

Joseph Saldivar

559-385-3832 | josephsaldivarg@gmail.com | [linkedin/josephasaldivar](https://linkedin.com/josephasaldivar) | github.com/Sque3ze | [portfolio](#)

EDUCATION

University of California, Berkeley

Berkeley, CA

Bachelor of Arts in Data Science

May 2025

- Data Structures, Data Engineering, Machine Learning, Data Mining and Analytics, Database Design, Discrete Mathematics and Probability, Structure & Interpretation of Programs, Principles & Techniques of Data Science

EXPERIENCE

Software Developer

May 2024 – Present

Data Annotation

- Partnered with AL/ML engineering teams to validate and refine LLM code outputs, improving execution accuracy by 20% and reducing defect rates across production-level workflows
- Designed and maintained automated test environments leveraging virtual environments and dependency managers, cutting environment setup time by 30%+ and ensuring consistent, reproducible builds.
- Debugged/Optimized AI-generated code (Python, Java, JavaScript, HTML, CSS) and authored eval frameworks to flag recurring model errors—informing retraining and reducing repeat issues 15%+.

Independent Full-Stack Developer

Sep. 2025 – Present

Spotify Data Visualizer (TrackGraph) — trackgraph.josephsaldivar.dev

- Shipped a full-stack Spotify analytics app used by 60+ users, analyzing 40k+ listens across a users entire listening history in a interactive React based frontend with bubble & leaderboard visualizations.
- Integrated Spotify Web API alongside a custom FastAPI to perform batched fetches, and a S3-backed metadata cache; cache-first hydration cuts external calls by 90%+, only requesting when new data is detected.
- Productionized on AWS (App Runner, ECR, S3/CloudFront, IAM) with a CI/CD pipeline, enabling safe, repeatable rollouts with automated docker container build/push, and cache versioning/invalidation.
- Optimized with vectorized Pandas aggregations and timeframe-keyed memoization, reducing recompute by 66%.

Math Tutor/Mentor

November 2021 – August 2022

State Center Community College District

- Provided 1:1 math tutoring by assessing learning styles, adapting tutor methods, and refining based on outcomes.

PROJECTS

Observability for TrackGraph | *Python, Pulumi, Docker, Grafana, Prometheus, AWS* October 2025 – Present

- Shipped a modular Pulumi (Python) IaC stack for my site on AWS(App Runner API, ECS Fargate collector, S3 buckets, ECR) backed by multi-stage Docker builds and GitHub Actions, CI runs tests + Pulumi preview on each push; deployments executed via pulumi up.
- Delivered end-to-end observability: OpenTelemetry traces to a custom App Runner collector; Prometheus metrics remote-written to AWS Prometheus; dashboards in Grafana (Prometheus/CloudWatch/X-Ray datasources).
- Built a CloudFront analytics toolkit (Lambda parser + aggregator + publisher) that ingests S3 log objects, handles gzip, and emits per-minute CloudWatch metrics (requests, bytes, 2xx/4xx/5xx, latency)

Database Design for Cal Field Hockey | *Python, MySQL, SQL, ML* September 2024 – December 2024

- Worked in an 8-member team to design and implement a relational database supporting player, event, and merchandise management; contributed to requirements gathering, schema design, build, and client delivery.
- Created 62 BCNF-normalized tables in MySQL Workbench; integrated historical and synthetic datasets to enable high-volume testing and ensure scalability.
- Developed SQL and Python analytics pipelines, including a Random Forest-based MVP prediction model and a lineup optimization tool, validated against historical results to support roster selection decisions.

TECHNICAL SKILLS

Languages: Python, JavaScript, TypeScript, Java, SQL, HTML/CSS

Stack: React, D3, FastAPI, Pandas, Uvicorn, Node.js/npm, PostgreSQL

Cloud/Dev Ops: Docker, Pulumi, Git/Github Actions, AWS (App Runner, S3, CloudFront, ECR, CloudWatch, X-Ray, Lambda, ECS Fargate), Prometheus, Grafana, OpenTelemetry, Playwright, Pytest