

FELIPE DUENAS

duenasfd@gmail.com | (510) 828-7750 | Oakland, CA | [LinkedIn](#) | felipeduenas.dev

EDUCATION

University of California, Los Angeles (UCLA)

Expected March 2027

B.S. Statistics & Data Science | Minor in Data Engineering

- Relevant Coursework: Software Engineering, Database Systems, Machine Learning, Data Mining, Generative Data Science, Cloud Computing, Neural Networks, Data Structures & Algorithms

EXPERIENCE

Amazon Web Services (AWS)

Seattle, WA

Software Engineer Intern, Data Platform

May 2026 – Present

- Engineered ETL pipeline transforming information from **3,000+ Tableau workbooks** and 10+ supply-chain sources into an AI-ready knowledge base, expanding queryable data by **85%** for an internal agentic-AI platform for data center planning
- Architected agentic dashboard generation system which cut dashboard development time **4 days to 3 minutes**, using LangGraph and new knowledge base to create context-specific and data-correct dashboards from NL prompts
- Developed agentic platform's improved data model and built its retrieval-evaluation framework, increasing agent **accuracy by 13%**, and driving adoption of platform's self-serve analytics to **5,000+ members** of AWS's infrastructure organization
- Analyzed usage patterns on legacy dashboards using Redshift and SQL, partnering with planning, infra, and BI stakeholders to identify **15 high-value features** guiding the platform's continued tool development

UCLA Trustworthy AI Labs

Los Angeles, CA

Backend & Data Engineer Intern – Agentic Data Science

January 2026 – May 2026

- Built an agentic data science application that autonomously plans and executes multi-step analysis pipelines from natural language prompts, achieving a **71% task completion** on the BIRD benchmark, outperforming GPT-4o by 8 percentage points
- Architected a serverless backend using Dockerized AWS Lambda behind API Gateway, provisioned via AWS CDK, to handle concurrent LLM inference requests across the agent's multi-step tool-calling pipeline
- Implemented an event-sourced persistence layer using Supabase (PostgreSQL) and S3 to solve distributed state management across the serverless multi-agent architecture, maintaining consistent agent state with **zero data loss** across **400+ test sessions**
- Created a data-centric evaluation harness spanning 34 table schemas and 120+ question types across 20+ specialized tool-calling functions, driving 3 architectural changes that improved task **accuracy by 11%** over baseline

Amazon Web Services (AWS)

Arlington, VA

Business Intelligence Engineer Intern

June 2025 – September 2025

- Developed an ensemble forecasting model to predict data center power utilization, achieving **1.9% MAE** vs. a 6.3% baseline and identifying under usage that informed a **\$400K** cost saving initiative across 12 AZs
- Engineered ELT pipelines using dbt and Redshift to process **500M+ daily telemetry events** from data center hardware, creating BI-ready data structures that helped reduce reporting time from **6 hours to 30 minutes** weekly for planners
- Designed a statistical anomaly detection system to flag capacity deviations in global S3 Glacier demand, replacing manual Excel-based monitoring workflows and catching **14 anomalies monthly** before they impacted capacity
- Built a centralized Tableau dashboard with a GenAI natural-language query interface for capacity planning, unifying outputs from 7 legacy sources and cutting data-gathering time by **4 hours** per planning project

Acer

San Jose, CA

Data Analyst Intern

June 2024 – September 2024

- Created a unified data model in PostgreSQL on 30+ years of legacy data, eliminating siloed systems and increasing data **accessibility by 125%** for Acer's legal team
- Automated monthly SLA reports for General Counsel by building Tableau dashboards connected to centralized data repository, reducing data reconciliation and report preparation time from **2 hours to 15 minutes** per process
- Identified a **\$22K** licensing discrepancy through analyzing data trends from merger and financial documents, leading to standardizing currency in reports and consolidating data collection across teams in a single platform as source of truth

SKILLS

- **Languages:** Python, SQL, Java, R, Bash
- **Pipeline & DevOps:** Spark, Airflow, dbt, Kafka, Docker, CI/CD, Git, Terraform, CDK, Data Modeling
- **Cloud:** AWS, GCP, Snowflake, Redshift
- **Generative AI:** Claude, LLM-as-a-judge, RAG, LangGraph, Prompt Engineering, MCP, Vector Databases, AI Agents
- **Machine Learning:** Regression, Classification, Clustering, NLP, Neural Networks, LLM Orchestration, PyTorch, scikit-learn