

Andrés Menor



Data Engineer · HPC & Cloud Infrastructure

andresmr997@gmail.com · 633 60 78 03 · andresmenor.com · linkedin.com/in/andres-menor · Cataluña / Híbrido

PERFIL

Data Engineer con experiencia en pipelines cloud-nativos, orquestación de flujos de datos y computación distribuida en entornos HPC. Ha diseñado e implementado sistemas RAG en producción sobre GCP, ejecutado cargas masivas en clústeres HPC con SLURM y construido pipelines ETL end-to-end. Sólida base matemática (media 9, honores en Probabilidad, Álgebra Lineal y ML). Busca rol con responsabilidad técnica real en un equipo de datos.

EXPERIENCIA

Data Engineer · Findspo (startup) 04/2026 – Presente

- Diseñé e implementé pipeline completo de ingesta desde PubMed (Python) para plataforma de investigación en longevidad basada en evidencia científica.
- Construí arquitectura RAG: búsqueda semántica con embeddings vectoriales + reranking con LLM para chatbot de respuesta clínica.
- Gestioné toda la infraestructura GCP mediante Terraform: Cloud Storage, BigQuery, Cloud Run y Vertex AI.
- Único ingeniero de datos responsable de la fiabilidad del pipeline, monitorización de calidad y mejora iterativa del sistema de recuperación.

Data Scientist Intern · CSIC – IBMB (Institut de Biologia Molecular de Barcelona) 09/2025 – 12/2025

- Procesé datasets de scRNA-seq a gran escala con Python y scanpy para predicción de edad biológica.
- Entrené y evalué modelos de regresión y clasificación (Elastic Net, XGBoost, Random Forest, SVM); mejor resultado con Logistic Elastic Net.
- Ejecuté todas las cargas computacionales en el clúster HPC del IBMB mediante SLURM.
- Construí workflows de experimentación ML reproducibles con preprocesado, selección de features y evaluación automatizados.

Junior Data & BI Analyst · Grup Boadella 10/2025 – 03/2026

- Automatiqué workflows de reporting SQL y pipelines de preprocesado en Python con validación y control de calidad de datos.
- Diseñé procesos de transformación de datos para soporte a modelos analíticos y KPIs de negocio.

HABILIDADES TÉCNICAS

Lenguajes	Python, SQL (PostgreSQL), R, Java
Data Eng.	ETL/ELT, Airflow, dbt, validación de datos, RAG, búsqueda semántica, reranking
Cloud & Infra	GCP (BigQuery, Cloud Storage, Cloud Run, Vertex AI), Terraform, Docker, Linux, Git
HPC	SLURM, ejecución distribuida en clúster HPC
ML & Análisis	scikit-learn, XGBoost, Elastic Net, Random Forest, Pandas, feature engineering, scRNA-seq

EDUCACIÓN

Grado en Ciencia de Datos — Universitat Oberta de Catalunya (UOC)

2023 – 2027

Media: 9,0 · Honores: Probabilidad y Álgebra Lineal · Matrícula: Machine Learning, Modelado Bayesiano

IDIOMAS

Español – Nativo · Catalán – Nativo · Inglés – Fluido (profesional)