

GOUTHAM SENAPATHI

Data Engineer & Analyst | AWS, Azure, Snowflake | Big Data & Cloud Solutions Expert

+1 (214) 55.....90 • goutham.senapathi...@gmail.com • Dallas, TX • <https://www.linkedin.com/in/goutham-senapathi>

PROFESSIONAL SUMMARY

Results-driven Data Engineer & Data Analyst with 4+ years of experience in designing, developing, and optimizing big data solutions, ETL pipelines, and cloud-based architectures. Proficient in AWS, Azure, Snowflake, and Apache Spark, leveraging scalable solutions to extract actionable insights. Expertise in data visualization, predictive analytics, and automation, reducing manual tasks by 30% and improving decision-making efficiency. Adept at working in Agile environments, collaborating with cross-functional teams to deliver secure, high-performing data solutions in industries like Finance, Healthcare, and E-commerce. Passionate about leveraging data-driven strategies to enhance business growth and operational efficiency.

Significant Achievements:

- Reduced data processing time by 40% through optimized Apache Spark jobs and AWS Glue pipelines.
- Enhanced revenue forecasting accuracy by 25% by implementing machine learning models on AWS SageMaker.

SKILLS

Programming Languages: Python, SQL, Java, Scala, C

Big Data Technologies: Apache Spark, Hadoop, Kafka, Hive, HBase, Cassandra

Cloud Platforms: AWS (S3, Redshift, EMR, Lambda, EC2, CloudWatch), Azure (Synapse, Data Factory)

ETL & Data Integration: AWS Glue, Apache NiFi, Talend, Informatica, Apache Airflow

Data Warehousing: Snowflake, Redshift, Azure Synapse, Google BigQuery

Visualization Tools: Power BI, Tableau, AWS QuickSight

DevOps & CI/CD: Jenkins, GitLab CI, Terraform, Docker, Kubernetes

Database Management: SQL Server, MySQL, PostgreSQL, MongoDB, DynamoDB

PROFESSIONAL EXPERIENCE

Big Data Engineer

Goldman Sachs, Dallas, TX | March 2023 – Oct 2024

Project: Scalable E-commerce Data Platform for Real-Time Business Insights

- Designed and optimized ETL pipelines using AWS Glue, reducing manual data transformation efforts by 30%.
- Developed real-time data ingestion pipeline using Apache Kafka & AWS Kinesis, decreasing data latency to under 2 minutes.
- Built a Lakehouse Architecture integrating Amazon S3, Redshift Spectrum, and Snowflake, enabling real-time analytics.
- Automated data validation and quality assurance with Great Expectations, ensuring 99.9% data accuracy.
- Improved cost efficiency by 35% by implementing S3 lifecycle policies and Intelligent-Tiering for data storage.
- Enhanced data security with AWS IAM, KMS encryption, and GDPR-compliant governance policies.

Data Engineer

Cotiviti Private Limited, Hyderabad, India | Oct 2019 - July 2022

Project: Healthcare Data Integration & Automation

- Developed end-to-end ETL pipelines using AWS Glue and Snowflake, reducing processing time by 40%.
- Designed automated workflows for data ingestion from AWS RDS and third-party healthcare sources.
- Improved query performance in Snowflake by 30% through optimized clustering keys and caching strategies.
- Integrated AWS Lambda functions to orchestrate real-time data transformation, reducing manual intervention.
- Built interactive dashboards using Power BI and AWS QuickSight, enabling real-time business intelligence reporting.
- Implemented data quality frameworks, ensuring compliance with HIPAA and healthcare industry standards.

EDUCATION

University of North Texas, Denton, TX

Master of Science in Data Science (Aug 2022 - May 2024)

ADDITIONAL INFORMATION

Certifications: AWS Certified Data Analytics – Specialty, Microsoft Certified: Azure Data Engineer Associate

Awards: Awarded “Excellence in Project Delivery” for successful implementation of real-time analytics platform