

Phaneendhra Patel

Portfolio | +1 908 549 3266 | Phaneendhrak0427@gmail.com | [GitHub](#) | [LinkedIn: Phaneendhra-Patel](#) | USA

Professional Summary

Results-driven Data Engineer with experience in designing and optimizing ETL/ELT pipelines using Spark, SQL, and AWS. I am skilled in data modeling, API data extraction, and workflow automation to support big data analytics. Strong collaborator, building data-driven insights, BI solutions, and Google Analytics integrations.

Technical Skills

Programming: Python, R, Scala, Java, SQL, UNIX Shell Scripting.

Data Engineering: Apache Spark, PySpark, Kafka, Hadoop.

Database: MySQL, PostgreSQL, SQL Server, MongoDB, HBase, DynamoDB.

Big data Tools: YARN, Hadoop, Spark Streaming, AWS Kinesis

ETL Tools: Talend, Informatica, Mulesoft, Apache NiFi, Alteryx.

Cloud: AWS (EMR, EC2, Glue, S3, Redshift, Lambda), Google Cloud Platform (BigQuery, Dataflow), Azure (Data Factory)

Visualization Tools: Tableau, Power BI, Looker.

CI/CD and Data Modeling: Docker, Kubernetes, Jenkins, Terraform, GitHub.

Certifications: [Wells Fargo - Data Modeling](#), [Databricks](#).

Project Experience

Data Engineer | Aug 2024 – Present

Walgreens | Remote

- Refactored Python-based ETL workflows for process automation, reducing data transformation latency and improving application efficiency.
- Leveraged AWS services, including EC2, S3, RDS, and Auto Scaling, to streamline resource utilization and cost-effectively manage cloud workloads.
- Designed and deployed PySpark applications for both streaming and batch ingestion, facilitating seamless data integration into HDFS Data Lake.

Data Engineer - Intern | Jan 2024 – April 2024

Discover Financial | Illinois

- Designed and optimized ETL pipelines using Python and Spark, improving data processing efficiency by 30%.
- Built and maintained data lakes and batch/real-time processing applications, ensuring data availability and reliability.
- Developed data models and reporting solutions that enabled business insights and analytics across teams.

Associate Data Engineer | Oct 2022 – Nov 2023

Saint Louis University | Missouri

- Developed and maintained scalable ETL pipelines using AWS Glue and Python, ensuring efficient and reliable data extraction, transformation, and loading
- Worked closely with cross-functional teams to design user interfaces and business applications, ensuring that database specifications and requirements were met for seamless integration and performance.
- Optimized large-scale data processing tasks by implementing partitioning and indexing strategies in cloud-based databases, reducing query execution time by 40%.

Software Engineer | Dec 2019 – Feb 2022

Value Momentum Pvt Ltd | Hyderabad, India

- Centralized data warehouse using SQL ensuring seamless and efficient data flow across departments.
 - Improved data transformation and loading times by optimizing SQL queries and fine-tuning data pipelines, reducing overall processing time by 30%.
 - Worked closely with cross-functional teams to migrate legacy data systems to cloud platforms like AWS and Azure, ensuring smooth integration and real-time data availability.
 - Implemented robust data governance practices, including automated validation checks and monitoring workflows, enhancing data quality and compliance across projects.
-

Education

Saint Louis University, Saint Louis, Missouri | 05/ 2024

Master of Science in Computer Science and Information Systems (GPA: 3.55)

- Coursework: Information Retrieval, Database and warehouses, Applied Analytics & methods, Information system strategy & management, Data Visualization & Analytics, Software Development and Ethical Evidence-based decision making.