SHWETA RANI

📞 +91-7668681196 | ✉ shwetanimesh700@gmail.com | 🔗 LinkedIn | ⊙ GitHub | ▶ YouTube

Professional Summary

Data Engineer with **2+ years** of experience building scalable ETL/ELT pipelines, real-time streaming solutions, and cloud-based data architectures. Expertise in **Azure, Python, SQL, Snowflake, Databricks** and **Apache NiFi**. Published researcher (Springer Singapore) and technical content creator building the Data Engineering community.

Technical Skills

Data Engineering: ETL/ELT Pipelines, Data Warehousing, Data Modeling, CDC, Real-time Streaming, SnowPipe
Cloud & Big Data: AWS (S3, Lambda, EC2), Azure (Data Factory, Functions), Snowflake, Databricks
Processing: Apache Spark, Apache NiFi, Apache Airflow
Programming: Python, SQL, PySpark
Databases: Snowflake, MySQL, Delta Lake
Tools & AI: Docker, Git, Linux, Jupyter, Jira, Generative AI, LLMs

Professional Experience

WNS <div align="right">**Jul 2024 – Present**</div>
Data Engineer <div align="right">*Gurugram, India*</div>
- Design and implement **end-to-end ETL/ELT pipelines** using Python, SQL, and AWS to process large-scale data volumes
- Build **real-time data streaming** architectures using Apache NiFi and Snowflake with significantly reduced latency
- Develop automated **data quality frameworks** with validation checks ensuring high data accuracy across pipelines
- Optimize SQL queries and data workflows resulting in enhanced performance and cost efficiency
- Collaborate with cross-functional teams to translate business requirements into scalable data solutions

Airtrik <div align="right">**Dec 2023 – Jan 2025**</div>
Data Engineer <div align="right">*Noida, India*</div>
- Built scalable data pipelines using Python, Apache NiFi, and **AWS (EC2, S3, Lambda)** for enterprise data integration
- Implemented **CDC using SnowPipe** and Snowflake streams/tasks for near real-time incremental data processing
- Designed data models and optimized Snowflake warehouse configurations for improved query performance
- Containerized data applications using **Docker**, significantly reducing deployment time across environments

Astrea IT Services <div align="right">**Jul 2020 – Feb 2022**</div>
Software Developer <div align="right">*Noida, India*</div>
- Developed **Salesforce solutions** including custom triggers, buttons, and Visualforce pages to enhance business processes
- Designed and implemented **Apex code** and **SOQL queries** for complex business logic and data manipulations
- Created **custom workflows** and **process automations** to streamline organizational operations
- Participated in code reviews and requirements gathering sessions ensuring high-quality deliverables

Projects

End-to-End GenAI Data Agent | *Databricks, Llama 3.3, Streamlit, Spark SQL, Python* | GitHub
- Built a **powerful GenAI Data Agent** leveraging Llama 3.3 model on Databricks Data Intelligence Platform
- Designed **Medallion Architecture** (Bronze/Silver/Gold) for robust data processing and transformation
- Implemented **custom AI Agent** hosted in Streamlit App that translates natural language questions into optimized Spark SQL queries
- Enabled business users to query massive e-commerce datasets without requiring SQL knowledge

Real-time Data Streaming Pipeline | *Apache NiFi, AWS, Snowflake, SnowPipe, Docker* | GitHub | Demo
- Architected end-to-end streaming pipeline: **EC2 → Apache NiFi → S3 → SnowPipe → Snowflake**
- Implemented **Change Data Capture (CDC)** using Snowflake streams and tasks for incremental data processing
- Containerized entire solution using Docker Compose for consistent deployment and scalability

Spotify ETL Pipeline | *AWS Lambda, S3, CloudWatch, Snowflake, Python* | GitHub | Demo
- Built **serverless ETL pipeline** using AWS Lambda to extract, transform, and load Spotify API data into Snowflake
- Automated pipeline execution with CloudWatch triggers; implemented data validation and error handling
- Designed **star schema data model** in Snowflake for efficient analytical queries

Automated Daily Content Ingestion Pipeline | *Python, Docker, GitHub, Azure ADF, Azure App Service* | GitHub
- Architected end-to-end **automated content ingestion pipeline** using **Azure Data Factory** to orchestrate daily web scraping jobs with **Container Registry** integration and automated deployment via GitHub
- Built **microservices architecture** with **Scraper** (Python web scraping) and **Viewer App** (Azure App Service) displaying articles by date; implemented CI/CD workflows enabling automated image pulls from registry

Education

Gautam Buddha University <div align="right">**Aug 2015 – Dec 2020**</div>
*Integrated B.Tech (Computer Science) + M.Tech (AI & Robotics) – **CGPA: 8.22 / 8.36*** <div align="right">*Greater Noida, India*</div>

Additional

- **Research Publication**: Data Imputation in WSN using Deep Learning – Springer Singapore (ICDAM 2021)
- **Community**: Building Data Engineering community on YouTube