**Sri Jana**

+1 (469)737-0562 | jlekhya1907@gmail.com |[LinkedIn](https://www.linkedin.com/in/j-sri/)

 **PROFESSIONAL SUMMARY**

* Experienced Data Analyst with 4+ years of expertise in building scalable ETL pipelines, designing data workflows, and delivering actionable insights using PySpark, SQL, PL/SQL, and Power BI to create scalable ETL pipelines, develop data workflows, and deliver actionable insights.
* Adept at implementing effective data processing and storage solutions by utilizing cloud platforms, notably Azure (Data Factory, Synapse, Databricks) and AWS (Redshift, S3, Glue).
* Skilled in statistical analysis and integrating data from APIs, using Power BI to create interactive dashboards and visualizations that provide actionable insights for stakeholders and support data-driven business decisions.
* Basic knowledge of machine learning for anomaly detection and predictive analytics using Python libraries like scikit-learn, Pandas, Matplotlib, and NumPy.
* Strong understanding of data modeling and integration, including dimensional modeling to improve database scalability and performance.
* Skilled in automating repetitive tasks and ETL workflows using PySpark, improving data processing efficiency.
* Become familiar with CI/CD pipelines and workflow orchestration using Airflow, Docker, and Kubernetes to streamline data engineering processes.
* Knowledgeable in data governance and security practices, ensuring compliance with GDPR and role-based access controls.
* Hands-on experience with Apache Spark and Kafka for real-time data streaming and processing.

**TECHNICAL SKILLS**

* Programming Languages: Python, SQL, PL/SQL, C, Java
* Data Management & Analytics: MySQL, Oracle, Snowflake, PostgreSQL, MongoDB, MS-SQL Server
* ETL & Data Processing: PySpark, Informatica PowerCenter, Apache Kafka, DataStage, Alteryx
* Business Intelligence & Visualization: Power BI, QuickSight, Tableau
* Cloud & Platforms:

AWS: EC2, S3, Redshift, Athena, Glue

Azure: Databricks, Synapse, Data Lake, ADF, DBFS, Azure Data Studio

Development Environments: PyCharm, IntelliJ, Jupyter Notebook

Operating Systems: Linux, Unix

* Architecture & Frameworks: OLAP, Microservices, Monolithic
* Version Control & DevOps: Git, GitHub, Jira, Putty, Docker, Kubernetes, REST API, CI/CD Pipelines
* Productivity & Collaboration Tools: MS Office Suite (Excel, PowerPoint, Word, OneNote, OneDrive), SharePoint

# PROFESSIONAL EXPERIENCE

**Circle K | Data Analyst** Tempe, Arizona | Sep 2023 - Present

* Automated ETL operations using Python, SQL, Snowflake, and Azure Data Lake to support and build a pricing application for retail products, maintaining scalability and efficiency in data integration.
* Developed and optimized PySpark scripts to process large-scale pricing data, improving ETL efficiency and reducing processing time.
* Implemented parallel processing in PySpark to handle high-volume data transformations, ensuring faster data ingestion and validation.
* Utilized PySpark DataFrames and RDDs to clean, aggregate, and transform raw data for real-time analytics.
* Integrated PySpark with Snowflake and Azure Data Lake, enabling seamless batch processing and data movement.
* Applied PySpark caching and partitioning techniques to optimize data retrieval, reducing query execution time by 30%.
* Designed PySpark-based ETL workflows to automate pricing model updates, ensuring accurate and timely insights for Category Managers.
* Built reusable PySpark functions for data transformation, reducing redundant code and enhancing pipeline scalability.
* Implemented error handling and logging mechanisms in PySpark, ensuring better monitoring and debugging of ETL pipelines.
* Developed and deployed ETL data pipelines using Azure Data Factory (ADF), Snowflake, and Python for seamless data processing.
* Extracted and integrated data from Azure Data Studio, Snowflake tables, Excel, and Azure Data Lake APIs to facilitate reporting and analysis.
* Used advanced DAX formulas in Power BI to build and deploy 13+ interactive pricing dashboards and calculate financial metrics, helping Category Managers make better decisions and improve pricing strategies at Circle K.
* Generated comprehensive monthly reports, including KPIs, financial metrics, and variance analyses, providing actionable insights to support strategic decision-making.
* Automated 90% of pricing application processes by integrating Snowflake’s robust SQL capabilities and Azure Data Lake pipelines, ensuring consistency and efficiency in pricing updates.
* Collaborated with cross-functional teams to gather requirements, align projects with business objectives, and enhance data-driven strategies.
* Partnered with data science teams to deliver processed data for machine learning models, improving predictive insights.
* Ensured regulatory compliance and data security by implementing best practices in Snowflake and Azure environments.

