

Nitish Songara *Data Engineer*

✉ nitishsongara75@gmail.com

☎ +91 98277 32220

📍 India

🌐 /nitishsongara1

🗣 /sonu7089

PROFILE SUMMARY

Data Engineer with **5+ years** of experience building and maintaining end to end **ETL/ELT** pipelines and analytics ready datasets across cloud platforms; I work with **Python, SQL, PySpark, Databricks, and Airflow** to ingest data from **files, databases, and external APIs**, model it for reporting, and ensure reliability through monitoring and data quality checks as well; I leverage **AI tools** in my workflow, have a keen interest in **AI**, and adapt quickly as technology evolves;

PROFESSIONAL EXPERIENCE

Associate Data Engineer, [Ascendium](#)

06/2025 – Present | Remote

- Built and automated end to end ETL pipelines processing **500M+ ad tech events/month** from Amazon S3 into MongoDB to power analytics dashboards used by **50+ stakeholders**
- Migrated batch processing workflows from **Pandas** to **PySpark**, achieving **~5x faster** processing on **200GB+** daily datasets while improving stability at scale
- Designed a reusable, **multi client data ingestion system** to fetch similar datasets from **multiple external APIs**, enabling faster onboarding of new clients with shared pipelines
- Orchestrated pipelines using **Apache Airflow** with clear dependencies, retries, and backfills to ensure reliable delivery of **ad and auction datasets** across environments
- Modeled raw event and ad auction data into analytic ready datasets and derived KPIs, enabling near real-time campaign optimization
- Partnered with business and strategy teams to translate goals into data driven ad recommendations and reporting, contributing to **12%–15% increase** in client ROI

Data Engineer, [Deloitte](#)

08/2021 – 05/2025 | Bengaluru, India

- Designed and maintained **Azure Data Factory** and **Azure Databricks** pipelines to ingest and transform data from **10+ sources**, processing **20M+ records** weekly (**~1TB/month**)
- Built reusable **PySpark** and **SQL** components (standardized transformations, parameterized logic), reducing pipeline development time by **30%** and improving maintainability
- Developed curated **SQL models** and **reporting views** for dashboards used by **200+ users**; improved query performance by **40%**
- Tuned **SQL Server** reporting workloads (query optimization, indexing, statistics maintenance), reducing report runtimes by **20–25%** and improving delivery for leadership metrics
- Improved data pipeline reliability to **99.9% uptime** using automated **data quality checks**, proactive monitoring/alerts, and **schema evolution** handling

SKILLS

Python | SQL | Java | PySpark | Databricks | MongoDB | ETL/ELT | Git Workflows | AI Tools | Docker | Data Modelling | Data Structures & Algorithms | AWS, Azure, GCP | BigQuery | Snowflake | PowerBi | Airflow | External Database and APIs | Problem Solving | Data Quality | Pandas | Query Optimization & Performance Tuning

EDUCATION

Master of Business Administration - Finance,

[Lakshmi Narain College of Technology](#) 

Completed alongside full time role at Deloitte

10/2021 – 08/2023 | Bhopal, India

Bachelor of Engineering - Information Technology,

[Institute of Engineering & Technology \(IET DAVV\)](#) 

07/2017 – 07/2021 | Indore, India