

# AMBICA DASARI

+1-408-802-2340   [ambdasa@iu.edu](mailto:ambdasa@iu.edu)   <https://linkedin.com/in/ambica-dasari>  
<https://github.com/ambdasa>   <https://portfolio-ambicas-projects.vercel.app/>

## PROFILE

A Data Engineer with 3 years of experience in designing scalable data pipelines and cloud architectures, specializing in AWS, Apache Kafka, and Python to optimize system performance. Passionate about leveraging cloud technologies and real-time streaming to build high-performance data solutions that drive efficiency and innovation.

## EDUCATION

**Indiana University, Bloomington**

**Aug 2023 - May 2025**

*Master of Science, Data Science*

**National Institute of Technology, Puducherry**

**Jul 2017 - May 2021**

*Bachelor of Technology, Computer Science*

## PROFESSIONAL EXPERIENCE

**Indiana University | AWS Cloud Developer**

**Jan 2025 - Present**

- **Cloud Infrastructure:** Designed and deployed a real-time data processing pipeline on AWS using Apache Kafka, efficiently handling 10,000+ seismic data points per second for ambient noise tomography under the FADS program.
- **AWS Solutions:** Engineered an auto-scaling, fault-tolerant cloud architecture, optimizing computational efficiency and ensuring high availability for large-scale geological data processing and analysis.
- **Performance Optimization:** Implemented AWS Lambda and S3-based event-driven workflows, reducing data ingestion latency by 40% and improving the scalability of geophysical data analysis pipelines.

**Larsen and Toubro Infotech | Data Engineer**

**Jul 2021 - Oct 2023**

- **Scalable Data Processing & Automation:** Designed and deployed a parallel processing ingestion framework handling 1,000+ data sources, reducing execution time by 50% for 10TB+ daily data while automating workflows with Python and Apache Airflow.
- **Security & Fraud Detection:** Implemented GDPR-compliant security protocols and ML-based anomaly detection, reducing fraud by 30% across 500+ transactions and securing 1M+ records.
- **Cloud & Performance Optimization:** Enhanced cloud infrastructure reliability and scalability, improving processing efficiency by 40%, ensuring seamless performance under high data loads.

**Lancius IT Solutions | Data Analyst Intern**

**Jun 2019 - Jul 2019**

- **Interactive Dashboards & Data Quality** - Designed and implemented Power BI dashboards with standardized data validation, improving business operations efficiency by 35% and reducing data inconsistencies by 40%.
- **Cross-Functional Leadership** - Collaborated with stakeholders to translate data insights into business objectives, improving project delivery timelines by 20% through effective communication and requirement gathering while increasing actionable insights.

## SKILLS

- **Programming Languages:** Python, C++, Java, R, C
- **Development and Data Tools:** Apache Kafka, Apache Spark, Git, Docker, Real-Time Data Processing, ETL & Data Pipelines, Airflow
- **Cloud & Tools:** AWS (ECS, Lambda, S3, Glue, Redshift CloudWatch, Kubernetes), GCP
- **Frameworks & Libraries:** Pandas, NumPy, PySpark, Matplotlib, Express, Flask, Power BI, Tableau
- **Databases:** PostgreSQL, MySQL, SQL, Supabase, Snowflake, MongoDB, DynamoDB
- **Others:** Data Warehousing & Governance, Big Data Processing, Machine Learning & Analytics, System Design, Agile & Scrum

## PROJECTS

**Stock Market Real-Time Data Analysis**

**Aug 2024 - Nov 2024**

- Architected and deployed a fault-tolerant real-time stock analysis pipeline using Apache Kafka and AWS MSK, implementing dead-letter queues and retry mechanisms for zero data loss.
- Integrated comprehensive monitoring using CloudWatch and Grafana dashboards, enabling real-time system health tracking and achieving 99.9% uptime

**YouTube Data Pipeline Optimization**

**Apr 2024 - Jul 2024**

- Designed and implemented a serverless ETL pipeline using AWS Step Functions and Lambda functions, reducing processing costs by 60% through optimized S3 storage patterns and Athena partition pruning.
- Integrated automated data quality checks using Great Expectations framework, achieving 99% anomaly detection rate in data processing workflows.

**Power BI ER Optimization Dashboard**

**Dec 2024 - Jan 2025**

- Created a Power BI dashboard to optimize patient flow, reduce wait times, and improve resource allocation.
- Delivered actionable insights to reduce wait times by 20% and improve satisfaction by 15%, enhancing ER efficiency.

## CERTIFICATES

- **AWS Certified Data Engineer- Associate:** AWS Certified Data Engineer- Associate