

# Lakshmi Kantha Reddy Attunuru

## Data Analyst

Herndon, VA | (202) 930-1699 | lakshmi.kantha.reddy01@hotmail.com | LinkedIn | GitHub



### SUMMARY

Data Analyst with almost 4 years of experience delivering data-driven insights through advanced analytics, statistical modeling, and scalable data pipelines. Strong expertise in SQL, Python, R, and BI tools to analyze complex datasets and build executive dashboards. Proven ability to design and optimize ETL workflows, data warehouses, and cloud-based analytics solutions across AWS, Azure, and GCP. Experienced in EDA, predictive modeling, A/B testing, and KPI tracking to support business growth, retention, and operational efficiency. Adept at collaborating with cross-functional teams to translate business requirements into actionable insights with measurable impact.

### SKILLS

**Programming & Analysis:** Python (Pandas, NumPy, SciPy, Statsmodels, scikit-learn, TensorFlow), R, SQL (Advanced Queries, Joins, CTEs, Window Functions), SAS, Jupyter Notebook, Data Cleaning, Exploratory Data Analysis (EDA)  
**Data Pipeline & Integration:** Airflow, dbt, Azure Data Factory, AWS Glue, Databricks, Data Modeling, Data Warehousing, Data Ingestion Pipelines, API Integrations, Workflow Automation, Feature Engineering, Hypothesis Testing  
**Databases & Clouds:** Snowflake, AWS Redshift, PostgreSQL, MySQL, BigQuery, MongoDB, AWS S3, CSV/JSON Processing  
**Data Visualization & Business Intelligence:** Tableau, Power BI, Excel (Advanced), Streamlit, Plotly, Matplotlib, Seaborn, Executive KPI Dashboards, Financial & Risk Reporting, Trend Analysis, Data Storytelling  
**Machine Learning & Predictive Analytics:** Regression & Classification Models, XGBoost, Time-Series Forecasting (LSTM), Clustering, Ensemble Methods, Model Validation, Cross-Validation, Model Evaluation (ROC-AUC, F1, Precision/Recall)  
**NLP & Text Analytics:** Sentiment Analysis (FinBERT, DistilBERT), Retrieval-Augmented Generation (RAG), Text Embeddings (Sentence-Transformers), Document Intelligence  
**MLOps & Model Management:** MLflow, Model Versioning & Monitoring, Performance Tracking, Experiment Logging  
**Data Governance & Quality:** Data Validation, Reconciliation, Audit Controls, Data Quality Checks, Data Drift Detection (Evidently AI), Regulatory Compliance (SOX, PCI-DSS), Governance Standards  
**Tools & Collaboration:** Git, GitHub, Docker, GitHub Actions (CI/CD), Azure ML, Python 3.11, Loguru, Agile/Scrum

### EXPERIENCE

#### Data Analyst | PNC Financials

May 2025 – Present

- Designed scalable ETL pipelines using SQL, Python, Airflow, and dbt to integrate multi-source financial, transactional, and risk data, improving data availability, accuracy, and analytics scalability by 35% across banking operations.
- Built executive dashboards using Tableau, Power BI, and Python (Matplotlib, Plotly), delivering real-time financial KPIs, risk metrics, and performance indicators for business teams, accelerating data-driven decision-making by 40%.
- Collaborated with risk and data science teams on feature engineering, validation, and A/B testing using Databricks, TensorFlow, and PostgreSQL, improving model performance by 20% while ensuring SOX and PCI-DSS compliance.
- Ensured enterprise-grade data quality, controls, and governance across Snowflake, Redshift, SAS, and MySQL by implementing automated validation, reconciliation, and audit checks, reducing financial reporting discrepancies.
- Performed advanced ad-hoc and exploratory financial analysis using Excel, SciPy, Statsmodels, AWS S3, and Redshift, generating actionable insights, improving customer profitability visibility and identifying key drivers of credit risk.
- Collaborated with finance and risk teams in Agile environments to convert regulatory requirements into KPIs and automated reports using APIs, R, and Azure Data Factory, reducing manual reporting effort by 40%.

#### Data Analyst | Avenir Technologies

Sep 2020 – Jul 2023

- Analyzed large-scale datasets using Python and R, applying statistical modeling, hypothesis testing, and segmentation, improving marketing targeting precision by 50% and accelerating retention response strategies.
- Designed and optimized SQL-based ETL pipelines and data models across Snowflake, AWS Redshift, and BigQuery, increasing data scalability by 45% and reducing reporting latency by 40%.
- Built and validated regression and classification models using Python, scikit-learn, TensorFlow, and XGBoost, deploying solutions on AWS and MongoDB, achieving 35% prediction accuracy and enabling real-time monitoring.
- Developed interactive dashboards and executive reports using Tableau, Power BI, SQL, and Excel, tracking KPIs such as churn, revenue, and fraud risk, driving 30% faster executive insights and improved decision-making.
- Performed advanced ad-hoc and exploratory analysis using Excel, SQL, and Python, uncovering hidden trends and growth drivers, reducing reporting cycles by 40% and identifying 15% more revenue opportunities.
- Automated data workflows and quality checks using AWS Glue, Airflow, SAS, Alteryx, and Azure ML, applying clustering and ensemble techniques, improving customer retention by 40% while enhancing analytics scalability.

## PROJECTS

### 1. End-to-End MLOps & Analytics Platform (Churn & Fraud Analytics)

- Designed complete analytics pipeline from synthetic data generation to model serving using XGBoost, PyTorch, and Scikit-learn, tracking metrics such as ROC-AUC, F1, and precision/recall.
- Built REST APIs with FastAPI and monitoring dashboards in Streamlit to deliver real-time model predictions and business KPIs.
- Implemented MLflow model registry, CI/CD (GitHub Actions), and drift detection, aligning with enterprise ML governance and lifecycle management standards.

### 2. RetailIQ – AI Retail Intelligence Platform

- Built a local-first retail analytics platform using FastAPI, Streamlit, SQL/Pandas, and ChromaDB, enabling natural language querying of inventory, supplier, and customer data; improved data accessibility for non-technical users.
- Developed LSTM demand forecasting models (PyTorch) and logged experiments in MLflow, delivering data-driven inventory planning insights and trend predictions.
- Implemented data quality & drift monitoring (Evidently AI) with automated reporting, improving data reliability and supporting production-grade analytics workflows.

### 3. Financial Market Intelligence Platform

- Engineered financial data pipelines using Pandas, NumPy, and yfinance, creating 15+ technical indicators (RSI, MACD, Bollinger Bands) for market trend analysis.
- Trained and evaluated LSTM time-series forecasting models (PyTorch) for stock price prediction with proper chronological validation to prevent data leakage.
- Integrated FinBERT sentiment analysis (Transformers) to analyze financial news and built interactive dashboards using Streamlit & Plotly for executive-ready visual insights.

## EDUCATION

**Master of Information Technology Management** | *Webster University*