Sathwik Hegde

Denver, CO | satwikhegde14@gmail.com |(617)-959-7167 | [linkedin](#) | [GitHub](#) |[Tableau](#)

Data Engineer

Accomplished Data Engineer with extensive experience in architecting, automating, and optimizing data analytics and reporting solutions to enhance operational efficiency and performance. Expert in leveraging advanced data engineering tools to support critical business decisions and improve data accuracy. Proven track record in delivering scalable workflows and actionable insights that drive superior business outcomes

KEY SKILLS

Programming Languages:	Python, SQL, R, Java, Html, CSS, JavaScript, ReactJS, Scala, C, C++
Databases:	MySQL, SQL Server, Oracle, Snowflake PostgreSQL, Fivetran, NoSql, Microsoft Azure Studio, SSIS
BI Tools:	Tableau, Power BI,Dbt, SSRS, Talend, Teradata, Alteryx Cloud, Knime Analytics
Cloud Platform:	AWS, Azure, GCP (BigQuery, Data Studio, Dataflow)
Libraries/Frameworks:	Matplotlib, Seaborn, Keras, AutoML, NLP, PyTorch, Elasticsearch
Operating Systems:	Windows, Linux, Ubuntu, macOS
Other Technologies:	Looker, Git, Airflow, Spotfire, Hive, Spark, Hadoop, Kafka, Map-reduce, Kubernetes, Docker

PROFESSIONAL EXPERIENCE

CloudData Technology LLC, Denver, CO | Data Engineer   Sep 2024 - present
- Designed and implemented ETL pipelines for healthcare claims data ingestion, transformation, and analysis using AWS Kinesis Firehose, AWS Glue, and AWS Lambda, improving claims processing time by 30%
- Designed dimensional and normalized data models in Amazon Redshift to optimize storage, improving query times by 20%
- Built and maintained workflows in AWS Step Functions to orchestrate data workflows, increasing pipeline reliability by 25%
- Collaborated with cross-functional teams to ensure compliance with HIPAA standards, enhancing data privacy and security across AWS data pipeline
- Conducted data validation and quality assurance processes, identifying and resolving issues in data with a 98% accuracy rate

Capgemini U.S. LLC, Denver, CO | Data Engineer   Sep 2023 - May 2024
- Developed and optimized Tableau dashboards for a leading American telecommunications and media company, equipping senior management with actionable insights on CBRS data usage, data offload relative to Verizon, HMNO eSIM provisioning, SIM consumption rates, SIM inventory, fallouts, key KPIs, and ticket metrics driving a 25% boost in operational efficiency
- Developed and maintained ETL pipelines using AWS Glue, Alteryx, and Knime Analytics to streamline data ingestion, transformation, and loading, improving data integration efficiency by 30%
- Automated data analytics and reporting processes using CI/CD pipelines and Python scripts, reducing manual interventions by 40% and improving data accuracy in mobile activations and mediation call processes

Capgemini India Pvt. Ltd., Mumbai, India | Data Engineer   Feb 2017 - Jul 2021
- Engineered ETL pipelines for an American wholesale corporation, integrating AWS RDS and on-premise data sources with AWS Glue for ingestion and transformation, and storing in Redshift and S3, increasing data accessibility and analytical efficiency by 25%
- Enabled faster business decisions for an American building materials manufacturer by optimizing export sales, shipments, freight management, carrier performance, and data security, resulting in project benefits valued between $10k-$15k each
- Reinforced data warehouse design and data modeling within Snowflake, optimizing pipeline performance by tuning cluster configurations and implementing parallel processing techniques improving data extraction speed by 30%
- Performed data analysis, cleansing, and testing on BI reports and databases, diminishing data inconsistencies and errors by 60% through bug fixes and optimizations Capgemini USA and Brazil's finance, HR, talent management, and payroll systems
- Coordinated with cross-functional teams to gather requirements, and validate data models on Oracle, MS SQL Server, and SAP HANA databases, supporting 1,000 concurrent users and 1 million records, with a 75% uptime SLA

PROJECT

User Management System (JavaScript, AWS)   2024
- Created and executed a collection of JavaScript APIs to support a versatile user management system utilizing an RDS backend
- Designed and orchestrated complete AWS infrastructure, encompassing VPC, subnets, EC2 instances, S3 buckets, RDS databases, Route 53 configurations, Auto Scaling policies, IAM roles and policies, Elastic Load Balancers, and CloudWatch monitoring
- Established automated CI/CD pipelines for code build and deployment and Integrated GitHub actions and Packer for AMI creation and EC2 instance launching via CloudFormation templates
- Reviewed CloudWatch CPU and memory metrics to establish ALB scale-in/out policies, and elevated system scalability to accommodate up to 10,000 concurrent requests
- Applied an email verification workflow deploying SES, SNS topics and Lambda Functions reducing spam accounts by 95%

EDUCATION | CERTIFICATION | TRAINING

Master of Science, Major In Information Systems | Northeastern University, Boston, MA   2023
Certifications: Databricks Lakehouse Fundamentals, dbt Fundamentals, Azure Data Fundamentals, AI Fundamentals