PRINCE OKAFOR

Data Analyst | Business Intelligence Analyst

chidex.po@gmail.com | +234 706 773 9461 | LinkedIn | GitHub | Portfolio

---

PROFESSIONAL SUMMARY

Results-driven Data Analyst with 4+ years of experience transforming complex datasets into strategic business insights across healthcare, finance, and operations. Expert in building end-to-end analytics solutions—from data extraction and modeling to interactive dashboards and automated reporting systems—using Python, SQL, Power BI, and Excel. Proven track record of delivering measurable business impact: **20% revenue increase** through data-driven sales optimization, **15% improvement in data accuracy** for health programs, and analytics solutions serving **500+ stakeholders**. Skilled at translating technical findings into actionable recommendations for non-technical executives, with experience collaborating cross-functionally with leadership, procurement, operations, and field teams.

KEY ACHIEVEMENTS

- Drove 20% revenue growth by analyzing regional sales data across 15 hospitals, identifying high-performing product lines and optimizing distribution strategies
- Improved data accuracy by 15% through design and implementation of standardized data collection framework for rural health programs serving 500+ beneficiaries
- Reduced stockouts by 12% by partnering with hospital procurement teams to analyze purchase patterns and align inventory with demand forecasts
- Built predictive ML model achieving 85% accuracy for loan repayment forecasting, enabling data-driven credit risk assessment
- Developed automated data pipeline refreshing 750+ products daily with zero manual intervention, eliminating operational bottlenecks
- Identified €51M budget variance in fintech project analysis, uncovering critical bottlenecks and resource inefficiencies across departments
- Increased health program participation by 35% by analyzing vaccination coverage data and targeting outreach to underserved communities

---

TECHNICAL SKILLS

Programming & Analytics: Python (Pandas, NumPy, Matplotlib, Seaborn, scikit-learn, TensorFlow), SQL (MySQL, SQL Server)

Business Intelligence & Visualization: Power BI (DAX, Power Query), Tableau, Microsoft Excel (Advanced: Pivot Tables, VLOOKUP, Power Query)

Data Science & ML: Exploratory Data Analysis (EDA), Predictive Analytics, Machine Learning (Regression, Classification, Clustering), Statistical Modeling, A/B Testing, Time Series Analysis

Data Engineering: ETL/ELT Processes, Data Cleaning & Transformation, Data Pipeline Development, Database Design, Data Quality Assurance, API Integration, Task Automation

Analytics Techniques: KPI Development & Tracking, Customer Segmentation, Cohort Analysis, Funnel Analysis, Trend Analysis, Churn Prediction

Soft Skills: Strategic Thinking, Stakeholder Management, Data Storytelling, Cross-functional Collaboration, Problem-Solving, Executive Presentation

---

EXPERIENCE

Medical Outreach Intern (Data Analytics Specialist) | Faith Alive Foundation | 2021 – 2022

- Designed data-driven health intervention strategy by conducting systematic analysis of community health data across 12 villages, identifying 40% lower vaccination rates among children under 5 in remote areas, directly informing mobile clinic scheduling and resource allocation
- Engineered standardized data collection framework with validation rules and quality control protocols, improving data accuracy from 78% to 93% (15% improvement) and enabling reliable program evaluation for government and donor reporting
- Performed quantitative analysis on 500+ patient records using Excel, uncovering geographic coverage gaps, demographic trends, and service utilization patterns that increased program participation by 35% within 6 months
- Trained and mentored 5 field staff on data collection best practices, form completion accuracy, and quality assurance procedures, reducing missing data by 25% and improving cross-team data consistency
- Collaborated with multidisciplinary team of 8+ nurses, community health workers, and logistics coordinators to ensure protocol adherence, resolve data discrepancies within 24 hours, and meet strict donor reporting deadlines
- Generated actionable insights from complex health datasets, synthesizing findings into clear presentations for non-technical audiences including village leaders, using data storytelling to drive community engagement

Sales Representative (Data & Analytics Focus) | Chuzzbond International Medical Company | 2016 – 2020

- Drove 20% revenue increase in Q3 2019 by analyzing regional sales performance across 15 hospitals, identifying top 3 high-margin product lines, and implementing data-driven distribution and customer engagement strategies
- Built interactive Power BI dashboards with drill-through capabilities to visualize sales trends, monitor KPIs (revenue, units sold, customer acquisition cost), and track product performance by region, supporting strategic planning across sales, marketing, and operations teams
- Optimized inventory management by partnering with hospital procurement teams to analyze 18 months of purchase data, forecast demand patterns, and align supply strategies with client needs, reducing stockouts by 12% and improving fill rates
- Developed SQL-based reporting infrastructure automating weekly sales reports, customer engagement metrics, and market trend analysis for executive leadership, reducing manual reporting time by 40%
- Performed customer segmentation analysis using RFM (Recency, Frequency, Monetary) methodology to identify high-value clients, enabling targeted account management that improved customer retention by 18%

- Conducted competitive market analysis tracking competitor pricing, product positioning, and market share trends across 20+ hospitals, providing strategic intelligence that informed pricing strategy and product portfolio decisions
- Created predictive sales forecasting model using Excel regression analysis to project quarterly revenue with 92% accuracy, enabling better resource planning and inventory optimization

Research Intern (Data & Energy Analytics) | National Centre for Energy Research and Development, University of Nigeria | Jan 2022

- Supported renewable energy research initiatives at national research institution, collecting and analyzing solar panel efficiency data across varying environmental conditions (temperature, irradiance, humidity)
- Maintained research databases using Excel with data validation, version control, and quality checks for longitudinal studies spanning 12+ months
- Conducted literature research on renewable energy technologies and sustainability frameworks, compiling annotated bibliographies for technical reports and grant proposals
- Collaborated with multidisciplinary research team including engineers and policy researchers, contributing to data summaries and progress presentations for stakeholder review

