

Luis Núñez

Data Engineer

Barcelona, Spain · luisfernando064@gmail.com · [LinkedIn](#) · [GitHub](#) · [Portfolio](#)

Data Engineer with 3+ years building production data pipelines, data warehouse models, and automated ETL/ELT systems at Amazon across 13 EU countries and 40+ external data sources. Strong SQL and Python applied daily to petabyte-scale datasets on AWS. Experienced across the full data lifecycle — from ingestion and transformation through quality validation to BI delivery. Significant measurable business impact through automated, data-driven systems. Open to permanent and contract roles in Switzerland, Spain, Ireland, and remote EMEA.

TECHNICAL SKILLS

- **SQL** — Advanced querying, query optimization, CTEs, window functions, complex joins, petabyte-scale transformations (Redshift)
 - **Python** — pandas, PySpark, FastAPI, scripting, pipeline automation, Jupyter notebooks
 - **AWS** — S3, Lambda, Redshift, Glue, Athena, DynamoDB, IAM, Secrets Manager
 - **ETL / ELT** — Pipeline design and maintenance, Prefect orchestration, idempotent backfills, incremental loads, event-driven ingestion, API integration
 - **Data Modeling** — Dimensional modeling (Kimball), medallion architecture (bronze / silver / gold), star schema, slowly-changing dimensions
 - **Big Data** — PySpark on AWS Glue, Apache Parquet, partitioned S3 data lakes, distributed processing
 - **Data Quality** — Automated anomaly detection, schema validation, KPI reconciliation, data integrity monitoring
 - **BI & Visualization** — QuickSight, Tableau, Preset / Superset
 - **Orchestration** — Prefect (production), familiar with Airflow concepts
 - **MLOps** — MLflow model tracking and registry, LLM integration (LLaMA 3.3-70B), prompt engineering
 - **Other** — Git/GitHub, CI/CD (GitHub Actions), REST APIs, PostgreSQL, MongoDB, InfluxDB
-

PROFESSIONAL EXPERIENCE

Business Intelligence Engineer — Amazon

Barcelona · Munich · Luxembourg — September 2022 to Present

- Designed and maintained scalable data pipelines integrating data from 40+ external carrier systems into the data warehouse across 13 EU countries.
- Built and optimized data warehouse schemas on Redshift powering analytics and reporting for 40+ stakeholders across 14 countries.
- Wrote production-grade Python and advanced SQL to transform and aggregate petabyte-scale datasets; built an automated SLA-claim pipeline with significant financial recovery impact.
- Implemented data quality frameworks (validation checks, anomaly detection, schema validation) ensuring accuracy and reliability of core financial and operational KPIs.
- Built SQL and Python analytics pipelines identifying underperforming carriers — driving measurable cost savings.
- Optimized complex SQL transformations on **petabyte-scale datasets** — achieved **40% performance improvement** through query rewrites and partitioning strategies.
- Developed automated KPI and volume-tracking frameworks comparing forecasts versus actuals — contributing to network optimization savings.
- Collaborated with data analysts, engineers, finance, and operations teams across multiple countries to translate business requirements into scalable data solutions.

Software Developer (Working Student) — ASTech (AUDI AG)

Ingolstadt, Germany — September 2020 to August 2021

- Developed software components in C++ (ADTF3 framework) for automotive driver-assistance systems.
 - Automated dependency visualization tooling using Python and GraphQL.
-

SELECTED PROJECTS

Full portfolio at luis-fer-333.github.io.

- **Medical Appointment Chatbot (Master's Thesis, 2025)** — Serverless application on AWS Lambda integrating LLaMA 3.3-70B for intent classification, DynamoDB for session management, Google Calendar API, and WhatsApp Cloud API. End-to-end pipeline with structured data validation and cost/latency benchmarking.
 - **Nobel Prize Data Lake** — Medallion-architecture data lake on AWS (S3 + Lambda + Prefect) with raw / bronze / silver ETL layers, scheduled backfill, and idempotent pipeline design.
 - **Amazon Reviews Spark Pipeline** — PySpark MLlib on AWS Glue processing **17M records** into a medallion S3 architecture with feature engineering and batch inference.
 - **Movie Database ETL Pipeline** — Multi-source ingestion (IMDb bulk files, TMDb REST API, MongoDB) into a normalized relational schema with data quality checks and deduplication.
 - **Energy Demand Forecasting** — Continuous time-series ingestion into InfluxDB via AWS Lambda with scheduled orchestration and Prophet-based forecasting.
 - **Airbnb Valencia BI** — SQL dimensional modeling and stakeholder dashboards on Supabase + Preset over 8K+ listings.
-

EDUCATION

M.Sc., Data Science & Cloud Data Engineering

Universidad de Castilla-La Mancha, Spain — 2024 to 2025

Focus: Python, AWS, PySpark, data pipeline design, MLflow, medallion data-lake patterns. Thesis on serverless cloud architectures and LLM integration.

B.Eng., Engineering and Management (Honors)

University of Applied Sciences Ingolstadt, Germany — 2019 to 2023

LANGUAGES

English C1 · Spanish Native · Catalan C1 · German B1 · Portuguese A2

AVAILABILITY

Open to permanent and contract positions. Available for relocation to Switzerland, Ireland, or other EU locations. Currently based in Barcelona, Spain. EU citizen — no work permit required.