TEJA KANDUKURI

Michigan — +1-346-358-7193 — Writetotejak@gmail.com — linkedin.com/in/kandukuri-teja

---

Professional Summary

Certified Data Engineer skilled in designing scalable, cloud-native data platforms for insurance, finance, and healthcare. Expert in GCP, AWS, real-time ingestion, and end-to-end ETL/ELT architecture. Adept at optimizing data quality, automating workflows, and delivering reliable, cost-efficient cloud solutions for analytics and machine learning.

Core Skills

Cloud Platforms & Data Infrastructure: BigQuery, Airflow, GCS, Dataflow, Cloud Functions, Pub/Sub, Dataproc, Vertex AI, Cloud SQL, S3, Glue, Redshift, Lambda, EMR, DMS, Kinesis, Data Warehousing Solutions, Data Lake Management, Data Storage Solutions

Data Engineering & Integration: ETL/ELT Pipeline Development, Batch and Streaming Data Integration (Apache Spark), CDC (Kafka, Debezium, Fivetran), Data Modeling Best Practices, Data Profiling Techniques, Airflow (Composer), GoAnywhere SFTP Integration/Testing/Automation Tools (FitNesse/Selenium)

Analytics & Reporting: Looker Studio, Tableau Dashboards, OBIEE, KPI Reports

Programming & Optimization: Python, PySpark, SQL Optimization Techniques

Project Delivery & Collaboration: Jira Project Tracking, Agile Methodologies, GIT Version Control

---

Work Experience

Data Engineer GCP   Mar 2023 – Present
Definity Financial Corp

- Built and optimized GCP-native ETL/ELT pipelines using Cloud Composer (Airflow), Cloud Functions, Dataflow, Pub/Sub, Python, and BigQuery to support enterprise reporting, analytics, and AI workloads.

- Designed integrations for Canada's first P&C claims system, ingesting data from Guidewire, Salesforce, Verisk, and external APIs into BigQuery with automated schema validation and schema evolution.

- Implemented real-time and batch CDC pipelines (AWS → GCP), improving data freshness and pipeline throughput by 25–30% while reducing pipeline incidents by ~40%.

- Engineered high-performance BigQuery datasets using partitioning, clustering, and optimized MERGE logic, improving query performance by ~25% and reducing compute costs by 20–25%.

- Optimized Airflow orchestration and schema automation, improving data availability SLAs by ~30% and lowering manual operational effort by ~40% through workflow automation.

- Built an AI Contact Center Analytics pipeline using Vertex AI, Contact Center AI (CCAI) Insights, Dialogflow CX, and Speech-to-Text, enabling automated transcription and sentiment/emotion detection across 100% of recorded calls.

- Reduced manual call review effort by 50–60% and improved customer experience metrics by 10–15% through full-population call analytics and AI-driven insights.

- Enabled actuarial and underwriting teams to reduce manual pricing analysis by 30–40% and improve pricing accuracy by 12–18% through curated analytical datasets and AI-ready data models.

- Designed a Dynamic Pricing & Risk Intelligence Engine using MCP + RAG + Vertex AI, analyzing underwriting rules, loss trends, vehicle telematics, and territory risk factors to recommend risk-adjusted pricing with SQL-backed evidence.

- Built RAG pipelines using Vertex AI embeddings + BigQuery Vector Search to index underwriting manuals, policy documents, and internal guidance for AI-driven retrieval and decision support.
- Built BigQuery SQL routines to automate ETL test case generation and validation at scale.
- Developed JavaScript-based BigQuery UDFs and scripts to compute data coverage and null percentage metrics across 200+ tables, improving data quality observability.
- Integrated data quality and test result metrics into Looker Studio dashboards for operational and stakeholder visibility.
- Conducted data profiling, anomaly detection, and validation using PySpark on Dataproc, improving pipeline reliability and reducing recurring incidents.
- Led production support, root-cause analysis, and cost optimization initiatives across GCP workloads.

ETL Engineering Analyst   **Aug 2018 – Apr 2021**
Accenture, India

- Developed large-scale data pipelines and lakes using AWS Glue, Lambda, Step Functions, and S3.
- Migrated on-prem Oracle workloads to AWS using DMS, Redshift, and EMR.
- Improved Redshift performance with distribution keys, sort keys, and query tuning.
- Built real-time and batch ingestion frameworks using Kinesis and S3 for analytics.
- Integrated Glue Data Catalog for governance; built automated data quality checks with Python.
- Created Power BI dashboards connected to Redshift/RDS for actionable insights.
- Resolved production pipeline issues to ensure SLA adherence and uptime.
- Collaborated with product owners to translate requirements into technical specifications under Agile delivery.

ETL Test Engineering Analyst   **Mar 2018 – Aug 2018**
Accenture, India

- Automated OBIEE UI regression testing with Selenium to reduce effort by 25%.
- Created test scenarios and SQL validation scripts; automated test executions using FitNesse-based SQL suites.
- Validated source-to-target mappings, transformation logic, and BI report outputs.
- Performed integration/regression testing; documented results; participated in Agile ceremonies.

Education

Master's in Big Data Solutions Architecture – 3.8 GPA   **2022**
Conestoga College

Certifications

GCP Professional Data Engineer (2025) — AWS Data Analytics Specialty (2023) — AWS Cloud Practitioner (2021)