

# KONDAPALLI SAI MAHATHI

Hyderabad, Telangana • +91 9391040109 • saimahathi99@gmail.com

<https://www.linkedin.com/in/sai-mahathi-kondapalli-b08b4b1a2/>

---

## SUMMARY

Experienced Data Engineer with nearly 4 years of proven success in designing, building, and optimizing scalable ETL/ELT pipelines and data workflows using Python, SQL, PySpark, Matillion, and Azure Data Factory. Skilled in developing data warehouse architectures on Snowflake, Azure, AWS, and handling structured, semi-structured. Strong hands-on knowledge in data quality frameworks, SQL performance tuning, and implementing real-time/batch pipelines. Adept at working in cloud-native environments with tools like Power BI, version control practices.

---

## SKILLS

- Languages: Python, SQL, HTML
- Cloud & Data Platforms: Snowflake, Azure, AWS
- ETL/Tools: Matillion, Alteryx, PySpark, Azure Data Factory (ADF)
- BI/Reporting: Power BI
- Data Modeling: Star/Snowflake schema, normalized/denormalized models
- Database: MS SQL Server

---

## CERTIFICATIONS

- AWS certified Developer Associate
- Matillion certified Associate
- Microsoft certified Azure Data Engineer
- Snowflake Snowpro Certified

---

## WORK EXPERIENCE

### Senior Software Engineer

Teksystems Global Services, Hyderabad, India

Aug 2021 – Present

- Built PySpark-based distributed data pipelines for transforming large datasets with millions of records, applying partitioning and caching strategies to improve performance.
  - Designed and developed high-volume ETL pipelines using Azure Data Factory, handling over 1M+ records daily.
  - Built and deployed batch and real-time pipelines integrating MS SQL Server, Snowflake, and Power BI.
  - Migrated 10+ Alteryx workflows to Matillion, optimizing performance by 25% and reducing licensing costs.
  - Created 50+ database objects including stored procedures and views for efficient metadata handling and automation.
  - Built and migrated Power BI dataflows across 5+ environments, improving analytics availability and reducing load times by 30%.
  - Delivered scalable pipelines that supported data governance, auditing, and custom quality validation checks.
-



- Implemented data modeling practices using star and snowflake schemas, enabling faster BI reporting.
  - Collaborated closely with data analysts and DevOps teams to ensure CI/CD-readiness and pipeline monitoring.
- 

## EDUCATION

### **Bachelor of Technology – Computer Science Engineering**

Sagi Rama Krishnam Raju Engineering College, Bhimavaram, Andhra Pradesh

**Jun 2017 – Jun 2021 | CGPA: 8.73**

---

## PROJECT HIGHLIGHTS

### **1. ETL Migration and Snowflake Optimization**

- Migrated ETL jobs from Alteryx to Matillion, using Snowflake as both source and target.
- Tuned Snowflake queries and implemented clustering strategies to reduce cloud resource usage.
- Improved operational efficiency by automating workflows and reducing manual intervention.

### **2. Real-Time Data Integration using ADF and On-Premises SQL Server**

- Developed ADF-driven pipelines that executed parameterized stored procedures from on-prem MS SQL Server.
- Ingested real-time incremental data using watermark-based strategies from on-prem to cloud environments.
- Integrated data into Power BI reports, improving the speed and accuracy of business dashboards.
- Reduced overall data processing time and improved traceability across environments.

### **3. PySpark Learning and Transformation**

- Gained hands-on experience with PySpark while supporting transformation of structured and semi-structured data.
  - Independently explored data partitioning and caching to improve performance on sample workloads.
  - Applied custom logic to prepare data for downstream processing in a limited production use case.
-