

# Chavula Shashidhar Reddy

[shashidharreddy0104@gmail.com](mailto:shashidharreddy0104@gmail.com) | (352)368-1270 | Gainesville, FL | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

## SUMMARY

Detail-oriented Data Analyst with over 4 years of expertise in building robust ETL pipelines and scalable cloud-based analytics solutions using AWS Glue, Apache Airflow, and Azure Databricks. Proficient in Python (Pandas, PySpark, Scikit-learn), SQL (PostgreSQL, T-SQL), and visualization tools such as Power BI and Tableau, effectively reducing manual processes and driving strategic insights. Adept at enhancing data governance and real-time analytical capabilities.

## EDUCATION

<b>Florida International University</b> Master of Science in Computer Science	<b>Jan 2023 – Dec 2024</b> Miami, FL, USA
<b>Vignan Institute of Technology &amp; Science</b> Bachelor's in Computer Science and Engineering	<b>Jun 2017 – Jul 2021</b> Hyderabad, India

## PROFESSIONAL EXPERIENCE

<b>McKinsey &amp; Company</b> Jr. Consultant/Data Analyst	<b>APR 2025 – Present</b> FL, USA
<ul style="list-style-type: none"><li>Developed PySpark validation scripts that helped resolve over 200 recurring data quality issues in the initial phase of pipeline deployment, improving early reporting reliability.</li><li>Contributed to the integration of 12+ disparate data sources by implementing schema normalization processes within AWS Glue and Azure Databricks.</li><li>Currently supporting the development of real-time ETL pipelines and streaming ingestion using Kafka, enabling near-instant analytics for supply chain and finance use cases.</li><li>Designing Building Power BI dashboards using SQL over Kafka-fed data warehouses, supporting real-time visibility for six departments.</li><li>Collaborating with cross-functional teams to align data pipeline logic, governance policies, and KPI reporting requirements for enterprise stakeholders.</li></ul>	
<b>Florida International University</b> Graduate Teaching Assistant	<b>AUG 2023 – DEC 2024</b> FL, USA
<ul style="list-style-type: none"><li>Managed 15+ Power BI dashboards, offering real-time visual insights into sales performance, customer behavior, and operational metrics to assist business teams.</li><li>Coordinated data analysis using SQL, extracting insights from datasets with more than 1 million records, supporting business decisions, and streamlining workflows.</li><li>Applied data mining techniques to analyze financial trends, identifying key insights from datasets containing 5,000+ financial transactions related to spending behaviors, risk assessment, and profitability, supporting strategies for financial product development.</li><li>Automated and maintained over two SQL Server backup, ensuring reliable data recovery and continuous availability of critical business data using T-SQL.</li></ul>	
<b>Infosys</b> System Engineer/ Data Analyst	<b>SEP 2021 – JAN 2023</b> India
<ul style="list-style-type: none"><li>Analyzed large-scale client datasets using SQL and Python, resolving 100+ data discrepancies quarterly and enhancing data integrity through automated validation workflows.</li><li>Developed interactive Power BI dashboards by integrating diverse enterprise systems, reducing ad-hoc reporting, and enabling faster decision-making for business units.</li></ul>	

- Streamlined ETL pipelines using Apache Airflow, automating ingestion and transformation tasks, and improving system reliability with proactive monitoring and debugging.
- Built predictive models using R(logistic Regression) to support risk assessment and operational planning, delivering actionable insights across various business scenarios.
- Implemented AWS Redshift for cloud-based data warehousing, scaling data infrastructure to support 15+ TB workloads and optimizing performance through query tuning.
- Audited metadata and access policies with Alation to strengthen data governance practices, standardize documentation, and ensure cross-platform consistency.
- Automated recurring reporting tasks using VBA macros, reducing manual effort and improving on-time delivery of stakeholder reports.
- Created geospatial dashboards using ArcGIS, visualizing regional performance and user behavior to inform expansion and service strategies.
- Supported critical Oracle databases by ensuring uptime, managing backups, and implementing disaster recovery procedures aligned with business continuity goals.

## Dell Technologies

**JUL 2020 - AUG 2021**

Data Analyst

India

- Optimized SQL queries by refining indexing strategies and structuring joins for faster data retrieval and aggregation, reducing report generation time, and boosting database performance for enterprise system analytics.
- Refactored complex SQL scripts to extract, transform, and analyze operational and infrastructure data across Dell's IT environment, improving accuracy and availability of critical metrics.
- Automated recurring reporting workflows in Excel using advanced functions like Pivot Tables, VLOOKUP, and Macros, streamlining internal process audits and inventory assessments.
- Strengthened database security in PostgreSQL by implementing role-based access controls, data encryption, and validation logic, ensuring compliance and safeguarding over 1TB of infrastructure and client data.
- Leveraged Python (Pandas) to cleanse and manipulate large volumes of system performance logs and telemetry data, enabling deeper insights for capacity planning and resource optimization.
- Conducted trend and variance analysis using Excel, generating actionable reports that supported product lifecycle decisions and operational efficiency improvements.

## SKILLS

- **Data Analytics & Business Intelligence:** Power BI, Tableau, ArcGIS, Excel (PivotTables, Power Query, VLOOKUP, Macros)
- **Programming & Scripting:** Python (Pandas, PySpark, Scikit-learn), R, T-SQL, VBA
- **Databases & Data Warehousing:** SQL (PostgreSQL, SQL Server, Oracle), AWS Redshift, Alation
- **ETL & Cloud Platforms:** Apache Airflow, Azure Databricks, PySpark, AWS (Glue, Redshift), Azure (Databricks), Apache Kafka
- **Data Science & Modeling:** Data Mining, Predictive Modeling, Scikit-learn
- **Data Governance & Automation:** Metadata Management, Role-Based Access Control, Python-based Automation, Excel Macros, SQL Jobs, Automated Validation Scripts
- **Tools & Methodologies:** Git, REST APIs, Docker