

Pranith Kumar Dongari

Boca Raton, Florida, USA | +1-561-757-8003 | pranithkumardongari3@gmail.com | LinkedIn

SUMMARY

Data Engineer with 2+ years of experience developing scalable ETL pipelines and optimizing data workflows for improved quality and performance. Proficient in SQL, Python, and Big Data frameworks, with a proven record of automating and streamlining data integration processes. Skilled in leveraging distributed computing to deliver reliable, real-time insights that empower data-driven decisions.

EDUCATION

Florida Atlantic University

10/2023 - 05/2025

Master of Science, Information Technology Management

- **GPA:** 4.0/4.0
- **Achievements:** Dean's List

EXPERIENCE

9thPixel Technologies India Pvt.Ltd

03/2021 - 06/2023

Data Engineer

Hyderabad, India

- Designed and implemented scalable ETL pipelines in Azure Data Factory (ADF) to ensure reliable data integration and consistency across multiple sources, reducing manual data handling by 20%.
- Optimized ADF activities, including copy, lookup, and metadata processes with stored procedure integration, accelerating data processing speeds by 35% and enhancing data quality.
- Developed reusable, parameterized ADF pipelines for scalable workflows, reducing future development time by 40%.
- Engineered PySpark transformations in Databricks to streamline large-scale data processing using distributed computing, ultimately enhancing BI reporting efficiency.
- Collaborated with cross-functional stakeholders to deliver real-time analytics solutions, supporting data-driven decision-making and contributing to a 5% revenue increase.
- Migrated legacy data systems to Azure Synapse, achieving a 25% improvement in query performance for downstream analytics teams.
- Developed robust logging and error-handling mechanisms in ADF pipelines to improve monitoring and reduce resolution time by 50%.
- Mentored team members on PySpark and Azure best practices, fostering productivity and effective cross-functional collaboration.

PROJECTS

Databricks End-to-End ETL Pipeline

GitHub

Personal Project

- Created dimensional data models, governing customer and product information using SCD Type 1 & 2, and enhanced the FactOrders table, leading to a 20% improvement in data retrieval speeds.
- Orchestrated ingestion to transformation using Databricks Workflows and parameterized notebooks.
- Fashioned core data warehouse components, applying SCD Type 1 & 2 to DimCustomer and DimProduct, and built FactOrders tables, completing 100% of assigned data modeling tasks ahead of schedule.
- Processed 100K+ records daily using Delta Live Tables (DLT) for real-time and incremental ETL.
- Transformed raw data into a clean, deduplicated, and surrogate-keyed data warehouse, accelerating analytics and BI reporting cycles by 40% and enabling faster decision-making across departments.

TECHNICAL SKILLS

- **Big Data & ETL Tools:** Databricks, Apache Spark (DataFrame, Spark SQL), Apache Hadoop (HDFS), Snowflake, Apache Airflow (Basic), Azure Data Factory, Unity Catalog
- **Cloud Platforms:** Microsoft Azure (ADF, Databricks, Synapse, Blob Storage), AWS Basic (S3), GCP Basic (Big Query)
- **Programming & Scripting:** Python (Pandas, NumPy, PySpark), SQL (Advanced), T-SQL
- **Data Modeling & Warehousing:** Dimensional Modeling, Star Schema, Snowflake Schema, Azure Synapse, Data Warehousing
- **Data Pipelines & Integration:** ETL/ELT workflows using ADF, PySpark, SQL, Real-time and batch processing, Data Engineering, ETL orchestration, Advanced SQL performance tuning
- **CI/CD & Version Control:** Git, GitHub Actions (Basic), Azure DevOps (Basic), SDLC Understanding
- **Databases:** MySQL, SQL Server, Azure SQL Database, NoSQL (Familiar)
- **Data Visualization:** Power BI (Basic), Looker (Familiar), MS Excel (Advanced)
- **Collaboration:** Cross-functional Communication, Agile (Scrum)

CERTIFICATIONS

- **Big Data Analytics:** Florida Atlantic University
05/2025
- **MS Excel - Advanced:** Cengage
04/2024