings into actionable recommendations for stakeholders.

**End-to-End Analytics & Business Intelligence:** Proven track record in designing, implementing, and delivering analytics projects—from data collection and preprocessing to statistical modeling, machine learning, and visualization. Skilled in analyzing KPIs, uncovering trends, and identifying optimization opportunities to drive business growth. Experienced in presenting insights to leadership through clear, compelling storytelling and data-driven decision frameworks.

**NLP & Text Analytics Specialist:** Expertise in Natural Language Processing (NLP), with a focus on extracting insights from unstructured text data. Proficient in sentiment analysis, topic modeling, named entity recognition (NER), and text summarization, enabling businesses to derive value from customer feedback, social media, reports, and other language-based sources. Passionate about transforming raw text into strategic intelligence for real-world applications.

**Time Series Forecasting & Predictive Modeling:** Strong background in time series analysis, including trend detection, seasonal decomposition, and forecasting using statistical and machine learning approaches. Experienced in building adaptive models for dynamic datasets, optimizing processes, and improving decision-making in finance, sales, and operational planning.

---

## I excel in the following:

- preprocessing and transforming raw data into actionable insights.
- Proficient in building and optimizing models for business solutions.
- Specialized in sentiment analysis, entity recognition, and text summarization.
- Skilled in identifying trends and seasonal patterns for dynamic data.
- Creates clear dashboards and visualizations to communicate insights.
- Designs and analyzes tests to optimize strategies.
- Bridges data science with marketing to drive ROI and campaign efficiency.
- Works seamlessly with technical and non-technical teams.
- Builds efficient pipelines and scalable models for real-time use.
- Combines technical and business acumen to address complex challenges.
- Translates technical findings into strategic recommendations.
- Quickly learns new tools and methodologies to stay ahead in dynamic fields.

## Experience

### - **Data scientist – X\_Advertising – Mansoura, Egypt**

January 2024 – Now

As a Data Scientist with 3+ years of experience and a bachelor's degree in data science, I specialize in cleaning, preprocessing, and analyzing large datasets to extract actionable marketing insights. Using Python (Pandas, NumPy, Scikit-learn) and SQL/NoSQL databases, I perform rigorous exploration data analysis (EDA) to identify customer behavior patterns, campaign trends, and operational efficiencies. My work in predictive modeling spans financial forecasting, recommendation systems, and customer lifetime value (LTV) prediction, leveraging frameworks like TensorFlow/PyTorch and AutoML tools to build high-impact solutions. I combine these technical skills with A/B testing and model explainability (SHAP/LIME) to ensure data-driven decisions align with business goals across healthcare, real estate, and e-commerce sectors.

My expertise covers the full ML lifecycle - from developing algorithms to deploying scalable solutions. I design machine learning pipelines using Scikit-learn and deep learning frameworks, then collaborate with engineering teams to integrate models into production via APIs or embedded AI systems. My strong programming skills and experience with cloud platforms (AWS/GCP) enable me to optimize model performance and resource efficiency. Beyond modeling, I translate complex results into executive strategy using data visualization tools (Matplotlib, Seaborn, Power BI, Tableau), creating dashboards that democratize insights across marketing, sales, and leadership teams. This holistic approach - combining predictive analytics, data engineering, and business intelligence - drives measurable improvements in campaign ROI and customer engagement.

Continuously advancing my skills, I stay ahead of emerging trends in AI, machine learning, and data engineering to deliver cutting-edge solutions. My work incorporates generative AI (LLMs) for content optimization and MLOps tools (MLflow, Kubeflow) for robust model deployment. As a cross-functional leader, I bridge technical and business domains - whether explaining model behavior to stakeholders using XAI techniques, mentoring teams on Python/R best practices, or prototyping innovative applications of computer vision/NLP. My strong analytical and problem-solving skills, honed through diverse projects from recommendation engines to operational analytics, position me to tackle complex marketing challenges while maintaining rigorous standards for data integrity, model accuracy, and business relevance.

### - **Apprenticeship – McKinsey & Company (Forward Program) – Cairo, Egypt**

November 2023 – April 2024

Engaged in McKinsey & Company's flagship global upskilling initiative, a rigorous, cohort-based program designed to cultivate next-generation leadership, data fluency, and structured problem-solving capabilities. Through a blend of interactive case simulations, peer collaboration, and expert-led modules, I refined my ability to deconstruct complex challenges, synthesize data-driven insights, and communicate strategic recommendations with clarity skills that directly aligned with and enhanced my third-year coursework in applied quantitative methods and systems analysis. The program's emphasis on adaptive thinking and digital transformation further equipped me to navigate interdisciplinary technical environments, reinforcing my ability to integrate analytical rigor with operational agility. This experience not only sharpened my technical and strategic competencies but also instilled a growth mindset essential for tackling advanced, solution-oriented mathematical projects.

