

SUMMARY

Data Engineer and Analytics Leader with 10+ years of experience in data analysis, reporting, and pipeline development. Skilled in building ETL workflows, maintaining SQL data warehouses, and using data to deliver actionable insights. Experienced with Python, SQL, and Power BI to automate data processes, streamline reporting, and improve data quality. Adept at solving complex challenges, creating scalable solutions, and mentoring teams to achieve technical excellence. I am seeking to expand into new domains, grow my skills as a data engineer, and deliver solutions that have a tangible, real-world impact.

SKILLS

Languages & Frameworks: SQL, Python, Spark (PySpark), Prefect, Pandas, Scikit-learn, Plotly, JavaScript

Data Engineering & Analytics: ETL/ELT pipelines, Delta Lake, API integration, Batch/Stream processing, Data modeling, NoSQL, A/B testing, Machine learning, Data validation

Tools & Platforms: Databricks, GCP (Cloud Run, Pub/Sub, Firestore, BigQuery), Git, Power BI, SQL Server

CERTIFICATIONS

MICROSOFT CERTIFIED: POWER BI DATA ANALYST ASSOCIATE | December 2022 | [Credly Badge](#)

EXPERIENCE

Data Strategy & Analytics Lead (Volunteer) | Jesus Said Go | Remote | August 2025 - Present

- Managed CRM workflows and data quality, automating processes and improving reporting for stakeholders.
- Provided technical guidance and coordination on data and CRM initiatives, ensuring alignment with organizational goals.
- Developed an automated WhatsApp reporting system using Deluge and API integrations to deliver personalized weekly performance stats and secure mobile verification for hundreds of disciple-makers across West Africa.

Principal Business Analyst / Data Lead | San Joaquin County Probation | Stockton, CA | March 2023 - Present

- Recruited and mentored three analysts in data engineering, analytics, and reporting, establishing team practices that improved pipeline development, automation, and reporting capabilities.
- Built an automated data orchestration system using Python and SQL, creating a library of jobs running ~30 scheduled scripts to ingest, transform, and load multi-source data into a centralized SQL Server warehouse.
- Produced over 20 multi-page Power BI dashboards department-wide, increasing accessibility, efficiency, and transparency in reporting [[Samples](#)].
- Designed and produced three annual evaluation reports with statistical analysis, trends, and performance metrics to support evidence-based decisions [[Report sample](#)].

Business Analyst II | San Joaquin County Probation | Stockton, CA | December 2014 – March 2023

- Developed business metrics, complex SQL queries, and Power BI reports to enable self-service analytics for management.
- Automated 20+ recurring reports and data extraction processes using Python, SQL, and VBA, reducing manual workload and improving accuracy.

- Maintained data integrity and quality across multiple systems during Oracle-to-Dynamics 365 CRM migration, advising on mapping and validating data.
 - Developed an NLP-based machine learning model using probation case notes to predict client outcomes, achieving 89% accuracy and delivering actionable insights for management decision-making [[Notebook](#)].
-

PROJECTS

NonProfit Website Clickstream Pipeline & Analytics | [[Github](#)]

- Designed and implemented, for a live production website, a real-time clickstream analytics pipeline using JavaScript, Google Cloud Run, Pub/Sub, Firestore, BigQuery, and Looker to capture anonymized website engagement events and deliver session-level analytics.

Earthquake ETL Lakehouse Pipeline | [[Github](#)]

- An end-to-end PySpark + Delta Lake pipeline in Databricks ingesting real-time USGS earthquake API data into Bronze/Silver/Gold medallion layers with incremental metadata tracking, optimized analytic tables, and visualization outputs for reporting.

Etsy Shop Data Pipeline | [[Github](#)]

- An ETL pipeline using Python + Databricks to ingest Etsy + Gelato API data, normalize nested JSON, and load curated analytic tables into Databricks SQL Warehouse with idempotent, modular engineering patterns.

Powerlifting Data Analysis & Visualization | [[Github](#)]

- Extracted, cleaned, and feature-engineered OpenPowerlifting competition data in Python, storing results as Parquet files. Explored data for insights and produced an interactive Power BI dashboard to analyze trends and participant performance.
-

EDUCATION

California State University - Stanislaus

B.S. Mathematics | Minor in Computer Science | December 2013