

# Lakshmi Kantha Monisha Pulavarthy

## Data Analyst

Hillsboro, OR, USA • Phone: 971-450-2699 • Email: plkmonisha@gmail.com • LinkedIn

### Professional Summary

Spearheaded transformative data analytics initiatives with over five years of experience in architecting robust ETL pipelines, developing advanced data visualizations, and conducting sophisticated statistical analyses. Optimized operational efficiency in healthcare, financial, and e-commerce sectors by leveraging Python, SQL, AWS, and Tableau, achieving efficiency gains of up to 20%. Committed to delivering actionable insights to drive strategic decision-making and business success.

### Technical Skills

- **Programming Languages:** Python (Pandas, NumPy, Scikit-learn, Seaborn, Matplotlib, Regex, BeautifulSoup), R (dplyr, ggplot2), SQL (T-SQL, PL/SQL), Shell scripting
- **Databases:** PostgreSQL, SQL Server, MySQL, Oracle, MongoDB, Firebase
- **Data Cleaning Tools:** Power Query, Excel (Pivot Tables, VLOOKUP, XLOOKUP, INDEX-MATCH), Python (outlier detection, regex)
- **Statistical Analysis:** T-tests, ANOVA, Linear Regression, Logistic Regression, K-means Clustering, PCA, Chi-square
- **Big Data Technologies:** PySpark, Hadoop, Hive, Kafka, Databricks
- **Machine Learning Frameworks:** Scikit-learn, XGBoost, Decision Trees, Random Forests, ARIMA, Prophet
- **ETL Tools:** Apache Airflow, Talend, SSIS, AWS Glue, Alteryx, Informatica, REST APIs
- **Data Warehousing:** Snowflake, BigQuery, Amazon Redshift, Teradata
- **Visualization Tools:** Tableau, Power BI, Looker, QlikView, Google Data Studio, Plotly, Dash
- **Cloud Platforms:** AWS (S3, Redshift, Athena, Glue, Lambda, QuickSight), GCP (BigQuery, DataPrep), Azure (Data Factory, Blob Storage)
- **Data Governance:** Great Expectations, SQL assertions, RBAC, data lineage tracking
- **Productivity Tools:** JIRA, Confluence, Trello, Notion, ServiceNow, MS Office, SharePoint

### Professional Experience

#### Data Analyst, Tenet Healthcare, USA

January 2025 – Present

- Engineered dynamic Tableau dashboards using Amazon Redshift and S3 datasets to monitor critical clinical KPIs, such as patient throughput and treatment outcomes, improving reporting precision by 20% and reducing decision-making latency by 12 hours weekly.
- Orchestrated sophisticated ETL pipelines with Python (Pandas, NumPy, Regex), AWS Glue, and Athena, processing over 750K healthcare records daily, integrating external data via FHIR APIs and advanced web scraping with BeautifulSoup and Selenium.
- Executed rigorous statistical analyses, including T-tests, ANOVA, and Chi-square, while developing machine learning prototypes with Scikit-learn and Prophet, reducing patient no-show rates by 15% and optimizing outpatient demand planning by 10%.
- Automated ETL workflows using AWS CloudWatch and Step Functions, implementing robust data validation with Great Expectations and SQL assertions, achieving 99.9% data accuracy for regulatory compliance.
- Strengthened data governance by enforcing RBAC, IAM policies, and PII masking across Redshift, DynamoDB, and QuickSight, reducing compliance risks by 35% and ensuring HIPAA adherence.
- Authored comprehensive data dictionaries, QA SOPs, and ETL specifications in Confluence and SharePoint, streamlining audit processes and cutting stakeholder alignment time by 18 hours per audit cycle.
- Collaborated with cross-functional clinical teams to integrate predictive analytics into workflows, reducing patient wait times by 15% and enhancing operational throughput by 12%.

