

# VIGNESH VENKATAKUMAR Data Engineer

Buffalo, NY, USA | +1(716)808-1681 | [vigneshv007@gmail.com](mailto:vigneshv007@gmail.com) | [LinkedIn](#) | [Portfolio](#) | Willing to Relocate

## SKILLS AND CERTIFICATIONS

- **Cloud Services and DevOps:** EC2, S3, Lambda, RDS, IAM, Glue, ECS, CloudFormation, EBS, DMS, Docker, GitHub, GitHub Actions, Terraform, AWS Codepipeline.
- **Data Engineering and Management:** PostgreSQL, MySQL, Oracle, NoSQL, RedShift, ETL pipelines, Microservices, Step Functions, Apache Airflow, Apache Spark, Databricks, Apache Kafka, DBT, Snowflake, Trino, Apache Iceberg.
- **Programming and Tools:** Python (Boto3, PySpark, Flask, Pandas, Scikit-learn), C, Visual Studio, TrackWise, Crystal Reports, Tableau.
- **Certifications:** AWS Certified Data Engineer Associate, AWS Certified DevOps Engineer Professional, AWS Certified Solution Architect Associate, Cloud Computing Certification from Great Learning, DataExpert.io

## EXPERIENCE

### Community Dreams Foundation | Data Engineer Intern

Sep 2024 – Mar 2025

- Designed and implemented a serverless architecture using AWS API Gateway, AWS Lambda, and DynamoDB for fast, secure storage and retrieval of student quiz data, performance metrics, and assessment results.
- Optimized DynamoDB database design for fast query performance and automated daily snapshots of quiz data into Amazon S3 data lake using AWS Glue and utilized Athena for trend analysis.

### LatentView Analytics | Senior AWS Data Engineer

Sep 2022 – Dec 2022

- Developed a Databricks-based analytics solution processing about 1TB daily vehicle telemetry data and delivered interactive dashboards to visualize traffic congestion, charging infrastructure needs, and regional driving insights.
- Built scalable ETL pipelines and APIs to analyze vehicle data across 100+ high-density customer regions, optimizing strategic decisions for service station placements and resource allocation.
- Optimized AWS infrastructure leveraging serverless architectures, AWS Glue workflows, Step Functions, and Codepipeline, enhancing urban mobility analytics while reducing data processing costs by 20%.

### Deloitte | Cloud Engineer

Aug 2020 – Sep 2022

- **Led a team of 3**, building automated tagging mechanisms for AWS S3 files and integrating ServiceNow via AWS Lambda and Glue, automating incident creation upon job failures.
- Delivered **POCs** for **containerizing** a Flask application on **AWS ECS**, enabling secure **cross-account S3 access**, and piloted **OHDSI Atlas software** on the cloud for healthcare analytics, which led to full-scale adoption.
- Architected **Data pipelines** and implemented transactional **data modelling** for vendor data using **Apache Airflow** with AWS serverless technologies reducing processing time by 25%. Received a **spot award** for flawless deployment.
- **Migrated** on-premises PubMed and Embase systems ETL processes to AWS, leveraging **Glue** and **Airflow**, enhancing data availability and streamlining operations for medical research systems.
- Enhanced **database performance** and availability by migrating on-premises systems to Aurora PostgreSQL using **AWS DMS**. Reduced operational costs by 50% and migrating historical data for vendor projects.
- Launched **CI/CD pipelines** using AWS Codepipeline, **CloudFormation**, and **GitHub Actions**, ensuring repeatability in testing and dependency management. **AWS SAM** builds and triggered deployments enhanced deployment speed.
- Designed Fact and Dimension tables for a pharmaceutical data warehouse on Amazon Redshift, reducing query response times by 30%, and leveraged DBT and Apache Airflow to automate department-specific data marts, accelerating business analytics.
- Built a Databricks batch analytics pipeline following the Medallion architecture for large-scale pharmaceutical sales and inventory data, leveraging Spark for high-volume data transformation, quality management, and governance.
- Conducted a POC comparing Snowflake and Amazon Redshift for pharmaceutical analytics, recommending Snowflake for its cost-effectiveness, scalability, and simplified management.

### Cognizant | Associate

May 2016 – Jul 2020

- Deployed TrackWise infrastructure via AWS CloudFormation utilizing EC2, ELB, RDS, for robust failover strategies.
- Configured **CloudWatch Agent** on EC2 to capture TrackWise logs and created Lambda and **CloudWatch Alarms** for **real-time monitoring** and alerts, reducing response time by 20% and improving system reliability.
- Orchestrated workflows using **AWS Step Functions**, integrating manual approvals and notifications with **SNS** for efficient data distribution ensuring regulatory compliance. Improved process efficiency by reducing manual efforts.

## PERSONAL PROJECTS

- **Car Rentals:** Flask web application leveraging SQL stored procedures and triggers to manage rentals on AWS using CI/CD.
- **Restaurant Analytics:** Machine learning app to predict restaurant ratings and online feasibility deployed on AWS.
- **Insurance Processing:** Streamlined claims workflow with AWS Step Functions Lambda, SNS, DynamoDB and API Gateway.
- **Cost Prediction:** Developed machine learning models for predicting medical insurance costs and generate insights.

## EDUCATION

University at Buffalo, Master of Science in Data Science

May 2024

Anna University, Bachelor of Engineering in Electrical and Electronics

May 2016