## - **Quantitative Analyst – The Global Webs – Cairo, Egypt**

Jul 2022 – Jan 2024 (1 year 7 months)

As a Quantitative Analyst specializing in marketing optimization, I leverage statistical arbitrage and econometric modeling to refine pricing strategies and customer segmentation. Using time-series forecasting (ARIMA, Prophet) and Bayesian inference, I assess market responsiveness to campaigns, optimizing budget allocation across channels. My expertise extends to stochastic modeling for risk assessment in marketing investments, ensuring minimal volatility in ROI. Unlike traditional data science roles, I focus on high-frequency data analysis from real-time bidding (RTB) platforms and programmatic ad exchanges, applying quantitative finance principles to maximize ad spending efficiency.

A key differentiator in my work is the application of quantitative market research techniques, such as conjoint analysis and discrete choice modeling, to decode consumer preferences on a scale. I design agent-based simulations to predict how micro-level user interactions influence macro-level campaign performance—a niche skill beyond conventional predictive modeling. My models incorporate game theory principles to anticipate competitor moves in digital ad auctions, giving clients a strategic edge. Additionally, I employ Monte Carlo methods to simulate worst-case scenarios, ensuring marketing strategies remain resilient under market fluctuations.

To push the boundaries of marketing analytics, I integrate alternative data sources (e.g., sentiment analysis from earnings calls, geospatial foot traffic data) into quantitative models, uncovering hidden correlations. My toolkit includes quantile regression for outlier-resistant insights and non-parametric statistics for skewed social media datasets. Unlike ML-centric roles, I prioritize structural equation modeling (SEM) to validate causal relationships in marketing funnels. By staying at the forefront of algorithmic marketing research—from attention economics to neuromarketing data quantization, I ensure my strategies align with evolving consumer psychology and market dynamics.

## - **Technical Instructor – Cisco (CNSC) – Cairo, Egypt**

February 2024 – June 2024 (5 months)

During my final undergraduate year, I served as an instructor for a global cybersecurity education program at a leading university, specializing in AI and data science training. My role involved designing and delivering modules on data processing (cleaning, feature engineering), machine learning pipelines, and statistical modeling, while reinforcing foundational analytical skills like mathematical optimization and data-driven problem-solving. The program's interdisciplinary structure, with teams spanning AI, cybersecurity, and full-stack development, allowed me to contextualize data science within broader technical ecosystems and refine my ability to communicate complex concepts to diverse audiences.

Beyond teaching, I developed scalable data solutions for real-world applications, employing regression analysis and ensemble modeling to bridge theoretical knowledge with practical implementation. This hands-on experience not only honed my technical communication and pedagogical adaptability but also demonstrated my capacity to transform academic concepts into deployable tools a skill I now seek to enhance through advanced research.

## Projects

### - **Forecasting stock market movement using sentiment analysis**

This project, deeply rooted in the principles of FinTech, exemplifies the synergistic application of financial technology and artificial intelligence. By developing a sophisticated predictive model that forecasts stock market trends through sentiment analysis of financial textual data, it directly addresses a core FinTech objective: enhancing financial decision-making through technological innovation. The project's use of advanced NLP techniques, including fine-tuned FinBERT models, automated data pipelines for financial news and social media, and the integration of alternative datasets like Reddit forums, showcases a commitment to leveraging cutting-edge technology for comprehensive financial analysis. Ultimately, by providing actionable insights for profitable trading opportunities, building risk models, and integrating sentiment signals into algorithmic trading strategies, this project delivers tangible solutions that demonstrate FinTech's potential to generate alpha and optimize trading performance.

### - **Identifying High-Value Customers for a Marketing Campaign**

This project was instrumental in optimizing marketing resource allocation by developing a data-driven framework for identifying high-value customers, segmenting them based on potential, and predicting their future behavior. By leveraging machine learning techniques such as logistic regression, random forests, and gradient boosting, the project successfully built predictive models to forecast repeat purchases and significant revenue generation. A key outcome was a direct improvement in marketing ROI through targeted campaigns, achieved by personalizing offers and recommendations based on customer segments and their calculated Customer Lifetime Value (CLV). The project also automated the customer targeting process, enhancing campaign efficiency and providing actionable insights that empowered the marketing team to fine-tune strategies, ultimately driving long-term revenue growth and improving customer loyalty.

### - **Optimize ad-spend across multiple social media platforms**

This project successfully optimized ad-spend across diverse social media platforms by developing a dynamic budget allocation framework and predictive model, resulting in a 30% improvement in Return on Ad Spend (ROAS) and a 15% reduction in Cost Per Acquisition (CPA). Leveraging Bayesian optimization, the solution continuously adjusted ad-spend based on performance predictions, automating the optimization process for scalability. A unified data pipeline aggregated metrics from platforms like Facebook, TikTok, and YouTube, providing real-time campaign insights through an interactive dashboard. The model further incorporated time-series analysis for ad fatigue and seasonality, enhancing prediction accuracy and enabling proactive planning of creative refreshes. This strategic initiative not only maximized ROI and minimized CPA but also fostered data-driven decision-making, setting a foundation for scalable, automated ad spend optimization crucial for competitive marketing.