---

ANALYTICS PROJECTS

Paystack Payment Analytics - Fintech SaaS Product Analysis| View Project
Technologies: Python (Synthetic Paystack API, scikit-learn, pandas), SQL Server, Machine Learning

Business Impact: Identified ₦50M+ annual revenue opportunity through churn reduction and fraud prevention

- Built end-to-end fintech analytics pipeline analyzing 17,962 payment transactions across 340 merchants, modeling Nigerian payment ecosystem (Paystack-style) to optimize merchant activation, reduce churn, and prevent fraud
- Developed Random Forest churn prediction model achieving 96% accuracy (90% recall, 0.997 ROC-AUC) identifying 33 high-risk merchants before they churned, with recency (26% importance) and transaction volume (23%) as top predictors
- Implemented Isolation Forest fraud detection flagging 360 anomalous transactions (₦4.13M fraud amount), discovering 67% of fraud occurs at midnight + weekends + amounts near ₦80K limits, enabling automated blocking rules
- Performed K-Means merchant segmentation identifying 3 behavioral clusters: Power Merchants (13.5%, driving 89% volume), Growing Merchants (73%), and Dormant Merchants (13%), enabling tiered retention strategies
- Engineered 20+ SQL analytical queries calculating MRR (₦6.3M), Net Revenue Retention (90%), logo churn (31.8%) vs revenue churn (25.5%), and merchant LTV (₦876K over 6 months), proving high-value merchants drive platform sustainability

Open Food Facts E-commerce Analytics & Automated Data Pipeline | View Project
Technologies: Python (API automation), SQL Server, Power BI, Machine Learning, Windows Task Scheduler

Business Impact: Built production-ready automated pipeline refreshing 750+ products daily with zero manual intervention

- Designed automated data pipeline using Python + Open Food Facts API + Windows Task Scheduler, executing daily at 6 AM with error handling, logging, and real-time monitoring, eliminating manual data refresh and ensuring 24/7 product catalog accuracy
- Built Random Forest churn prediction model achieving 83% accuracy (89% ROC-AUC) on 3,900 customers across 855 transactions, identifying discount usage (12.5% importance) and promo code behavior (11.5%) as top churn predictors
- Performed K-Means customer segmentation across 1,170 customers identifying 5 behavioral clusters: VIP High Spenders ($66.57 avg, 74 customers driving 22% revenue despite <2% of base), Premium Regular, Frequent Bargain Hunters, Occasional Shoppers (53% of customers, low frequency), and Budget Newcomers
- Developed content-based recommendation engine using cosine similarity for cross-selling and personalized product suggestions, analyzing 10,002 products to optimize inventory and email marketing campaigns
- Created 4 interactive Power BI dashboards integrating SQL analytics (20+ business queries) with ML predictions: revenue trends, customer insights, churn probability tables, segment characteristics, and confusion matrix visualizations for executive decision-making
- Discovered actionable insights: 27% predicted churn rate, customers without discounts 12.5% more likely to churn, VIP segment (<2%) contributes 22% revenue, potential $9,480 annual revenue saved through targeted retention

League of Legends Player Retention Analytics | View Project
Technologies: Python (Riot Games API, pandas, scikit-learn), SQL Server, Power BI, Machine Learning
Business Impact: Identified strategies to reduce 97% early churn rate, projected $161K annual revenue opportunity

- Built end-to-end analytics pipeline integrating Riot Games API to extract 100,000+ matches across 12,760 players, designing Python scripts with rate limiting, error handling, and automated daily refresh capabilities
- Engineered normalized SQL database with 4 tables (players, matches, match_participants, events), implementing complex queries for cohort retention analysis, engagement funnels, and player-level aggregations
- Developed K-Means clustering model to segment players into 3 distinct behavioral groups: Extreme Players (12.3 matches/day, 35% retention), Core Players (8.7 matches/day, 88% retention), and Casual Players (2.1 matches/day, 45% retention), achieving 0.72 silhouette score
- Built churn prediction model using Logistic Regression with 84% accuracy (0.86 AUC-ROC) to identify at-risk players within 14 days, discovering that reaching 3 matches increases retention by 8.6x (15.4% vs 1.8%)
- Created interactive Power BI dashboard with advanced DAX measures and Power Query transformations, visualizing retention funnels, ML insights, and segment-specific performance metrics for executive and product teams
- Identified catastrophic 97% first-match drop-off through cohort analysis, discovering that first-to-second match conversion (2.96%) represents the critical retention bottleneck with massive revenue impact potential

- Quantified business opportunity**:** Improving first-to-second match rate from 2.96% to 10% could generate $161,280 annual revenue based on $15 average lifetime value calculations
- Delivered 5 actionable recommendations including optimizing matchmaking quality (high-performers retain 2x better), implementing 3-match activation gamification, and deploying segment-specific retention strategies to reduce churn by 30-40%
- Solved technical challenges including API rate limiting (20/sec, 100/2min) using exponential backoff, DAX context issues in retention calculations using Power Query aggregation tables, and dashboard performance optimization (45s → 8s refresh time)

EDUCATION

Bachelor of Science in Biochemistry | University of Nigeria, Nsukka | 2017 – 2023

CERTIFICATIONS

- **SQL for Data Analysis** | View Certificate
- **Microsoft Excel for Data Analytics** | View Certificate
- **Python for Data Science and Machine Learning** | View Certificate