**Southern Arkansas University | Graduate Assistant**Magnolia, Arkansas | August 2022 - August 2023

* Developed and deployed AI-driven applications and data analytics models, integrating machine learning, NLP, and cloud-based APIs to solve real-world problems.
* Built a single-view AI image generator using Xcode, SwiftUI, and OpenAI’s DALL-E API, providing smooth user experience for text-to-image generation on a mobile.
* Conducted exploratory data analysis (EDA) and machine learning modeling on a 10GB dataset, streamlining security clearance evaluations using Decision Trees and NLP techniques such as stop-word removal and sentiment analysis.
* Designed and implemented feature engineering strategies using Pandas, NumPy, and Scikit-learn, visualizing insights with Matplotlib and Seaborn.
* Developed a real-time security clearance prediction model with an interactive HTML-based interface, improving decision-making efficiency.
* Created a containerized network security framework using Docker, Linux, and Python, enabling real-time packet analysis with Wireshark and Scapy for cybersecurity research.
* Simulated and analyzed security vulnerabilities in MySQL databases using Metasploit, enhancing understanding of database protection strategies.
* Assisted 70+ freshmen in learning web programming (HTML, CSS, JavaScript, PHP), improving project completion rates and technical proficiency.
* Graded over 200 student assignments with strict adherence to grading guidelines and provided constructive feedback for academic growth.
* Helped Professor in designing an e-learning platform for RDBMS courses, increasing student engagement and accessibility through interactive database management tools.

**Cognizant | Programmer Analyst**  *Bangalore*, India | December 2020 – May 2022

* Developed and implemented software applications using enterprise data warehousing techniques to efficiently track and manage product orders, enhancing data management and lowering manual tracking mistakes.
* Created and optimized ETL processes utilizing Informatica PowerCenter transformations, workflows, mappings, and mapplets to improve data processing and integration.
* Created and optimized database structures, including materialized views and SQL queries, to improve query performance and system efficiency.
* Developed and maintained Informatica SCD Types I, II, and III solutions to improve data history tracking and reporting accuracy.
* developed and deployed scalable data solutions for cloud-based processing and data migration on AWS, involving Glue, Lambda, EC2, VPC, and Redshift.
* conducted data migration and distribution testing to verify data accuracy and ensure smooth transitions across business systems.
* Worked with flat files, Oracle DB, MySQL, and PostgreSQL as data sources, ensuring accurate execution of data extraction, transformation, and loading (ETL) processes.
* Developed PL/SQL scripts and optimized Oracle SQL queries for high-performance data retrieval.
* Worked with the Data Warehouse team to create dimensional models and examine ER diagrams to improve database design and created PL/SQL triggers, stored procedures, and functions to help with database automation and operational efficiency.
* Built Shell Scripts to automate database tasks and system monitoring.
* Workflows were optimized by strategically positioning filters, optimizing join processes, and assuring effective session monitoring.
* Provided technical mentoring, tested, and fostered knowledge sharing throughout the team.
* Created interactive Power BI dashboards for tracking product orders and warehouse KPIs, enabling real-time reporting and analytics.

# EDUCATION

# Southern Arkansas University – 4.0 CGPA May 2022 – August 2023

Master of Science, Computer and Information Science Magnolia, Arkansas

Coursework: Data Analytics, Advanced Programming, Soft. Engineering, Computer Networking, Machine Learning

* Mahatma Gandhi Institute of Technology – 3.7 CGPA June 2017 – May 2021

Bachelor of Technology, Electrical and Electronics Engineering Hyderabad, India

Relevant Coursework:Data Structures, C, OOP in JAVA, Linear algebra, MATLAB/Simulink, Probability and Statistics

# CERTIFICATIONS

* DP-300: Administering Microsoft Azure SQL Solutions (DBA), Udemy, Aug 2024
* AWS Cloud Practitioner and Foundation, AWS, Feb 2024
* SQL Programming, LinkedIn Learning, Aug 2023
* Microsoft SQL Server 2022 Essentials, LinkedIn Learning, Aug 2023
* Oracle SQL Certification, Udemy, Mar 2022
* Power BI Essentials, LinkedIn Learning, Dec 2021
* Introduction to NLP-AI, Udemy, Dec 2021
* Introduction to Cloud Computing, Udemy, Dec 2021
* Complete PL/SQL, Udemy, Aug 2021