### - **High Viewership with low conversion Rates**

This project addressed the critical challenge of converting high social media viewership into actionable outcomes by conducting an in-depth analysis of user behavior and optimizing the conversion funnel. The objective was to enhance social media content and campaign effectiveness by identifying conversion bottlenecks and refining engagement strategies, ultimately leading to higher sales, sign-ups, or leads. Key achievements included meticulous Conversion Rate Analysis to pinpoint drop-off points, rigorous Audience Engagement Optimization through A/B testing diverse content formats and Targeting Refinement to ensure content reached the most relevant audience segments. The project also encompassed revamping Call-to-Actions (CTAs) for increased compellingness, launching conversion-focused campaigns with tailored offers, and integrating efforts across platforms for seamless user experiences. By continuously adjusting strategies based on data-driven insights, this initiative successfully unlocked the full value of high traffic, maximized marketing ROI, and provided valuable insights into audience behavior for future digital marketing strategies.

## Technical Skills & Expertise

**Programming & Development:** Python, R, SQL, HTML, Bash, Git, Jupyter Notebook, RStudio, VS Code, GitHub, FastAPI, Flask, TensorFlow Serving, Docker, Kubernetes

**Data Analytics & Business Intelligence:** Data Cleaning, Wrangling, Transformation, Validation, Automation, SQL/NoSQL, ETL Pipelines, Data Warehousing, Big Data, KPI Analysis, EDA, Ad-hoc Reporting, Descriptive Statistics, Inferential Statistics, DDDM (Data-Driven Decision Making), A/B Testing, Causal Inference, Hypothesis Testing, Feature Engineering, Dimensionality Reduction (PCA, t-SNE, UMAP), Time Series Cross-Validation, Outlier Detection.

**Machine Learning & AI:** Linear & Logistic Regression, Decision Trees, Random Forests, SVM, KNN, XGBoost, Naïve Bayes, Clustering, Ensemble Learning (Stacking, Voting), Recommender Systems, Deep Learning, AutoML, Transfer Learning, Few-Shot & Zero-Shot Learning, Model Explainability (SHAP, LIME), Model Evaluation & Tuning, Imbalanced Classification (SMOTE), Knowledge Graphs, Prompt Engineering, LLM Fine-Tuning (LoRA, PEFT), Retrieval-Augmented Generation (RAG).

**Natural Language Processing (NLP):** Text Preprocessing, Tokenization, Lemmatization, Named Entity Recognition (NER), Topic Modeling, Sentiment Analysis, Text Summarization, Text Classification, FinBERT, Transformers, Hugging Face, Prompt Engineering, LLMs, Multi-modal AI, Social Listening Analytics.

**Time Series & Forecasting:** ARIMA, Prophet, Seasonal Decomposition, Trend Modeling, Anomaly Detection, Bayesian Inference, High-Frequency Forecasting, Quantile Regression, Time Series Cross-Validation.

**Data Visualization & Reporting:** Power BI, Tableau, Matplotlib, Seaborn, Canva, Dashboards, Executive Presentations, Visual Analytics, Data Storytelling, Report Automation, Microsoft PowerPoint.

**Statistics, Math & Quantitative Methods:** Probability, Statistics, Linear Algebra, Calculus, Discrete Mathematics, Monte Carlo Simulation, Bayesian Inference, Structural Equation Modeling (SEM), Game Theory, Optimization, Quantile Regression, Econometrics.

**Cloud Platforms & Data Engineering:** AWS (S3, Lambda), Google Cloud Platform (GCP), API Integration, REST APIs, Web Scraping (BeautifulSoup, Selenium), JSON, XML, Cloud Functions, CI/CD (GitHub Actions, GitLab CI), Data Versioning (DVC), Workflow Orchestration (Airflow, Prefect), Real-time Data Processing (Kafka, Spark Streaming).

**Marketing & AdTech Analytics:** Marketing Mix Modeling, Customer Segmentation, CLV Prediction, Funnel Optimization, Campaign Attribution (First-Touch, Data-Driven), Budget Allocation Algorithms, Conversion Rate Optimization, Uplift Modeling, Ad Fatigue Prediction, Programmatic Advertising, RTB (Real-Time Bidding), Neuromarketing, SEO, Hashtag Strategy, Behavioral Segmentation, Growth Metrics (CAC, MRR, Retention).

**Tools & Platforms:** PostgreSQL, SQL Server, Microsoft Excel, Microsoft Office Suite, CRM Systems, AI Tools (OpenAI APIs, GPT), Cybersecurity Platforms (Cisco context), Canva.

**Soft Skills & Communication:** Strategic Thinking, Project Scoping, Cross-Functional Collaboration, Technical Communication, Mentorship, Teaching & Instruction, Agile Teamwork, Curiosity & Fast Learning, Systems Thinking, Design Thinking, Storytelling, Public Speaking, Business Acumen.

---

## Education

- **The Open University (UK)**
  - Bachelor of Science - BSc, Data Science
    - Mar 2021 – Dec 2024
- **Arab Open University (EG)**
  - Bachelor of Science - BSc, Data Science
    - Mar 2021 – Dec 2024
- **Politecnico Di Milano**
  - Master of Science, BSc, Geoinformatics Engineering

## Languages

- Arabic
  - Native or bilingual proficiency
- English
  - Professional working proficiency
- Italian
  - Professional working proficiency
- Germany
  - Elementary proficiency