

# NITHEESH CHANAMBATLA

## SENIOR DATA ENGINEER

Dayton, Ohio | (937) 907-2325 | [nitheeshc2325@gmail.com](mailto:nitheeshc2325@gmail.com)

[linkedin.com/in/nitheesh2325](https://linkedin.com/in/nitheesh2325) | [github.com/Nitheesh2325](https://github.com/Nitheesh2325)

### PROFESSIONAL SUMMARY

---

Senior Data Engineer with 10+ years of experience building batch and streaming pipelines, cloud data platforms, and analytical data models across healthcare, financial, and enterprise environments. Hands-on background with Python, SQL, Spark, Airflow, Hadoop, PostgreSQL, and AWS data services, supported by frozen portfolio projects with measured local pipeline and data-quality evidence.

### TECHNICAL SKILLS

---

Python, SQL, Java, Scala, Shell | Apache Spark, Hadoop, Hive, Presto | Airflow, AWS Glue, Informatica | AWS S3, Athena, Kinesis, Data Firehose, Lambda, Redshift, IAM, KMS, CloudWatch | PostgreSQL/PostGIS, SQLite | Dimensional Modeling, SCD Type 2, ETL/ELT, Data Quality, CDC | Git, Jenkins, Docker, ELK

### PROFESSIONAL EXPERIENCE

---

#### **Data Engineer | Discover (Client Engagement) | Chicago, Illinois | Jan 2021-Present**

- Built and maintained Python, Airflow, and Spark ETL/ELT pipelines supporting analytics and machine-learning datasets.
- Designed and deployed an AWS data lake integrating S3, Glue Data Catalog, and Athena across structured and unstructured sources.
- Implemented real-time ingestion with Kinesis, Data Firehose, and Lambda for analytics and alerting workflows.
- Migrated enterprise data systems to Amazon Redshift and optimized schemas for BI query efficiency.
- Automated pipeline testing and delivery with Git, Jenkins, and Docker; monitored failures, latency, and data quality through CloudWatch and ELK.
- Applied IAM, KMS encryption, and role-based access controls while supporting GDPR- and HIPAA-aligned data handling.

#### **Data Engineer | Cigna (Client Engagement) | Bengaluru, India | Aug 2018-Dec 2020**

- Developed Spark, Hive, and Presto batch pipelines for large-scale financial transaction data in a healthcare client environment.
- Designed star and snowflake schemas and applied partitioning, bucketing, and indexing for analytical workloads.
- Built Python frameworks for ETL validation, error handling, and performance monitoring.
- Deployed Hadoop workloads across HDFS and YARN and integrated CDC with Debezium and Kafka Connect.
- Embedded cataloging, classification, and retention controls and supported production incident response and root-cause analysis.

#### **Data Engineer | Wipro Technologies | Chennai, India | Jul 2016-Jul 2018**

- Developed Informatica and SQL ETL workflows across enterprise systems in finance and healthcare domains.
- Designed reporting data models and implemented Python validation for large-scale batch loads.
- Supported legacy-to-Hadoop migrations and on-premises-to-AWS initiatives using S3 staging and Redshift warehouse layers.

## SELECTED PROJECTS

---

### **NASA Earth Observation Event Intelligence Platform | Flagship, Frozen, Verified**

- Built and verified a local NASA FIRMS pipeline that processed 1,000,000 replay events with Spark in 149.502 seconds, measuring 6,688.88 events/second.
- Validated 1,000,000 events in PostgreSQL/PostGIS and independently verified deterministic generation and read-back at 10,000,000 records; designed and locally validated AWS infrastructure without deployment.

### **Customer360 Analytics Warehouse | Supporting, Frozen, Verified**

- Built a PostgreSQL dimensional warehouse loading 400,000 sales facts across five source domains with tested SCD Type 2 customer and product dimensions.
- Implemented surrogate-key resolution, validation, rejected-record quarantine, constraints, and repeat-safe loading; executed 20 analytical SQL queries and passed 20 automated tests plus hosted CI.

### **RetailSync Data Platform | Supporting, Frozen, Verified**

- Built a deterministic local batch pipeline with a separately verified 100,000-order default run and transactional SQLite full-snapshot replacement.
- Added non-numeric validation and quarantine tests preventing invalid measures from reaching SQLite; passed six local tests and hosted CI on a 1,000-row fixture.

## EDUCATION

---

Campbellsville University | Master's in Data Science | Dec 2024

VIT University | Bachelor's in Science, Multimedia and Animation | Mar 2021

## CERTIFICATIONS

---

- AWS Certified Data Analytics - Specialty
- AWS Certified Solutions Architect - Associate
- Google Cloud Professional Data Engineer