Daniel Pineda

DATA ENGINEER | FULL-STACK BACKGROUND

+57 300 510 3421 | danielpinedasalazar19@gmail.com | github.com/Danielpinedasalazar
| portfolio-eta-kohl-24.vercel.app | linkedin.com/in/daniel-pineda-s

PROFILE

Systems Engineering student (9th semester, EIA University) transitioning into Data Engineering, currently completing the DataTalksClub Data Engineering Zoomcamp (Docker, Postgres, Terraform, GCP/BigQuery, Kestra, dbt, Spark, Kafka). One year of production experience building and deploying cloud-based systems on AWS, plus hands-on machine learning work fine-tuning and serving a medical imaging model. Looking for a remote junior Data Engineer role.

FEATURED PROJECT — SECOP II

Public Procurement Analytics — GCS, BigQuery, dbt, FastAPI, Next.js

- Built a daily ELT pipeline over Colombia's SECOP II public-procurement data, landing raw extracts in GCS and transforming them in BigQuery with dbt.
- Designed an ELT warehouse so failed transformations never require costly re-extraction; pre-aggregated analytical marts for fast dashboard queries.
- Built FastAPI and Next.js layers for procurement analytics, retaining supplier concentration as the only live fact-table query (~163 MiB per entity).
- Represented missing values as first-class, hatched gaps—not zeros—and used texture rather than colour for accessible, print-safe visualizations.
- Served row-count metrics live from the API health endpoint, keeping dashboard figures aligned with daily mart rebuilds.

EXPERIENCE

Full-Stack Engineer — APPU Store *1 yr*

- Managed cloud data infrastructure on AWS: DynamoDB for persistence, S3 for storage, and IAM for access control, running on ECS.
- Built and maintained backend services in Django processing transactional data for a POS system used by 2,000+ active users, containerized with Docker.

OTHER PROJECTS

OncoVision — Breast Histology Classifier *Python, PyTorch, FastAPI*

- Built a multi-task EfficientNet-B0 classifier over 7,909 breast histology slides, predicting benign/malignant status and 8 clinical subtypes from a shared backbone.
- Chose a single multi-task head to share visual features across both tasks and prevent contradictory predictions, while EfficientNet-B0 fit free-tier inference constraints.
- Added Grad-CAM heatmaps so every prediction includes the image region driving it; served the model through FastAPI.
- Built an interactive 3D comparison of AlphaFold2 PALB2 structures (wild-type vs. mutant) to complement the classifier's clinical context.

SKILLS

Data & ML: Python, PyTorch, CNNs / EfficientNet, Grad-CAM, FastAPI, model serving
Data Engineering: Airflow/Kestra, dbt, Terraform, BigQuery, GCS
Cloud: AWS (ECS, DynamoDB, S3, Cognito, IAM), GCP (BigQuery, Terraform, Cloud Run, GCS)
Databases: PostgreSQL, MySQL, DynamoDB, Prisma ORM, Supabase
Backend & APIs: Django, NestJS, RESTful API Design, Docker

EDUCATION

Systems Engineering — EIA University *2022 – Present (9th semester)*

LANGUAGES

Spanish — Native | English — B2