AISHWARYA RAO KALLEPU

Aishurao2021@gmail.com| (656) 565-9217 | Tampa, FL, USA

Summary

Results-driven **Junior Data Engineer** with hands-on experience in building and maintaining data pipelines, managing cloud-based storage, and curating data elements to support analytics. Proficient in **Python, SQL, Snowflake, DB2**, and **ETL development** using DataStage and DBT. Skilled in data lineage documentation, relational/NoSQL databases, and cloud data platforms including **AWS**. Adept at translating business requirements into technical solutions and ensuring data quality, governance, and compliance. Passionate about continuous learning and delivering scalable, clean data systems.

Technical Skills

- Programming & Scripting: Python, SQL, Bash, Java (basic)
- ETL Tools: DataStage, DBT, Azure Data Factory (basic)
- Cloud Platforms: Snowflake, AWS (S3, RDS, EC2), Azure (basic)
- Big Data Technologies: Hadoop (basic), Kafka (basic), Spark (basic)
- Databases: DB2, MySQL, PostgreSQL, MongoDB
- Data Modeling & Governance: Data lineage, curation, data mesh architecture
- Version Control & Tools: Git, Jira, VS Code
- Other: Excel (PowerQuery), Power BI, Tableau

Professional Experience

Graduate Research Assistant *University of South Florida, Tampa, FL • Aug 2023 – Present*
- Built and maintained ETL pipelines using Python and SQL to support academic data analysis, reducing manual reporting efforts by 35%.
- Documented data lineage and curated datasets across 3 major research modules using Snowflake and Excel-based metadata repositories.
- Optimized SQL queries to extract insights from structured and semi-structured sources; improved data retrieval efficiency by 40%.

Research Assistant – Bioengineering Department *Moffitt Cancer Center, Tampa, FL • May 2024 – July 2024*
- Engineered an NLP-based staging classification system in Python to analyze pathology reports.
- Validated system accuracy (95%) using regex-driven pipelines across 500+ clinical cases.

Freelance Full Stack Developer
- Designed and deployed a cloud-based cancer staging application using Python, SQL Server, and DB2, improving data accuracy and accessibility.
- Built ETL pipelines using DataStage and integrated with Snowflake; automated ingestion from on-prem systems to AWS S3, saving 15+ manual hours/week.
- Applied validation scripts to ensure data governance standards and integrity before reporting to analysts.

Undergraduate Software Developer Intern – IT Projects *KITS, Warangal, India • Jul 2019 – May 2021*
- Assisted in maintaining and querying relational databases using MySQL and PostgreSQL.
- Contributed to data migration from legacy Excel systems to structured databases, improving accessibility and reducing redundancy.
- Developed data transformation scripts in Python for downstream analytics and reporting in academic departments.
- Participated in data quality checks, indexing improvements, and performance tuning of SQL scripts.

Education

Master's in computer science
University of South Florida   08/2023 - 05/2025 | Tampa, Florida, USA
Bachelor of Engineering in Computer Science
Kakatiya Institute of Technology and Sciences, Warangal   08/ 2019 - 07/2023 | Waranagl, India

Projects

ACH Transaction Monitoring Pipeline
- Built a mini ACH simulation dataset and constructed an ELT pipeline using DBT and Snowflake to support fraud analytics.
- Delivered a dashboard using Power BI to visualize transaction volumes and failure rates.

Cloud Data Lake Integration (AWS + Snowflake)
- Developed a POC to ingest, store, and query large datasets from on-prem PostgreSQL into Snowflake via AWS S3 and Python scripts.
- Enabled 60% faster access to cleaned data across distributed analytics teams.