

GABRIEL ALVES

Data Engineer

Data Engineer · Software Engineer · Test Automation · AI-Assisted Development

Viana do Castelo, Portugal · profile.gabrieldaes.com · Open to remote (worldwide) · hybrid (North Portugal)

SUMMARY

Computer Engineering graduate and Software Engineering Intern building production data pipelines, backend services, and automated test infrastructure. Comfortable owning a project end-to-end — from an undocumented Django codebase to a merged, tested, production-deployed feature. Engineering priorities: quality, security, optimization, clean code, velocity. Uses AI tools (Copilot, Claude, Cursor) as a reviewed and validated accelerant, never a substitute for understanding code.

SKILLS

- **Data Engineering:** PySpark, Airflow, dbt, ETL/ELT design, star-schema modeling, data quality validation
- **Backend / SWE:** Django, REST APIs, PostgreSQL, JDBC, Git/PR-based review
- **Testing / QA:** Pytest, Playwright (suite architecture + reusable helpers), E2E automation
- **AI / ML Tooling:** Copilot, Claude, local LLMs (dev); GPT, Gemini, Perplexity (research/chat); feature engineering & model evaluation
- **Cloud & DevOps:** AWS (S3, IAM, Cognito), Docker, CI/CD; learning Kubernetes & Terraform
- **Languages:** Python, SQL, Java

EXPERIENCE

Software Engineering Intern | Valuedate.io

Jun 2026 – Present | Viana do Castelo, Portugal

- Fixed a production bug in an inherited, undocumented Django codebase (no DB access), shipped with Pytest coverage.
- Introduced Playwright as the company's first automated E2E testing tool, independently building a full test suite for a production app; approach has since been requested by senior engineering staff for use on a second production project.
- 4 production PRs merged in first 3 weeks, each reviewed and approved by the engineering lead.

PROJECTS

- **World Bank GDP Pipeline | Personal — Data Engineering**

Paginated REST API → S3 (Hive partitioning) → PySpark → PostgreSQL star schema (JDBC) → Airflow → Docker.

- **Customer Churn Prediction Pipeline | Personal — ML/DE**

ML-integrated ETL, feature engineering, profit-based threshold optimization; deployed on AWS with Cognito auth.

- **NLP Sentiment Analysis Pipeline | Personal — Data/AI**

Large-scale text processing pipeline with a Streamlit dashboard.

- **SQL Workforce Analysis | Personal — Data Analytics**

CTE-first analytical SQL surfacing workforce metrics.

EDUCATION

Sep 2020 - Jul 2024

Bachelor Degree in Informatics Engineering

- ESTG - Instituto Politécnico de Viana do Castelo, Portugal

DataCamp: Data Engineer Professional, Data Scientist Professional, Python Data Associate, SQL Associate, AI Engineer for Data Scientists Associate