

Rohan Reddy Bandi

+1 (813) 893 5627 | brohanreddy24@gmail.com | [Tableau](#) | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

SUMMARY

Experienced Data Engineer with 5+ years working across cloud and on-premise environments to build reliable, scalable data pipelines and warehouse solutions. Skilled Python, SQL, Snowflake, Airflow, and Kafka to support everything from reporting to real-time analytics. Known for collaborating closely with analysts, engineers, and business teams to deliver clean, well-modeled data that drives decision-making.

PROFESSIONAL EXPERIENCE

Data Engineer, CVS Health

January 2024 – Present

- Designed and implemented scalable ETL pipelines using **Databricks**, **Kubeflow**, and **Snowflake**, integrating with **Azure Data Lake Storage (ADLS)**, **SFTP**, and **Azure Kubernetes Service (AKS)** to automate data ingestion from JDA servers.
- Built an end-to-end ML scoring pipeline using **Snowpark** and Snowflake compute, replacing Databricks and **Airflow**, which reduced overall compute costs by 20%.
- Developed robust data migration workflows from **Oracle** to Snowflake using Databricks and **PySpark**, resolving schema mismatches and ensuring data integrity.
- Created reusable **Argo Workflows** and modular ETL components within **Kubeflow**, optimizing resource usage and cutting infrastructure costs by 50%.
- Tuned and optimized **Apache Spark** and PySpark jobs to enhance pipeline performance, reducing runtime by 20% and improving system reliability.
- Delivered over \$10M in business impact, including \$500K in value from enhanced precision tracking enabled by Kubeflow-based ML pipelines.

Data Engineer, Accenture PVT LTD

August 2020 – July 2022

- Engineered scalable ETL/ELT pipelines using Python, SQL, and **AWS Glue** to process **EHR** and clinical datasets containing over 10 million records, loading into **Redshift** for integrated **HIV** registry analytics.
- Structured data models using SQL views, **stored procedures**, and automated workflows with **Airflow**, **Glue**, and **Terraform**, reducing reporting time by 40% and ensuring consistent outputs across studies.
- Performed data transformation and analysis with **Pandas** and **tidyverse**, cutting manual preparation by 30% and improving collaboration across research sites.
- Provisioned and fine-tuned **AWS infrastructure** with Terraform, applying partitioning and query optimization techniques to reduce processing time and cloud costs by 30%.
- Maintained centralized **metadata** systems and data dictionaries, ensuring secure, compliant data assets in alignment with **HIPAA**, **IRB**, and **HL7** protocols.

Associate Data Analyst, Tech Mahindra

May 2019 - July 2020

- Implemented scalable data pipelines using **AWS Glue**, **S3**, and **Redshift** to process over 5 TB of daily data, supporting both operational reporting and analytics.
- Configured 15+ business dashboards in **AWS QuickSight** and **Tableau**, supporting performance monitoring and team decision-making.
- Executed advanced **SQL** queries in Amazon Redshift and used **Python (Pandas)** for data exploration, delivering insights that improved resource planning accuracy by 25%.
- Streamlined ETL processes using **AWS Lambda** (Node.js 14.x) and Python, reducing repetitive tasks and improving delivery timelines.
- Outlined key data assets, workflows, and transformations in internal wikis, accelerating team onboarding by 20% and boosting self-service reporting adoption.

TECHNICAL SKILLS

Programming & Scripting: Python, SQL, Java, JavaScript, Bash, Scala.

Data Engineering & Processing: ETL/ELT Pipeline Design, Apache Airflow, dbt, Apache Kafka, HL7 Integration, Data Integration, Data Modeling, Data Lake Architecture.

Databases & Data Warehousing: PostgreSQL, MySQL, Snowflake, Oracle, Azure Cosmos DB, AWS DynamoDB

Cloud Platforms: AWS (S3, Redshift, Lambda, Kinesis EMR Glue, QuickSight), Azure (Synapse, Data Factory, Key Vault, Databricks).

Data Preparation & Automation Tools: Alteryx, Salesforce Integration, Talend, Informatica.

Machine Learning: Predictive Modeling, Anomaly Detection, A/B Testing, Snowpark ML Integration, TensorFlow.

Visualization & Reporting: Tableau, Power BI, Looker,

Version Control & DevOps: Git, Azure DevOps, CI/CD Pipelines, Jenkins, Kubernetes, Docker, Terraform.

EDUCATION

University of South Florida Tampa, Florida <i>Master of Science, Business Analytics and Information Systems</i>	GPA:3.94
SR Engineering College India <i>Bachelor of Technology, Computer Science</i>	GPA:3.65

ACADEMIC PROJECTS

Real-Time Stock Market Analysis

Key Technologies: Python, Kafka, Boto3, AWS Glue, S3, Athena, real-time processing, pipeline architecture.

- Developed a stock market simulation app in Python to generate continuous CSV-based trading data, streamed in real-time to Apache Kafka hosted on Amazon EC2 using Boto3, and stored in Amazon S3 for downstream access.
- Orchestrated a complete data pipeline automating real-time ingestion of stock data into AWS S3, cataloging with AWS Glue, and querying with Athena; reduced query processing time by 60%.

Airbnb Data Modeling & Insights

Key Technologies: SQL, dimensional modeling, schema design, analytics, customer behavior analysis.

- Designed a relational database in SQL Developer to mirror Airbnb’s business model, capturing entities like listings, reservations, customers, and reviews using a dimensional modeling approach with clearly defined fact and dimension tables. enabling efficient querying and reducing report generation time by 40%.
- Wrote SQL queries to analyze booking frequency, customer behavior, and regional demand patterns, uncovering insights that improved targeting strategies and increased regional marketing efficiency by 25%.

CERTIFICATIONS

<ul style="list-style-type: none">Microsoft Certified: Azure Fundamentals (AZ-900)Google Data Analytics Specialization	<ul style="list-style-type: none">USF Muma college of business Tableau bootcampAWS Data Engineer Certification
---	---

HONORS & AWARDS

- Led a team of 4, successfully secured State Funding (\$300), patent application and participated in competitions (SIH, Tie-Grad, Boot Camp’s) and State-Level recognition (top 60 Telangana Innovation Yatra in January 2020).
- Achieved first place in the internal hackathon for the Smart India Hackathon 2019, held in January 2019.