Piyush Badule

Data Engineer

✉ piyush.badule30@gmail.com   📞 9322479517   📍 Pune, India   🔗 piyushbadule.com

○ Github   in LinkedIn

PROFESSIONAL SUMMARY

Data Engineer with more than 6 years of experience, specializing in designing and optimizing ETL pipelines using Azure, AWS, and Databricks. Proficient in Python and SQL for data processing, querying, and manipulation, with expertise in tools like Pandas, Numpy, and PySpark. Skilled in managing data workflows with Apache Airflow and ensuring efficient large-scale data processing. AWS Certified Data Engineer, AWS Certified Developer, and Azure Data Engineer Certified, with strong experience in cloud deployments on both Azure and AWS. Proficient in version control with Git and containerization using Docker, delivering high-performance, scalable data solutions.

PROFESSIONAL EXPERIENCE

T-Systems ICT India Pvt. Ltd.   06/2022 – Present
Software Engineer   Pune, India

Key Responsibilities:

- Engineered scalable ETL pipelines using PySpark on Databricks and AWS Glue for structured and semi-structured data processing.
- Maintained and optimized Databricks and Glue notebooks for data ingestion, cleansing, transformation, and orchestration.
- Built RESTful APIs with FastAPI to support cloud-based translation of documents, images, text, and websites.
- Integrated Azure and AWS services for secure data storage, access control, and enterprise-scale dataset management.
- Authored modular Python utilities and crafted PySpark UDFs to enable reusable transformations and reduce technical debt.
- Applied data validation and quality frameworks using pytest, AWS CloudWatch, and Step Functions for robust monitoring.
- Designed interactive dashboards and semantic layers to facilitate self-service BI and reporting for business stakeholders.
- Partnered with cross-functional teams (Finance, HR, Sales, Portfolio) to analyse needs and deliver domain-specific data products.
- Contributed to CI/CD workflows using Azure DevOps and Docker, ensuring streamlined and automated releases.
- Delivered production support, ensured data integrity, improved performance, and mentored junior team members.

MindNerves Technology Services Pvt. Ltd.   12/2019 – 06/2022
Software Engineer   Pune, India

Key Responsibilities:

- Designed batch ETL pipelines using AWS Glue and PySpark to transform structured data across customer support, invoicing, logistics, and procurement domains.
- Implemented a multi-zone Lakehouse architecture (Raw, Trusted, Shared) on Amazon S3 with Parquet and Iceberg formats, enabling schema evolution and time travel.
- Enabled ad-hoc analytics through Amazon Athena, delivering curated datasets for BI and downstream systems.
- Created interactive dashboards using Qlik to provide insights into vehicle lifecycle and after-sales operations.
- Orchestrated automation workflows with AWS Step Functions, Lambda, and monitored processes via CloudWatch.
- Applied Lake Formation for fine-grained access control (row/column-level) and role-based authorization.
- Promoted Data Mesh principles by collaborating with domain teams to establish decentralized ownership and governance.
- Enforced schema validation and version control mechanisms to ensure data consistency and pipeline auditability.

SKILLS

Programming Languages
- Python
- SQL

Version Control
- Git
- GitHub

AWS
- IAM
- CloudWatch
- AWS Monitoring Services
- Step Functions

Cloud Storage
- Amazon S3
- Azure Blob Storage
- ADLS

CI/CD Pipelines
- GitHub Actions
- Azure DevOps Pipelines

Containers
- Docker

Backend Frameworks
- FastAPI
- Django
- Flask

ETL Tools
- AWS Glue
- Azure Data Factory (ADF)

NoSQL
Cosmos DB

Azure
- Azure Active Directory
- Event Grid
- Log Analytics
- KQL

Data Processing
- Pandas
- NumPy
- PySpark

Frontend Development
- HTML
- CSS
- JavaScript

Architecture
- Delta Lake
- Data Mesh
- Lakehouse Architecture

Data Warehouses
- Amazon Redshift
- PostgreSQL

Workflow Engines
Apache Airflow

Big Data & Lakehouse
- Databricks
- Apache Spark

EDUCATION

Bachelor of Engineering - Electronics & Telecommunication   2016 – 2019
G. H. Raisoni College of Engineering   Nagpur, India
CGPA: 8.73/10

Diploma in Electronics & Telecommunication   2013 – 2016
Government Polytechnic, Nagpur   Nagpur, India
Percentage: 72.31%

Grade - X   2012 – 2013
K. John Public School   Nagpur, India
CGPA: 8.8/10

AWARDS

Certification Of Appreciation   01/09/2024
T-Systems ICT India Pvt. Ltd.

Certification Of Appreciation   01/05/2024
T-Systems ICT India Pvt. Ltd

Pat on the Back   01/02/2024
T-Systems ICT India Pvt. Ltd

Stay Curious & Grow   01/04/2023
T-Systems ICT India Pvt. Ltd

Pat on the Back   01/12/2022
T-Systems ICT India Pvt. Ltd

CERTIFICATES

Azure Certified Data Engineer | AWS Certified Data Engineer | AWS Certified Developer

LANGUAGES

- English
- Hindi
- Marathi