

[Early/Mid-Career Data Engineering role]

FirstName LastName

Arlington, VA • name@gmail.com • XXX-XXX-XXXX • github.com/name • linkedin.com/in/name

Data Engineer with strong foundation in cloud-based data pipelines, distributed systems, and data warehousing. Experienced in designing scalable ETL workflows using Python, SQL, Spark, and Databricks on AWS and Azure.

EDUCATION

The George Washington University	Washington, DC
Master of Science, Data Science (GPA: 3.9)	May 2025
Relevant Coursework: Data Warehousing, Big Data Analytics, Cloud Computing, Distributed Systems, Database Design	
Manipal Institute of Technology	Manipal, India
Bachelor of Technology, Electronics, and Communication (GPA: 8.75/10.0)	August 2023

TECHNICAL SKILLS

Programming & Scripting: Python, SQL, Bash
Data Engineering & Cloud Tools: Apache Spark, Airflow, Azure Databricks, AWS Sagemaker, Azure Data Factory
Databases & Storage: PostgreSQL, Redshift
Cloud Platforms & DevOps: AWS, Azure, Git, Docker
Visualization & Reporting: Power BI
Certifications: AWS Academy – Cloud Foundations & NLP, SAS Certified Specialist – Visual Modeling

PROFESSIONAL EXPERIENCE

George Washington University School of Business	Washington, DC
Data Consultant	Jan 2023 – Present
<ul style="list-style-type: none">Streamlined faculty research workflows by designing and automating data pipelines in Python; advised on database querying strategies and implemented reproducible project templates using Git and cloud-hosted notebooks.	
ABC Analytical Organization, Inc.	Mumbai, India
Data Engineer (marketing and sales focus)	Sept 2020 - Jul 2022
<ul style="list-style-type: none">Led Image Recognition team; earned Superstar Award for Client Success by expanding markets from 18 to 25.Reduced dashboard refresh errors by 95% and improved SLAs across 25 markets by building and maintaining data pipelines for store analytics using Power Platform and Amazon Redshift.Independently developed and deployed Power BI Data Integrity Dashboard with backend automation via Python and Redshift queries, improving anomaly detection and reducing response time by 60%.Increased data throughput by 40% and saved \$5,000/month in compute costs by re-engineering data ingestion pipelines using SQL optimization and parallel processing.	

PROJECT EXPERIENCE

Cloud Data Warehouse Migration	Jan - May 2024
<ul style="list-style-type: none">Rebuilt company's internal reporting system in cloud using sample business data to improve speed and scalability.Reduced query time by 70% by migrating 2TB from SQL Server to AWS Redshift using dbt; automated ETL in Airflow.	
Marketing Analytics Data Lake Pipeline	Sept - Dec 2023
<ul style="list-style-type: none">Consolidated marketing data sources, including ad performance, email engagement, and web traffic, into unified reporting structure.Automated daily data loads using Azure Data Factory and Databricks, stored in Delta Lake, and visualized results in Power BI with 50% faster dashboard refresh times.	
Real-Time Retail ETL Pipeline	Sept - Nov 2023
<ul style="list-style-type: none">Built real-time data pipeline to track transactions and monitor sales and inventory patterns across 10 store locations.Developed Kafka–Spark streaming workflow to ingest and clean 50K+ sales records per hour; stored in PostgreSQL and reduced processing time from 15 to 2 seconds.	