

# FRANCISCO MINGUEZ

Independent Data Scientist & ML Engineer

Madrid, Spain | Remote Europe | Available for independent projects

[franminguezg23@gmail.com](mailto:franminguezg23@gmail.com) | [linkedin.com/in/francisco-minguez](https://www.linkedin.com/in/francisco-minguez) | [franciscominguez.dev](https://franciscominguez.dev)

---

## PROFESSIONAL SUMMARY

Independent Data Scientist and ML Engineer based in Madrid, available for remote projects across Europe. Experience building and evaluating RAG systems, machine learning models and data pipelines with Python, PySpark and Databricks, alongside data quality and business analytics work. Background across technology, financial services, insurance and European regulation. Experience collaborating with product, engineering and business teams, including client-facing environments.

## AREAS OF WORK

- RAG design, retrieval evaluation, grounded generation and failure analysis
- Applied machine learning, NLP and model evaluation
- Data and ML pipelines with Python, PySpark and Databricks
- Data-quality assessment, analytics and reporting automation
- External collaboration with European data, product and engineering teams

## PROFESSIONAL EXPERIENCE

Independent Data & AI Practice | Data Scientist & ML Engineer

Jul 2026 - Present | Madrid / Remote Europe

- Building an independent practice focused on applied analytics, data engineering, machine learning and reliable generative AI systems.
- Publishing case studies on RAG reliability and analytical pipeline modernisation at [franciscominguez.dev](https://franciscominguez.dev).
- Developing MigaDigital, a parallel initiative for scoped data, AI and process-automation pilots for Spanish SMEs.

### Multiverse Computing | ML Engineer

Jan 2026 - Jun 2026 | Madrid, Spain | Fixed-term project contract

- Designed and evaluated RAG pipelines, focusing on retrieval quality, context sufficiency, abstention and answer reliability.
- Developed evaluation workflows for LLMs and vision-language models, and worked with supervised fine-tuning, model compression and post-compression performance recovery under confidentiality obligations.
- Took functional Product Owner responsibility for a technical workstream, coordinating priorities and acceptance criteria.
- Presented technical progress and results in sessions with clients and stakeholders while collaborating with engineering, product and research teams.

### Experian | Data Scientist

Sep 2024 - Jan 2026 | Madrid, Spain

- Developed supervised and unsupervised ML workflows for data management, segmentation and automation use cases.
- Migrated and parallelised an income-estimation workflow from SQL Server to Databricks/PySpark, improving its execution and maintainability through reusable Spark components.
- Built reusable PySpark components with automated data quality checks and statistical validation.
- Optimised ETL workflows using PySpark and SQL Server/SSIS, improving execution efficiency and maintainability.
- Developed segmentation workflows using K-Means and DBSCAN, with PCA for dimensionality reduction and analysis.

PROFESSIONAL EXPERIENCE

EIOPA - European Insurance and Occupational Pensions Authority | Data Scientist Trainee

Feb 2024 - Sep 2024 | Frankfurt, Germany

- Built an NLP and OCR pipeline with spaCy and rule-based components to support ESG document review and prioritisation.
- Measured recall and reviewed the workflow's effect on manual document screening.
- Developed predictive models with LightGBM and PyTorch and evaluated them against an established baseline.
- Migrated SAS-based analytical workflows to Python and Azure, improving traceability and maintainability.
- Automated statistical-analysis and anomaly-detection workflows using Isolation Forests and autoencoders.

AXA Seguros | Data Analyst Intern

Sep 2023 - Jan 2024 | Madrid, Spain

- Migrated analytical processes from SAS to PySpark and Databricks.
- Standardised business metrics and introduced more consistent analytical workflows.
- Developed exploratory analyses and KPI visualisations using Python, R and Power BI for business teams.

Banco Sabadell | Business Manager / Insurance Specialist

Feb 2021 - Jan 2023 | Madrid, Spain

- Managed a portfolio of corporate clients across banking and insurance products.
- Used segmentation, reporting and client analysis to support acquisition, follow-up and commercial activity.
- Automated recurring reports and developed experience in client communication, negotiation and regulated financial products.

SELECTED PORTFOLIO WORK

**Analytical Pipeline Modernisation:** Case study on migrating legacy analytical processing to modular PySpark workflows with data quality checks, logging, tests and performance analysis.

**RAG Reliability Study:** Synthetic evaluation study covering retrieval quality, citations, abstention and failure analysis for document-based question answering.

**Live case studies:** [franciscominguez.dev](https://franciscominguez.dev)

TECHNICAL SKILLS

**Programming & Data:** Python, SQL, PySpark, R

**ML & AI:** scikit-learn, LightGBM, PyTorch, NLP, RAG, LLM and VLM evaluation, embeddings, OCR, model compression

**Data Engineering:** Databricks, Spark, ETL, SQL Server, SSIS, data quality checks

**Analytics & BI:** Power BI, Streamlit, Plotly, KPI design, analytical reporting

**Cloud & Delivery:** Azure, Docker, Git, GitLab

**Collaboration:** Client presentations, stakeholder communication, functional product ownership

EDUCATION

**Master's programme in Data Science and Business Analytics, with dual specialisation in Data Science and Business Intelligence:** IMF / UCAV / Indra, 2023 - 2024. Master's thesis: predicting bank deposit subscriptions with classification models and a deep polynomial neural network in PyTorch.

**MBA:** ENEB, 2020 - 2021

**Bachelor's Degree in Marketing, taught in English:** Universidad Rey Juan Carlos, 2015 - 2020

LANGUAGES, CERTIFICATIONS & AVAILABILITY

**Languages:** Spanish - Native | English - C1, TOEIC Listening & Reading: 910/990

**Certifications:** Docker, OpenWebinars (2024) | Scrum / Agile Methodologies, IMF (2023) | Microsoft Data Analysis Fundamentals (2023)

**Availability:** Immediate | Remote Europe | Independent projects | Fixed-term contracts | Consultancy collaboration