

ISMAEL BOVINGDON CASTILLEJO

Junior Data Engineer | Azure Data Factory • SQL • Python • Databricks • Spark

Madrid, Spain • +34 692 871 279 • ishmaelbovingdon@gmail.com

linkedin.com/in/ismael-bovingdon • github.com/I-Bovingdon • Portfolio: i-bovingdon.github.io/DataEngineer.com

PROFESSIONAL SUMMARY

Junior Data Engineer with hands-on experience building ETL pipelines on Azure Data Factory and SQL Server, and a 4-year background in high-volume financial systems (BBVA, DB2/SQL). Currently delivering BI and ETL solutions for financial and industrial clients at Babel, while completing an MSc in Data Science, Big Data & Business Analytics (UCM, September 2026). I designed and operate a 24/7 production ingestion pipeline on Linux with a layered raw to Parquet architecture, processing over 320,000 records per day for a real-time delay detection project on Madrid's commuter rail network.

PROFESSIONAL EXPERIENCE

Junior Data Engineer | Business Intelligence • Babel

Madrid (Hybrid) | Apr 2026 - Present

- Build end-to-end ETL pipelines with Azure Data Factory and SQL Server (SSMS), integrating heterogeneous data sources for financial and industrial clients.
- Develop and maintain around a dozen Power BI dashboards and reports that track operational KPIs and support business decision-making across the company.
- Translate business logic between DAX and SQL and validate the quality and consistency of corporate data.

Technologies: Azure Data Factory, SQL Server, SSMS, Power BI, Python, DAX, Git

Health & Safety Manager • Construction and industrial projects

Nidus, Ibersys, ARPO • Madrid | Jun 2022 - Aug 2025

- Produced technical reports and tracked compliance KPIs to support decision-making in multidisciplinary projects, including a water-supply pipeline renovation for Canal de Isabel II.
- Standardized document control and reporting processes across multiple stakeholders, ensuring information traceability.

Backend & Data Processing Developer, Financial Systems • BBVA Technology

Madrid (Hybrid) | Nov 2017 - Jan 2022

- Developed and maintained batch and online financial applications in business-critical, high-availability environments.
- Queried and processed DB2/SQL databases with millions of records, ensuring the consistency and reliability of financial processes through data validation and analysis.
- Optimized operational processes and resolved incidents in enterprise systems under corporate data-quality and version-control standards.

Technologies: SQL (DB2), JavaScript, COBOL, Control-M

TECHNICAL PROJECTS

24/7 Real-Time Data Pipeline • Madrid Commuter Rail Delay Prediction (MSc Thesis, UCM)

2026 • github.com/I-Bovingdon

- Designed and operate a 24/7 multi-source ingestion pipeline on a Linux VPS (Ubuntu) capturing RENFE GTFS-Realtime feeds and AEMET weather data: over 320,000 records per day.
- Implemented a layered raw (immutable JSON) to processed (daily Parquet) medallion-style architecture, with nightly compaction (cron), automated backups to Google Drive (rclone) and resilience via systemd (auto-restart verified).
- Performed the data-quality audit: 100%-null feed fields, ± 24 h delay anomalies and a join strategy with static GTFS to recover the line identifier.

Technologies: Python, pandas, PyArrow/Parquet, Linux, systemd, cron, rclone, REST APIs, GTFS-RT

Machine Learning Pipeline • Water Infrastructure Maintenance (DrivenData “Pump It Up”)

2026 • github.com/I-Bovingdon

- Built an end-to-end ML pipeline (EDA, feature engineering, target encoding, ensembles) to predict water-pump operational status; Stacking Ensemble scoring 0.8230 with dedicated strategies for a 7.3% minority class.
- Integrated an LLM-based EDA agent (LangGraph/LangChain, Anthropic API) and an interactive Gradio prediction interface.

Technologies: Python, scikit-learn, LangGraph, LangChain, Gradio, Git

EDUCATION

MSc in Data Science, Big Data & Business Analytics • Universidad Complutense de Madrid | 2025 - 2026

BEng in Industrial Mechanical Engineering • Universidad de Huelva | 2011 - 2017

CERTIFICATIONS

- Snowflake University Platform Skills Badge • Snowflake | 2026
- AI Skills Fest 2026 • Microsoft | 2026

SKILLS

Languages: Python (pandas, scikit-learn), SQL (SQL Server, DB2), DAX, JavaScript • Git/GitHub, SSMS, VS Code, Jupyter

Data Engineering: Azure Data Factory, ETL/ELT, Databricks, Apache Spark (PySpark), Snowflake, Data Warehouse, medallion architecture, Parquet, Linux, systemd, orchestration (cron, Control-M)

BI & Analytics: Power BI, data modeling, reporting, data visualization

ML & LLM: scikit-learn, TensorFlow/Keras, feature engineering, ensembles, LangChain/LangGraph, Anthropic API, LLM agents (EDA/RAG), Gradio

MLOps & DevOps: Git/GitHub, Linux, systemd, cron, process automation (Python/Bash), monitoring and resilience, model serving (REST API), model lifecycle and productionization

LANGUAGES

Spanish: Native • English: B2+ (Professional Working Proficiency)