

Francisco Almeida

📍 Lisbon, Portugal • ✉ franciscoal95@hotmail.com • 🌐 franciscodalmeida.com • 📺 francisci-co • 🎧 francisci-co

Summary

Data Scientist with 4+ years of experience delivering analytics, machine learning, and data driven solutions for global consumer-panel clients. With a background in Bioinformatics, I combine scientific thinking with strong programming skills to build scalable data products, modernise workflows, and translate complex data into meaningful business outcomes. Driven by curiosity, I continuously seek opportunities to learn, innovate, and improve the tools, technologies, and methods behind my work.

Experience

- NielsenIQ — Senior Data Scientist** — Lisbon, Portugal June 2025 – present
- Developed and enhanced products and processes, including fraud monitoring systems, data normalization, and fuzzy matching.
 - Improved data imputation methodologies for consumer panel data completeness and reliability.
 - Delivered ad-hoc analysis and investigations that guide decision-making on complex business problems.
- Kantar Worldpanel — Data Scientist** — Lisbon, Portugal May 2023 – Nov 2024
- Re-engineered legacy SAS products into R, cutting runtime and establishing R development and packaging best practices.
 - Developed and managed the ETL data pipeline for the Central and South American regions powering the global advanced analytics products.
 - Delivered advanced analytics for top global clients, including clustering, segmentation, classification, and time-series.
- Kantar Worldpanel — Data Analyst** — Lisbon, Portugal Apr 2022 – May 2023
- Built predictive models on consumer panel data to uncover brand-growth insights.
 - Produced customer, brand, advertising, and media analytics, defining KPIs behind client brand strategies (MMM).
- i3S — Bioinformatician (Trainee)** — Porto, Portugal Jan 2020 – Jan 2022
- Built NGS pipelines for germline variant calling from WES and WGS data across multiple callers, with clinical applications.
 - Automated bioinformatics workflows on a Unix-based HPC cluster.

Education

- MSc University of Porto** Sept 2019 – Sept 2021
MSc in Bioinformatics and Computational Biology – Porto, Portugal
- Thesis: *Analysis and comparison tool for multiple variant calling results from germline NGS data with clinical applications* — Expression Regulation in Cancer group, i3S.
- BSc University of Trás-os-Montes e Alto Douro** Sept 2016 – July 2019
BSc in Genetics and Biotechnology – Vila Real, Portugal

Technical Skills

Programming: Python, R, SQL, SAS, MATLAB
Platforms: Databricks, Azure
Visualisation: Shiny, Plotly
Data & ML: PySpark, arrow, pandas, NumPy, scikit-learn, PyTorch, caret, tidymodels, MLflow
Tooling: Git, Docker, Bash
Spoken Languages: Portuguese (native), English (professional), Spanish (intermediate)

Publications & Conferences

- Co-author: *Journal of Global Antimicrobial Resistance* ([10.1016/j.jgar.2019.05.017](https://doi.org/10.1016/j.jgar.2019.05.017)) and *European Journal of Clinical Microbiology & Infectious Diseases* ([10.1007/s10096-019-03709-6](https://doi.org/10.1007/s10096-019-03709-6)).
- Multiple poster and oral communications at international conferences in Genetics and Bioinformatics. More info [here](#).

Certifications

Python Essentials for MLOps (3L7CJZVICLOU) / Exploratory Data Analysis for Machine Learning (WZNM54SUKPEC) / Machine Learning with Python (XSCWRJX8FU3W)