#### Data Analyst, Hexaware, Hyderabad, India

December 2022 – July 2023

- Developed advanced Excel-based financial models using Power Query, INDEX-MATCH, and pivot tables, automating Snowflake-to-Excel pipelines, reducing AR/AP reconciliation time by 30% and saving 12 hours weekly.
- Integrated treasury and billing datasets using SQL (CTEs, window functions), dbt, and Python, orchestrating daily extracts of 150K+ records via AWS Data Pipeline and Lambda, enabling real-time financial reporting.
- Formulated validation checks in Excel and Snowflake, applying chi-square tests and regression analyses to uncover profitability insights, increasing ARR by 10% and visualized through impactful Power BI and QuickSight dashboards.
- Bolstered data security with AWS Lake Formation and Snowflake RBAC policies, designing KPI dashboards (ARR, MRR, churn) in

QuickSight, improving query performance by 45% using Athena optimizations.

### **Data Analyst, IGT India Pvt Ltd, Hyderabad, India**

March 2019 – January 2022

- Engineered scalable ETL pipelines using AWS Glue, Lambda, and Step Functions, enabling real-time customer experience analytics in Snowflake data marts, processing 1.2M+ records daily for operational and financial reporting.
- Automated complex Excel reconciliations with Power Query and dynamic charts, applying K-means clustering and logistic regression in Python, improving customer retention by 12% and upselling revenue by 8%.
- Developed interactive Power BI and Google Data Studio dashboards with Row-Level Security, achieving a 65% increase in stakeholder engagement and reducing report generation time by 15 hours monthly.
- Conducted in-depth ANOVA and ARIMA analyses on loyalty program data, validating 99.95% of event logs with SQL assertions and Great Expectations, ensuring compliance with industry regulations.
- Managed A/B testing pipelines using AWS Batch and Athena, standardizing governance with RBAC and PII masking across Snowflake and S3, reducing data breach risks by 30%.

### **Data Analyst, Amazon India Pvt Ltd, Hyderabad, India**

August 2017 – August 2018

- Analyzed 3M+ e-commerce datasets with SQL, Python (Pandas, NumPy), and Excel, processing data from SAP HANA, Redshift, and S3, optimizing operational efficiency by 18% and reducing delivery delays by 10%.
- Automated ETL workflows with AWS Glue, Step Functions, and Data Catalog, delivering QuickSight and Tableau dashboards to track KPIs like returns and delivery delays, cutting reporting time by 25 hours weekly.
- Processed clickstream and engagement logs using PySpark on AWS EMR, optimizing Snowflake data models with RBAC and partitioning, improving query performance by 35%.
- Performed anomaly detection and data validation with Python and SQL, documenting comprehensive QA SOPs in Confluence, reducing audit preparation time by 12 hours per cycle.

## **Education**

### **Master of Science, Information Systems**

September 2023 – May 2025

Northeastern University, Boston, MA

- Relevant Coursework: Advanced Database Systems, Machine Learning, Cloud Computing

### **Bachelor of Science, Computer Science and Engineering**

October 2013 – May 2017

Jawaharlal Nehru Technological University, India

- Relevant Coursework: Data Structures, Algorithms, Database Management Systems

## **Certifications**

- Google Data Analytics Professional Certificate, Coursera, 2023
- Microsoft Power BI Data Analyst Associate, Microsoft, 2023
- AWS Certified Data Analytics – Specialty, Amazon Web Services, 2022
- IBM Data Science Professional Certificate, Coursera, 2022

## **Key Projects**

- Kansas City Service Requests** (Jan 2024): Transformed 75K+ raw service records using Alteryx and SQL, delivering Power BI and Tableau dashboards that reduced query response time by 40% and supported data-driven urban planning decisions.
- Sports League Management System** (Dec 2023): Designed a SQL database with triggers and stored procedures, integrating Tableau and Streamlit for real-time performance tracking, improving analytics access by 45% for league administrators.
- Food Establishments Inspections** (Feb 2024): Built a Talend-based data pipeline processing 12K+ inspection records, visualizing compliance trends in Tableau and Power BI, reducing analysis time by 25 hours monthly and enhancing regulatory reporting.
- Zomato Data Analysis** (Apr 2024): Conducted exploratory data analysis on 20K+ restaurant records using Linear Regression and Decision Trees in Python, identifying cost-rating trends that improved predictive accuracy by 20% and informed business strategies.