

Rajeev Tripathi

Cincinnati, OH | (513)-266-3982 | rajeev.tripathi0207@gmail.com | [LinkedIn](#) | [Tableau](#) | [Portfolio](#)

PROFILE

- Insight-driven Data Analyst with 3+ years of experience applying statistical modeling and visualization techniques to uncover business trends, optimize performance, and guide strategic decisions across different domains.
- Proficient in **Python, SQL, and Power BI** for data analysis and insight generation, with experience using **Azure Databricks** to build scalable analytical workflows that support data-driven decisions and improve business performance.

EXPERIENCE

| | |
|--|------------------------|
| Data Analyst Intern Lightship Foundation, Cincinnati, OH | May 2025 – Present |
| • Designed structured streaming pipelines in Databricks using Pyspark to integrate data from multiple sources for KPI reporting. | |
| • Created data validation scripts in Python and SQL to ensure consistency, completeness, and accuracy across data layers. | |
| Data Analyst University of Cincinnati, Cincinnati, OH | August 2024 – May 2025 |
| • Performed EDA to detect error occurrence patterns and device feature correlations influencing ticket volume. | |
| • Collected and cleansed large-scale customer and ticket datasets using PySpark to prepare for analysis. | |
| • Built Power BI dashboards to track ticket volume, error types, and streaming issues in real-time. | |
| • Delivered actionable insights to senior management, contributing to a 19% reduction in support tickets . | |
| Business Intelligence Developer COGNIZANT, Delhi, India | June 2021 – July 2024 |
| <i>Client: Global Security Operations Center</i> | |
| • Designed and deployed ETL pipelines in Azure Data Factory to integrate and transform data from multiple on-prem and external sources, improving overall data load performance by 34% . | |
| • Automated ingestion workflows using ADF pipelines, enabling incremental loads and reducing manual intervention. | |
| • Implemented robust data quality checks and logging mechanisms in ADF to monitor pipeline health and performance. | |
| • Designed Power BI dashboards to monitor attack and phishing incidents across 170+ member firms. | |
| • Implemented Row-Level Security (RLS) in Power BI dashboards to restrict data visibility based on region and country. | |
| • Collaborated with business stakeholders to define key metrics, ensuring analytics outputs aligned with requirements. | |
| <i>Client: SBI (General Electric)</i> | |
| • Implemented the ETL process in Alteryx reducing data extraction processing time by 22% . | |
| • Optimized SQL queries and database structures to improve performance across high-volume transactional tables. | |

SKILLS AND CERTIFICATIONS

Tools: Databricks, Streamlit, Power BI (DAX, Power Query), SQL (CTEs, Window Functions, Query Optimization), Python (Pandas, NumPy, Scikit-learn), PySpark, R, Alteryx, Advanced Excel (Power Pivot, VBA), MongoDB, Tableau

Analytics & Machine Learning: Time series Analysis, Forecasting, Clustering, A/B Testing, Data Modeling, Data Visualization

Certifications: Power BI Data Analyst ([Microsoft](#)), Data Analytics ([Google](#)), Azure Data Fundamentals ([Microsoft](#))

PROJECTS

| | |
|---|--------------------------|
| Multiple Disease Prediction System <i>Python, Streamlit, Scikit-learn, NumPy, Pandas, Pickle</i> | (Link) |
| • Developed predictive machine learning models achieving 88%+ accuracy and deployed a Streamlit-based application with CI/CD pipelines to automate model updates and releases.. | |
| AI-Powered Multi-Agent Research Assistant <i>Python, Langflow, AstraDB, Git, Streamlit, RAG</i> | (Link) |
| • Built a multi-agent LLM app that ingests health data, runs a RAG pipeline, and answers real-time research queries. | |
| • Implemented “Ask AI” functionality with agent routing through LangFlow and real-time data retrieval using Astra DB. | |
| Data Wrangling: Data-Driven Growth Opportunity Analysis <i>R, SQL, Tableau, Business Intelligence</i> | (Link) |
| • Developed predictive machine learning models to identify diabetes, heart disease, and Parkinson’s disease with over 88% accuracy via data preprocessing and deployed the app on Streamlit, ensuring CI/CD automation. | |

EDUCATION

| | |
|---|---|
| University of Cincinnati, Carl H. Lindner College of Business (GPA: 3.95/4) <i>Master of Science, Business Analytics and Data Science</i> | Cincinnati, Ohio August 2024 – August 2025 |
|---|---|

- **Relevant Courses:** Data Management, Predictive Analytics, Data Modeling, Data Mining, Intelligent Data Analysis
- **Scholarships:** Lindner College of Business International Outreach (2024), UC Global Welcome Scholarship award (2024)