Elnaz Watkins

SQL/ETL Developer | Python · SQL · AWS (S3, Lambda) · ETL Pipelines · Machine Learning

Austin, TX · (413) 801-4763 · elnaz.watkins@gmail.com · linkedin.com/in/elnazwatkins · https://elnazwatkins.github.io/

Summary

Data engineer and former data scientist with 9+ years in software, data and machine learning roles. At Seismos, built ETL pipelines loading MATLAB, SQL Server, CSV and Google Sheets data into AWS Aurora PostgreSQL, wrote stored procedures invoked by AWS Lambda functions to process real-time sensor streams, and led a schema and data migration with integrity checks and no interruption to ingestion. At Amdocs, built predictive ML and time-series forecasting models in Python, automated data preprocessing to cut manual work 90%, and cut an 8-hour SQL job to 5 seconds. M.S. in Computer Science with published ML research.

Technical Skills

- **ETL and data pipelines:** ETL design, multi-source extraction (CSV, SQL Server, MATLAB, Google Sheets), data cleaning, transformation and validation, integrity checks, training-data preparation
- **SQL and databases:** complex query writing and optimization, execution plans, indexing, stored procedures, data modeling; PostgreSQL (AWS Aurora), SingleStore (distributed SQL), SQL Server
- **Machine learning:** supervised and unsupervised learning, predictive modeling, time-series forecasting, XGBoost, feature engineering, model evaluation
- **Programming:** Python (Pandas, NumPy), SQL, R, SQLAlchemy, FastAPI, Java
- **ML and AI frameworks:** TensorFlow, Keras, LangChain, LLM fine-tuning
- **Cloud (AWS):** S3, Lambda, Aurora, SageMaker, DynamoDB, IoT Core, EC2, CloudWatch
- **Tools:** Git, GitHub, Bitbucket, Jira, Alembic, CI/CD (CircleCI), Grafana, Splunk, Tableau, Matplotlib, seaborn, Linux, Claude Code, Cursor

Professional Experience

Seismos, Inc.

Database & Data Engineer · Austin, TX · Sep 2024 – Oct 2026

- Owned production databases and data pipelines end to end: a distributed SingleStore cluster ingesting 20 Hz and 1 Hz DAQ (data acquisition) sensor time series, plus the relational databases.
- Built ETL pipelines that extract, transform and load historical data from MATLAB (.mat) files, SQL Server backups, CSV files and Google Sheets into AWS Aurora PostgreSQL.
- Developed stored procedures in SingleStore and Aurora PostgreSQL, invoked by serverless AWS Lambda functions, to transform and load real-time sensor data streams.
- Designed relational tables and data pipelines that prepared training data for the data science team, drawing on a data science background to shape the data around how models would use it.
- Led the database platform change end to end: evaluated options, chose AWS Aurora PostgreSQL for relational and 1 Hz time-series data, kept 20 Hz data in SingleStore, and migrated schema and data on schedule with no interruption to ingestion.
- Resolved data-integrity problems caused by SingleStore's lack of foreign-key support with a project-by-project migration and integrity checks, so orphaned and inconsistent records never reached Aurora.

- Fixed data and system errors; cut SingleStore CPU from 80% to 65% by indexing a join column and tuning queries.
- Cut SingleStore costs 50% by removing idle branches vendor support had missed and right-sizing compute.
- Introduced Alembic version-controlled DDL with PR review and CircleCI, letting application teams change data models independently and safely.
- Provided on-call support one week per month, meeting SLA targets, and validated backup restores.
- From July 2026, also took on backend work: added endpoints and fixed bugs in a FastAPI app (SQLAlchemy, Pydantic) through PR review and CircleCI.
- Built an AI-assisted workflow (Claude Code) that takes Jira tickets to tested pull requests.

Hill Country Community Ministry

Freelance Data Analyst · Leander, TX · Jan 2024 – Sep 2024

- Cleaned and preprocessed raw data; automated validation and reporting dashboards (Excel VBA, Google Apps Script) and ran statistical analyses of food demand.

Amdocs Inc. · Apr 2017 – Mar 2023

Data Scientist · Remote · Sep 2020 – Mar 2023

- Built predictive ML models in Python forecasting revenue loss, promotion impact and server overloads.
- Implemented time-series forecasting models that quantified how effectively promotions drove revenue.
- Automated cleaning and preprocessing of large datasets for modeling, cutting manual work 90%.
- Rewrote core analysis processes as set-based SQL and Python, cutting run time from 8 hours to 5 seconds.
- Analyzed server logs to forecast overloads, improving capacity headroom 30% on Black Friday; found 15% in billing revenue leaks.
- Led a 6-person team automating a data-preprocessing and ingestion project.

Software Engineer · Champaign, IL · Apr 2017 – Aug 2020

- Supported backend and frontend of CRM and inventory applications for a major U.S. telecom provider.
- Diagnosed data errors affecting application performance using advanced SQL, Splunk, and log analysis, reducing open tickets by 30%.
- Root-caused production issues and shipped Java fixes during critical-incident bridge calls, keeping apps within SLAs.

Southern Illinois University · Carbondale, IL · 2014 – 2017

- **Instructor and Graduate Teaching Assistant:** taught Linux administration, Oracle/SQL, SQL databases, and information storage and management.
- **Published ML research:** classification model in R to predict social-network check-in locations, with a novel noise-handling method (CATA 2018).

Education

- **M.S. Computer Science** (GPA 4.0/4.0), Southern Illinois University, Carbondale, IL · May 2017
- **B.S. Computer Software Engineering, University of Qazvin, Iran**
- **Certificates:** Deep Neural Networks (hyperparameter tuning); Unsupervised Learning and Reinforcement Learning; Data Lakes on AWS