Antony J

senior data engineer, 7+ years experience
bangalore, India
https://augmntx.com/profile/senior-data-engineer-antony-j/axo30668

AugmntX
+91 9820045154 | hr@augmntx.com

A highly accomplished and results-driven Data Engineer with experience in designing, implementing, and optimizing highly scalable and secure data pipelines across both AWS and Azure environments. Specialize in Big Data technologies (Hadoop Ecosystem, Spark, PySpark, Databricks) and Cloud ETL services (AWS Glue, Step Functions, Lambda), consistently driving efficiency and cost savings. Key achievements include reducing data processing time by 30% through Spark optimizations in AWS Glue, leading to a $50K annual cost reduction, and automating workflows to reduce manual intervention by 80%. Expertise spans the entire data lifecycle, focusing on robust data governance, validation checks, and security protocols to ensure 99.9% uptime and reliable data quality for critical business reporting and analytics. Excel at leading cross-functional teams to deliver tailored data solutions and drive significant improvements in data accuracy and reporting efficiency.

INDUSTRIES: Medical, business consulting

TECHNICAL SKILLS

• Pyspark	6 years & 1 months
• MySQL	6 years & 1 months
• MongoDB	6 years & 1 months
• AWS	6 years & 1 months
• Databricks	6 years & 1 months

PROJECT DETAILS

• Patient 360 Healthcare — August-2024 - September-2025

Project Description: Led the design and implementation of a comprehensive healthcare data platform consolidating patient information from multiple sources, architecting a robust 5-layer data pipeline (Landing Raw Trusted Enriched Business using AWS with real-time data consumption capabilities for performance monitoring.

Responsibilities: Achieved 88 percent performance improvement through advanced Spark optimization techniques, reducing critical ETL job execution time from 50 minutes to 6 minutes via broadcast joins strategic caching and intelligent resource allocation techniques applicable to realtime ad bidding optimization. Orchestrated complex data integration from MongoDBbacked e-commerce platform and healthcare systems implementing comprehensive data governance with multi-tier validation, quality monitoring and automated error handling frameworks for ensuring data accuracy similar to CTR validation processes. Enabled multi-channel data consumption through Tableau, react dashboards, AWS Quick Sight, and Amazon Athena for real-time performance monitoring, delivering comprehensive insights for operational efficiency and strategic decision-making

Technologies: h Tableau, react dashboards, AWS Quick Sight, and Amazon Athena

Industry: Medical

• Enterprise June-2022 - August-2024

Project Description: Designed enterprise data engineering solution for workforce and performance analytics processing colleague headcount

Responsibilities: Designed enterprise data engineering solution for workforce and performance analytics processing colleague headcount absence data and performance metrics across 3,000 retail stores globally handling 500K records daily using Apache Spark and Azure services with real-time monitoring capabilities Developed data processing applications with secure data transformation pipelines implementing UUID mapping and encryption for sensitive performance data, ensuring data security and compliance with enterprise standards similar to ad tech privacy requirements Optimized performance and implemented partitioning strategies reducing data processing time by 30% for real-time analytics, established monitoring with Splunk and Azure Monitor ensuring 99.9 percent pipeline reliability and 20 percent improvement in data accuracy for performance metrics validation

Technologies: Spark, Python, Hadoop, Tableau, Azure Stacks, Databricks, PySpark, Spark SQL, Real-time Analytics, Marketing Data Processing, Jira, Splunk

Industry: business consulting

EMPLOYMENT HISTORY

• senior data engineer September-2025 - September-2025

Description: ❖ Designed and implemented highly scalable real-time data pipelines using AWS Glue ETL, Step Functions, Lambda, and S3 for processing ad tech transaction data and performance metrics, ensuring seamless data movement across layers while maintaining 99.9% uptime. ❖ Engineered Spark optimizations within AWS Glue for real-time event processing that reduced CTR calculation processing time by 30% and limited worker usage, leading to improved efficiency and $50K annual cost savings for performance analytics workloads. ❖ Led cross-functional teams including data scientists and marketing analysts to deliver robust data solutions tailored to various business use cases, enhancing collaboration and project outcomes with focus on performance marketing optimization. ❖ Ensured data quality by implementing comprehensive validation checks and security protocols for real-time bidding data at each stage of the pipeline, guaranteeing accurate CTR, CPM, and revenue tracking in the business layer. ❖ Automated workflows with Step Functions and Lambda functions for ad performance monitoring, reducing manual intervention by 80% and enhancing reliability through robust error handling and real-time alerting systems.

EDUCATION

• Bachelors degree in B.Tech from Mahatma Gandhi University 2013 - 2017

LANGUAGES

• English - advanced

END OF DOCUMENT
This PDF is automatically generated by AugmntX on 20 Nov 2026
Profile URL:https://augmntx.com/profile/senior-data-engineer-antony-j/axo30668