

PHANEENDHRA KATAKAM

Phaneendhrak0427@gmail.com | (908) 549 3266 | github.com/Phani-kp/ | linkedin.com/in/phaneendhra-katakam/

SUMMARY

- Dynamic and results-driven Data Engineer with 2 years of experience in designing and implementing robust **data pipelines** and **ETL** processes. Proficient in **SQL**, **Python**, and big data technologies to extract, transform, and load data for analytics and business insights.
- Strong background in database management, data warehousing, and cloud technologies.
- Effective communicator with a passion for solving complex problems and optimizing data workflows to drive organizational success.
- Strong expertise in SQL-based technologies, ensuring seamless integration and accessibility of both structured and unstructured data for comprehensive analytics.
- Skilled in developing Python-based data pipelines, utilizing libraries such as **Pandas** and **NumPy** to optimize data extraction, transformation, and loading tasks for data accuracy and efficiency in large-scale environments.
- Experienced Data Engineer with a talent for crafting impactful data visualizations using tools like Tableau and **Power BI**, translating complex datasets into compelling visuals to facilitate data-driven decision-making.
- Proficient in handling big data technologies, including **Hadoop** and **Spark**, to design and implement scalable solutions for processing and analyzing vast datasets.
- Experienced in orchestrating complex data workflows and scheduling tasks with **Apache Airflow**, ensuring efficient and reliable execution of ETL processes within a scalable and maintainable framework.
- Expertise in troubleshooting various desktop and web-based applications to ensure optimal performance and reliability.
- Conducted complex data analysis using **Databricks SQL** and **PySpark** to derive actionable business insights.
- Proven ability to design and develop applications from the ground up, ensuring robust and scalable architecture.
- Strong analytical skills to assess and improve the performance of data processing systems.
- Proficiency in using data visualization tools to communicate insights effectively and support strategic business decisions.
- Knowledgeable in integrating diverse data sources to provide a unified view for better business intelligence.
- Skilled in maintaining and optimizing databases, ensuring high availability and performance of critical systems.
- Adept at collaborating with cross-functional teams, providing technical expertise and support to achieve project goals.
- Expertise in Databricks for building, managing, and optimizing data pipelines and workflows.
- Extensive experience with Microsoft Azure cloud platform, utilizing services such as **CosmosDB**, **Azure Storage**, **Application Services**, and **Azure Data Factory** for comprehensive data solutions.
- Skilled in optimizing cloud PaaS resources, adhering to best practices to reduce costs and enhance performance.

TECHNICAL SKILLS

- Data Engineering : ETL development, data pipeline design, data modeling.
- Programming Languages : Python, SQL, C, HTML, CSS, Java.
- Tools : Git, Tableau, Power BI, Databricks.
- Big Data Technologies : Apache Spark, Hadoop, Spark SQL.
- Database Management : PostgreSQL, MySQL, MongoDB.
- Cloud Platforms : AWS (Amazon Redshift, S3), Google Cloud Platform (BigQuery).

EXPERIENCE

Amazon Connect (Intern)

Aug 2023 – Oct 2023

Client : Saint Louis University, Saint Louis, MO

Responsibilities:

- Configured and optimized Amazon Connect, designing contact flows, queues, and routing profiles to enhance call handling and agent efficiency.
- Integrated Amazon Connect with **AWS Lambda**, enabling custom functionality and utilized Amazon **DynamoDB** for efficient storage and retrieval of customer information.

Software Engineer

April 2021 - Aug 2022

Company: Value momentum Software Services Pvt Ltd, Hyderabad, India

Client: Geico, USA

Responsibilities:

- Collaborated with cross-functional teams to ensure data design and analytics requirements were met, maintaining data integrity and consistency across all projects.
- Designed, built, and maintained data pipelines and **ETL** processes, optimizing data workflows for maximum efficiency.
- Implemented scalable ETL processes using **Apache Spark** and **Python**, reducing data processing time by 30%.

Data Engineer Intern

May 2020 - July 2020

Company: PWC Pvt ltd, Bengaluru, India

- Managed and optimized databases (**PostgreSQL**, **MongoDB**) for performance, including schema design and index optimization.
- Developed data pipelines for batch and real-time data ingestion from diverse sources, ensuring data consistency and accuracy.
- Supported data governance initiatives by documenting data pipelines, processes, and data dictionaries.

EDUCATION

Saint Louis University

Masters in computers and information sciences.

Vaagdevi Engineering College

Bachelors in Electronics and Communication's.

Saint Louis, MO

Graduated, 2024

Warangal, India

Graduated, 2021

PROJECTS

Portfolio: <https://phani-kp.github.io/Portfolio/>

- Showcased expertise in web development through the creation of a dynamic and visually engaging online portfolio using **HTML**, **CSS**, **React**, **JavaScript**, and **PHP**.
- Deployed the portfolio project on **GitHub**, demonstrating adeptness in version control and collaborative work within a development team.