Nandakishor Reddy Pulagam

Lead Data Engineer| Dallas, Texas

215.452.9969| [pnkr.492@gmail.com](mailto:pnkr.492@gmail.com) | [LinkedIn](https://www.linkedin.com/in/kishor-pulagam/)

# Professional Summary

Senior Data Engineer with 10+ years of experience designing end-to-end data solutions for established corporations. Proven ability to deliver value by designing scalable data pipelines, crafting predictive models, and unlocking insights that empower businesses to succeed in today's data-driven world.

# Skills & abilities

* **Programming Languages**: Python, PySpark, SQL, Spark-SQL, Scala, R, PL-SQL, T-SQL
* **Frameworks**: Spark, Spark Structured Streaming, Kafka
* **Cloud Skills**: Azure, Databricks, Azure Synapse, AWS
* **Databases**: Mongo DB, Azure SQL DB, SQL Server, Oracle, Mongo DB, Data Lakes, Data warehouses
* **Tools**: ADF, Datadog, Airflow, Power BI, Tableau, Power Automate
* **Version Control:** Azure DevOps, Git, Confluence, Jira
* **Professional Experiences**: Team Lead, Data Engineer, Data Scientist, Data Architect, Data Analyst, Scrum Master.

# Experience

## DXC Technology | Senior Data Engineer Jun 2021 – Present

* Led a team of data engineers in developing and implementing a new data lake house, improving data performance and scalability by 30%.
* Architected and implemented end-to-end data pipelines using Azure and Databricks, facilitating large datasets efficient extraction, transformation, and loading (ETL) for multiple retailing clients.
* Optimized data pipelines using Spark, achieving significant performance gains, reducing runtime by 40% and cost by 20%.
* Developed and implemented a data ingestion strategy to support real-time and batch data ingestion using Spark structured streaming, improving timeliness and accuracy of data reporting.
* Implemented the CI-CD pipelines in Azure DevOps to integrate seamlessly with Azure Databricks and ADF, ensuring efficient and reliable deployment of data processing pipelines.
* Practices involve leveraging scalable Azure services such as Azure Databricks for distributed data processing, Azure Synapse Analytics for high-performance analytics, and Azure Data Lake Storage for scalable data storage.
* Expert in implementing comprehensive monitoring and alerting solutions to ensure data pipeline reliability and availability.
* Implemented delta tables and scheduled automated jobs in Databricks, ensuring timely and accurate data extraction from diverse sources.

## DXC Technology | Senior Data Science & Engineering Consultant Sep 2019 – Jun 2021

* Led the development of a predictive maintenance model for an Aviation client, reducing unplanned aircraft downtime by 20%.
* Created and open-sourced a Python package, DXC-Industrialized-AI-Starter, to help data scientists with end-to-end data science tasks.
* Developed an algorithm that improved the accuracy of estimating man-hours needed to fix Airbus defects by 15%.
* Successfully implemented the Medallion Architecture, providing a unified and structured approach to store data from multiple sources, facilitating easy access and analysis.
* Practices include leveraging Microsoft's integration capabilities such as Azure Logic Apps and Azure Functions for orchestrating data workflows and implementing custom data transformations. Utilizing industry-standard formats and protocols for data interchange
* Conducted thorough data analysis and quality checks, ensuring the reliability and consistency of data, and implemented industry best practices to optimize data processing and analysis workflows.

## Computer Sciences Corporation | Data Engineer Aug 2014s – Sep 2019

* Designed, developed, and maintained an Enterprise Data Warehouse using Microsoft SQL Server to support healthcare data management and analytics.
* Created and managed staging (stg) and Operational Data Store (ODS) environments to ensure data integrity, consistency, and accessibility.
* Created multiple Oracle Apex applications for a healthcare client and provided database support for different applications.
* Optimized database performance by implementing indexing, partitioning, and query optimization techniques, resulting in faster data retrieval and processing times.
* Experience in driving architectural discussions, making design decisions, and providing technical guidance to ensure alignment with business goals and industry best practices.
* Collaborated with cross-functional teams to design the databases and enhance the overall system performance and data retrieval efficiency.
* Worked on several healthcare and manufacturing domain automation projects and developed multiple solutions that helped automate several daily tasks using RPA tools and reduced costs.

# Education

## Drexel University, Philadelphia | MS Data Science | CGPA-3.9 June 2021

* Student leader in Drexel Graduate IEEE and Pragati Associations.
* Member of Drexel AI, Graduate Student Association, Drexel Blockchain, Drexel International Student Association, and Dragons Teach Students Teaching Engaged Minds.

# Certifications

* Databricks Certified Data Engineer Professional.
* Microsoft Certified: Azure Solutions Architect Expert
* Azure Data Engineer Associate certified by Microsoft.
* Certified in Databricks Fundamentals
* Certified in Kafka Fundamentals
* Certified in Azure AI, Data Fundamentals
* Google-certified Data Analytics Professional
* Certified in MongoDB Basis and MongoDB for SQL Pros from MongoDB
* IBM Certified in Machine Learning with Python and Big Data

# Achievements

* Selected by DXC Technology to participate in their prestigious master's sponsorship program in the USA, earning a master's degree in data science debt-free.
* Instructed and developed the skills of over 250 employees in data engineering and data science through immersive boot camp training.