

Narasimha Guda

Jersey City, NJ ✉ narasimhaguda99@gmail.com ☎ +1 (972) 654-5411 🌐 [Narasimha Guda | LinkedIn](#)

SUMMARY

Certified Data Engineer with a Master's in Data Science and a Bachelor's in Computer Science, with 5+ years of experience in architecting and deploying efficient Big Data solutions. Proficient in leveraging SPARK, AZURE, AWS, Snowflake and Python to design and manage robust data pipelines, ensuring seamless extraction, transformation, and loading of data. Demonstrated expertise in data warehousing, employing test-driven development methodologies, and tackling intricate challenges to translate complex business requirements into actionable insights. Skilled in SQL, ETL processes, and Statistical Analysis, with a commitment to fostering data-informed decision-making and contributing to organizational success through innovative data strategies.

EXPERIENCE

Data Engineer | Innovers Inc | Remote, USA | January 2023 – Present

- Developed and designed various layer **Datalake** infrastructure for optimal extraction, transformation, and loading of data from various data sources using Microsoft Azure services, Python, and SQL technologies
- Developed data ingestion processes using **Azure Data Factory**, integrating structured and unstructured data from on-premises and cloud sources into Azure storage solutions (**Azure Blob Storage and Azure Data Lake Storage**)
- Designed and implemented scalable and efficient data pipelines on Azure, leveraging services such as Azure Data Factory, **Azure Databricks**, and **Azure Functions** to extract, transform, and load (ETL) data from diverse sources (Salesforce, Personify, Piano, Snipcart)
- Implemented data transformations and data cleansing using **Azure Databricks**, **PySpark**, and SQL, ensuring data quality and consistency for downstream analytics and reporting
- Designed and implemented **Azure SQL Data Warehouse** and **Azure Synapse Analytics** solutions, optimizing data storage, partitioning, and query performance for large-scale analytics workloads
- Implemented data security and compliance measures on Azure, including encryption at rest and in transit, **Azure Active Directory** authentication, and RBAC (Role-Based Access Control) configurations
- Collaborated with cross-functional teams (Marketing, Product, Engineering) to gather requirements and translate them into scalable and performant data engineering solutions, aligning with business needs and objectives
- Conducted performance tuning and optimization of **Azure data solutions**, including query optimization, indexing strategies, and resource management, to improve overall system efficiency and cost-effectiveness
- Implemented monitoring and alerting solutions on Azure, utilizing **Azure Monitor**, **Azure Log Analytics**, and **Azure Functions** to track pipeline performance, detect anomalies, and proactively address issues
- Worked closely with data scientists and analysts to operationalize machine learning models on Azure, integrating them into the data pipeline for real-time or batch predictions
- Worked on API integrations and FTP data extractions from different data formats like JSON, XML, CSV, TXT, etc.
- Worked on data governance to provide operational structure to previously ungoverned data environments
- Build analytical reports using **PowerBI** to utilize the data pipeline, providing actionable insight into key business performance metrics including operational efficiency and customer acquisition and Retention
- Worked in an Agile environment and used the Kanban tool to maintain the user stories and tasks
- Utilized **Azure Key Vault** as a central repository to store and manage confidential information accessed by Azure Data Factory and Databricks notebooks
- Conducted training sessions for junior data analysts, engineers, and non-technical audiences on Big Data and **PowerBI** dashboard design
- Developed automated data pipelines for eBook URL generation project using **AWS Lambda**, **S3**, **Amazon Redshift** to create downloadable eBook URLs for users

Data Engineer | LTI Mindtree Ltd | Chennai, India | Jan 2019 - August 2022

- Transformed healthcare analytics by optimizing data processing and enhancing efficiency using Azure cloud-based data warehousing solutions, such as **Azure SQL Data Warehouse (Synapse Analytics)**, resulting in a 30% reduction in data processing times.
- Implemented BI concepts and cloud-based solutions with hands-on expertise in architecting and managing data warehouses on **Azure Synapse Analytics**.
- Designed and implemented scalable data models on Azure Synapse Analytics, ensuring optimal performance for complex analytical queries.
- Created insightful dashboards in **Power BI** using data sourced from the Data Engineering team's **Azure SQL Data Warehouse**, providing actionable insights to stakeholders.
- Tailored large volumes of healthcare data according to business requirements, incorporating additional views and conducting data modeling on facts and dimensions to support diverse analytical needs.

- Implemented row-level security in **Power BI** to ensure data confidentiality and compliance with regulatory standards, safeguarding sensitive healthcare information.
- Collaborated closely with **stakeholders**, including healthcare providers and regulatory bodies, to understand business requirements and translate them into actionable insights through intuitive dashboards and reports.
- Played a key role in testing and validating dashboards to ensure accuracy and reliability **in data visualization**, conducting thorough QA processes to maintain data integrity.
- Developed custom data processing scripts in **Python** to automate repetitive tasks and increase code efficiency by 15% within healthcare analytics solutions.
- Leveraged advanced analytics techniques, such as predictive modeling and machine learning, to uncover hidden patterns and trends in healthcare data, driving informed decision-making and improving patient outcomes.
- Participated in knowledge sharing sessions and workshops to disseminate best practices in BI development and cloud-based analytics across the organization, fostering a culture of continuous learning and innovation.

EDUCATION

Master's in Data Science | 3.75 | Stevens Institute of Technology | Hoboken, New Jersey | 2023

Related Courses: Statistical methods, Analysis Review, Applied Machine Learning, Deep Learning, Marketing Analytics, Big Data Technologies, Supply Chain Analytics, Advanced Optimization methods, Time Series Analysis.

Bachelor's in Computer Science | JNTUA| Chittoor, India

Related Courses: Data Structure & Files, Advanced Database Management System, Computer Programming I, Computer Programming II, Analysis of Algorithm & Design, Web Engineering, Data Warehouse & Mining.

INVOLVEMENT

Business Intelligence & Analytics Club | Stevens Institute of Technology | Vice President | Jan 2023 - Dec 2023

Graduation Academics And Student Success | Stevens Institute Of Technology | Graduate Student Assistant | Aug 2022 - Jan 2023

SKILLS

- **Programming Languages:** Python, PySpark, Java, C++
- **Key Skills:** Data Modeling, Data Integration, ETL/ELT Processes, ER Modeling, Data Warehousing, Data Pipelines, Data Analytics, Data Modeling, EDA, Forecasting, Predictive/Statistical modeling, Time Series Analysis, Machine Learning, Text & Visual Analytics, A/B Testing
- **Databases Technologies:** MSSQL, T-SQL, MySQL, NoSQL
- **Cloud Technologies :** Microsoft Azure (Azure SQL, Datalake, Data Factories, VM, Synapse, Databricks), AWS (Glue, EMR, EC2, S3, Redshift, Lambda, Glue, Databricks, Athena, Kinesis, Aurora)
- **Tools & other Skills:** Tableau, Power BI, Snowflake, Hadoop, Spark, Anaconda, Jupiter Notebook, UNIX shell Scripting, MS Excel, MS PowerPoint, Project Management, Git, Jira, Kanban, Slack, Agile, UML, System Development Life Cycle(SDLC), Waterfall Model.

CERTIFICATIONS

- **AWS Data Engineer Associate** – 2024
- Tricentis Tosca Fundamentals – Automating web application testing (AS1)
- Tricentis Tosca Fundamentals – Optimizing test automation with centralized test data (AS2)