

# PHANEENDHRA KATAKAM

Portfolio | +1 908 549 3266 | Phaneendhrak0427@gmail.com | GitHub | LinkedIn – Phaneendhra Katakam | New Jersey, USA

## TECHNICAL SKILLS

**Programming:** Python, R, Scala, Java, SQL, Pyspark, GitHub Copilot, Bash.

**Database:** MySQL, PostgreSQL, Snowflake, SQL Server, Oracle, DBA, Db2, MongoDB, HBase, DynamoDB.

**Big data Tools:** HDFS, Hive, Delta Lake, Databricks, Amazon EMR, Spark, Hadoop, Airflow, Kafka, Snowflake.

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

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

**Visualization Tools:** Tableau, Power BI, QlikView, AWS QuickSight, Matplotlib.

**CI/CD and Data Modeling:** Docker, Kubernetes, Jenkins, Terraform, GitHub, Snowflake Schema, Data Vault.

**Certifications:** Wells Fargo - Data Modeling, Databricks - Fundamentals, Gen AI - Fundamentals

## EXPERIENCE

### Data Engineer

Walgreens, Illinois

May 2024 – Present

- Designed and analyzed ETL pipelines using python, reducing data processing time by 30% and enhancing application performance.
- Deployed scalable AWS cloud solutions leveraging EC2, S3, RDS, EBS, Elastic Load Balancer, and Auto Scaling Groups, improving system reliability by 50% and reducing operational costs by 20%.
- Built batch and real-time data processing applications using PySpark, enabling seamless ingestion of multi-source data into HDFS Data Lake, accelerating data availability by 40%.
- Created and maintained DDL/DML scripts in SQL and HiveQL to support analytics applications across RDBMS and Hive, resulting in a 25% increase in query performance.
- Led a team of 5 data engineers in the development of a data integration platform, reducing data latency by 35% and improving data consistency across departments.

### Junior Data Engineer

Discover Financial, Illinois

Jan 2024 – April 2024

- Developed ETL frameworks using Python and optimized SQL queries, processing over 500K weekly transactions for internal analytics and reporting, enhancing data processing efficiency by 30%.
- Automated data extraction and transformation workflows, increasing data loading efficiency by 25% and reducing manual intervention.
- Maintained and optimized a PostgreSQL data warehouse by writing complex SQL queries to clean, aggregate, and analyze customer and transaction data, improving reporting accuracy by 20%.
- Designed data validation scripts in Python, ensuring 99.9% data accuracy for business intelligence and reporting purposes, reducing data errors by 15%.

### Associate Data Engineer

Saint Louis University, Missouri

Oct 2022 – Nov 2023

- Built and managed scalable ETL pipelines using AWS Glue and Python, ensuring reliable data extraction, transformation, and loading processes, reducing data integration time by 30%.
- Collaborated with cross-functional teams to design database schemas and user interfaces, ensuring seamless integration and optimal performance, increasing user satisfaction by 25%.
- Conducted training sessions on data engineering best practices for over 20 team members, and maintained comprehensive documentation for team workflows and processes, improving team productivity by 15%.
- Optimized large-scale data processing by implementing partitioning and indexing strategies, reducing query execution time by 40% and enabling faster data retrieval.

### Software Engineer

Value Momentum Pvt Ltd, India

Dec 2019 – Feb 2022

- Centralized enterprise data into a unified data warehouse using SQL, improving data accessibility and flow across departments, resulting in a 30% improvement in decision-making processes.
- Enhanced data transformation and loading efficiency by optimizing SQL queries and fine-tuning pipelines, reducing processing time by 30%.
- Led the migration of legacy data systems to AWS and Azure, ensuring real-time data availability and seamless integration, reducing system downtime by 40%.
- Established data governance frameworks, including automated validation checks and monitoring workflows, to ensure data quality and compliance, increasing data accuracy by 20%.

## EDUCATION

Saint Louis University, Saint Louis, MO

May 2024

Master of Science in Computer Science and Information Systems

Vaagdevi Engineering College, India

Sep 2019

Bachelor of Technology in Electronics and Communication Engineering