

Smit Shah

[in linkedin.com/in/smit-shah01](https://www.linkedin.com/in/smit-shah01) | [Github](#) | [Portfolio](#) | [✉ shahsmit661@gmail.com](mailto:shahsmit661@gmail.com) | [📞 \(857\) 395-9925](#)

WORK EXPERIENCE

Data Engineer - Code Logic Infotech, Houston

July 2025 - Present

- Used **AWS** services like **EC2**, **S3**, and **CloudWatch** for improving cloud testing; revamped scripts and alerts, boosting uptime up to 90% to 95%.
- Tasked with streamlining QA, I integrated **Databricks** and **Python** into AWS CodePipeline and CodeBuild, **reducing manual testing** effort by 80%.
- Optimized **AWS infrastructure** using **EC2**, **Athena**, **S3**, and **Lambda**, resulting in a 60% improvement in data processing workflows and a 50% reduction in processing time.
- Analyzed **1M+** NYC Taxi records; built **data models** and **data pipelines** with **AWS S3**, **Mage**, and **Python**, improved analysis throughput by 65% and reduced runtime from **40 to 15 minutes**.
- Utilized **AWS Glue** and **EMR** to process **1+ GB** of raw data daily with optimized workflows.

Data Analyst – Rebecca Everlence Trust, Chicago

August 2024 – July 2025

- Gathered **business requirements** for an internal project management tool, improving the efficiency by 30% and **streamlining workflows** through team collaboration and tailored data insights.
- Mapped **data requirements** and built a data validation process, ensuring data accuracy by 40%.
- Automated a **data pipeline** integrating inputs from **3 departments** and standardized data formatting across sources, reducing **report creation time** by 65% and minimizing manual errors.
- Developed 4 **Tableau dashboards** for the Talent Enablement team, providing **real-time insights** into program expansion by defining **10+** key metrics.

Cloud Data Analyst – AKS Infotech, New Jersey

May 2020 – March 2021

- Utilized **SQL** to analyze over **100K** records, ensuring optimal performance.
- Implemented **data partitioning** and **data modeling** to enhance accessibility, designing data-driven reports and dashboards in **Tableau**, which improved reporting efficiency by 55%.
- Prepared and analyzed **5 reports** based on **supply chain data** for the sales team by tracking **KPIs using Tableau**, leading to a 30% increase in product sales.
- Guided a **team of 3** junior analysts through iterative analytics cycles to deliver revenue insights, streamlining **reporting workflows** and **increasing efficiency** by 70%.

EDUCATION

MS in Information Technology - **University of Massachusetts, Boston**

May 2024

BE in Information Technology - **University of Mumbai, Mumbai**

May 2022

SKILLS

Programming Languages/Web: SQL, Python, C, R, HTML.

Relational Database: MySQL, SQLite, Oracle, Microsoft SQL Server, ETL.

Data Visualization/BI Tools: Tableau, Power BI, Excel, Snowflake, DAX, Looker, BigQuery.

Tools: Amazon Web Services, Databricks, Azure, GCP, dbt, Mage, JIRA, Git.

PROJECTS

Banking Risk Analytics Dashboard | PowerBI | [Github](#)

May 2025 – June 2025

- Built an **end-to-end banking** analytics project using **Python**, **SQL**, and **Power BI**, analyzing **10K+** **client** records and **\$8M+** in financial data.
- Designed dashboards with **12+ KPIs** to assess **\$3M+ in loans**, **\$5M+ in deposits**, and client engagement for data-driven decisions.

Sales and Customer Performance Analysis | [Tableau](#) | [Github](#)

November 2024 – December 2024

- Created an interactive sales dashboard using **Tableau**, visualizing over **9K orders** and nearly **\$2M+ in sales** to provide a clear view of total profit, product performance, and sales trends.
- Developed a comprehensive customer **analytics dashboard** that analyzed a customer base of **750+ unique customers** to identify patterns, geographic distribution, and different KPIs.