

LAKSHMI VARAHA KRISHNA CHITTELLA

DATA ENGINEER

Sioux Falls, SD · krishna.chittella@gmail.com · [Portfolio](#) · [GitHub](#) · [LinkedIn](#)

PROFESSIONAL SUMMARY

Data Engineer with 2.5 years of experience developing Python and SQL data pipelines, Snowflake transformations, and Spark-based processing. Delivers dependable operational reporting through incremental loads, data modeling, reconciliation, and production troubleshooting; independently launched the live Cover Next data product.

TECHNICAL SKILLS

Programming & Processing: Python, SQL, Java, Apache Spark, PySpark, Spark SQL, pandas

Data Engineering: ETL/ELT, Data Warehousing, dbt, Data Modeling, Incremental Loads, Data Quality, Reconciliation, Airflow

Databases & BI: Snowflake, BigQuery, PostgreSQL, Oracle, MongoDB, DynamoDB, Tableau, Power BI

Cloud & Streaming: AWS (S3, Glue, EMR, Lambda, Kinesis, EC2, CloudWatch), Azure (Data Factory, Databricks, Blob Storage, Synapse, Event Hubs), GCP (Cloud Storage, Cloud Monitoring), Kafka

Engineering Tools: Git, GitLab, GitHub Actions, CI/CD, Docker, Kubernetes, Terraform, Linux, REST APIs

PROFESSIONAL EXPERIENCE

Data Engineer | Softtechers LLC

Dec 2025 – Present

- Engineered Python and SQL pipelines integrating case, customer, and activity records from PostgreSQL and REST APIs for operational analytics.
- Built paginated API extractors with authentication, bounded retries, exception handling, and structured execution logs.
- Integrated case assignment, status-change, and resolution events into consistent datasets for downstream reporting.
- Modeled current case records alongside event histories to support point-in-time analysis and operational reporting.
- Created staging transformations using SQL CTEs, joins, window functions, and aggregations to prepare analytical datasets.
- Standardized customer identifiers, case statuses, timestamps, and source fields across disparate application systems.
- Implemented null, duplicate-key, and source-to-target reconciliation checks to prevent invalid records from reaching reports.
- Replaced full reloads with timestamp-based incremental extraction and SQL upserts for new and modified records.
- Maintained scheduled runs, dependency checks, execution logging, and controlled reprocessing for failed data refreshes.
- Published SQL reporting views for case volume, open backlog, resolution trends, and service-level performance.

Data Engineer | TechRoots Inc.

Jan 2025 – Nov 2025

- Supported 4 production ETL/ELT pipelines in Snowflake delivering customer, sales, and billing datasets for reporting.
- Developed SQL transformations for staging and reporting tables using joins, aggregations, and business-rule validations.
- Applied PySpark and Python to clean source fields, standardize data types, and process structured datasets.
- Supported incremental data-loading workflows and validated source updates before loading reporting tables.
- Reconciled source and target records with SQL counts, aggregates, key comparisons, and duplicate/null checks.
- Diagnosed pipeline failures involving SQL errors, missing inputs, and dependency issues; verified controlled reruns.
- Documented recurring failure scenarios, validation queries, and recovery steps to streamline troubleshooting.

Associate Systems Engineer | Tata Consultancy Services

Oct 2022 – May 2023 · Kolkata, India

- Validated Oracle batch ETL loads of customer and transaction data into operational reporting tables.
- Wrote SQL joins and aggregations; checked record counts, null fields, duplicate identifiers, and key consistency.
- Investigated batch failures with senior engineers, verified completeness after reruns, and supported historical reprocessing.
- Maintained ETL mappings, test evidence, and operational runbooks for troubleshooting and team knowledge transfer.

SELECTED PROJECTS

Cover Next | [Live site](#) | [GitHub](#) — Built a live creator analytics site comparing Wikimedia 7-day trends with recent YouTube coverage and IGDB game identity.

- Automated daily Python collection and Next.js deployment via GitHub Actions; built dbt/BigQuery models and Terraform-managed GCP infrastructure.

Power-Line Fault Classification | [GitHub](#) — Applied transfer learning to classify faults in insPLAD power-line inspection images.

EDUCATION & LEADERSHIP

M.S. Computer Science · University of South Dakota · Dec 2024 · GPA 3.6/4.0

B.Tech Computer Science & Engineering · BVC College of Engineering (JNTUK) · 2022

Student Lead · University of South Dakota Campus Retail Food Services · 2024 · Staff training and operations