

David Zumaquero

✉ zumaquero david@gmail.com | in linkedin.com/in/davidzuma | 🌐 davidzuma

ABOUT ME

I am a proactive data engineer with 4 years of experience building scalable ETL/ELT pipelines, managing large-scale datasets, and designing cloud-native data platforms. Expertise in Python, Azure, Spark, Databricks and orchestration tools like ADF and Airflow.

EXPERIENCE

Data Integration Engineer
EON Grid Solutions (Contractor via Accenture and EnergyAI)

Berlin, Germany
Feb 2024 – Apr 2025

- Designed and implemented scalable ETL pipelines using Databricks and Spark to integrate diverse data sources into a unified energy network data model.
- Developed a GenAI-powered application for automated mapping of source system data to target models, reducing ETL work time by 50%.
- Established standardized health checks for ETL pipelines, including logging, alerts, and automated validation for improved reliability.
- Collaborated with multiple grid operators to deliver high-quality data enrichment and analytics solutions.
- Technologies: Databricks, Delta Lake, Azure Data Factory, Spark, Python, SQL, GitLab CI/CD, Terraform.

Data Engineer
USA Consulting Firm (Contractor via Cathedral Software and Deep Kernel Labs)

Málaga, Spain
Jan 2022 – Jan 2024

- Created a comprehensive energy data lake aggregating information from multiple public sources, including ENTSO-e and National Grid.
- Develop ETL pipelines for electric vehicles adoption and battery degradation forecasting models.
- Designed and implemented backend systems and data processing workflows for a benchmarking platform that compares energy companies across scenarios.
- Technologies: Python, Databricks, Airflow, Kedro, dbt, PySpark, Pandas, Azure, Django, Docker, Kubernetes, GitHub Actions, PostgreSQL, Git.

Data Engineer
PricewaterhouseCoopers (PwC)

Málaga, Spain
May 2021 – Jan 2022

- Developed a Django-based backend system for a supply chain management platform.
- Built scalable ETL pipelines using PySpark on Azure for data integration and analytics.
- Enhanced data mapping integration with NLP, enabling AI-powered assistance for users
- Technologies: Django, PySpark, Pandas, Azure, Jenkins, PostgreSQL, Git, Spacy.

EDUCATION

Masters Degree in Mathematics
Universtity of Málaga, grade 8.3/10

Andalucía, Spain
Sep 2020 - Sep 2022

Bachelor Degree in Mathematics
University of Málaga and Marie Curie-Slodowska University, grade 7.1/10

Málaga, Spain
Sep 2016 – Jul 2020

LICENSES AND CERTIFICATIONS

Data Engineer Associate | Databricks

Jun 2025

Generative AI Engineer Associate | Databricks

Dec 2024

Scrum Master | Scrum Manager

Oct 2021

SKILLS

- Languages and frameworks:** Python, SQL, PySpark, dbt, FastAPI, Django
- Data Platforms:** Databricks, Azure Synapse, Airflow, ADF
- Infrastructure:** Azure, Docker, Kubernetes, GitLab CI/CD, Terraform