

# Iker Aguirre Etxebarria

## Ingeniero de Datos Junior

Graduado en Ingeniería Informática por la Universidad del País Vasco (UPV/EHU) con nota media de **8,3/10**. Realicé **prácticas de 6 meses en BBVA** (equipo de data engineering) donde diseñé **4 pipelines ETL** en producción que procesan **12 millones de registros diarios**. Experiencia con Python, SQL, Apache Spark, Airflow y AWS. Certificación AWS Cloud Practitioner.

### HABILIDADES

Python (pandas, PySpark, scripting ETL) • SQL avanzado (PostgreSQL, Amazon Redshift) • Apache Spark y PySpark • Apache Airflow (orquestración de pipelines) • AWS (S3, Redshift, Glue, Lambda) • dbt (data build tool) • Docker y contenedores • Git y control de versiones • Great Expectations (calidad del dato) • Modelado dimensional (Kimball)

### IDIOMAS

Español (nativo) • Euskera (B2 - EGA) • Inglés (C1 - Cambridge Advanced)

### EXPERIENCIA

Becario en Data Engineering, BBVA, Bilbao Ene 2025 - Jun 2025  
Incorporación al equipo de ingeniería de datos dentro de la unidad de Data & Analytics, responsable de la infraestructura de datos para banca digital.

- Diseñé e implementé **4 pipelines ETL** con Apache Airflow y Python que procesan **12 millones de registros diarios** desde fuentes transaccionales hacia el data lake
- Optimicé **8 consultas SQL** en Amazon Redshift que redujeron los tiempos de carga del reporting en un **35%**
- Desarrollé **3 módulos de validación de calidad del dato** con Great Expectations, detectando y corrigiendo anomalías en un **2,3% de los registros** procesados
- Documenté la arquitectura de **6 flujos de datos** en Confluence, facilitando el onboarding de nuevos miembros del equipo

### FORMACIÓN

Título y especialidad, Universidad del País Vasco (UPV/EHU), Bilbao Sept 2021 - Jun 2025  
Nota media: **8,3/10**. TFG sobre optimización de consultas distribuidas en Apache Spark para procesamiento de datos financieros, calificado con **Sobresaliente (9,2)**.

- Especialización en Ingeniería del Software y Sistemas de Información
- Proyecto de asignatura: diseño de un data warehouse para análisis de movilidad urbana con **5 millones de registros** del Ayuntamiento de Bilbao

### CERTIFICACIONES

AWS Certified Cloud Practitioner, Amazon Web Services Oct 2024 - Oct 2024

### ACTIVIDAD EXTRACURRICULAR

Colaborador - Grupo de Investigación en Big Data UPV/EHU Ene 2024 - Jun 2025  
Participación en proyecto de investigación sobre procesamiento de datos de movilidad urbana.

- Contribuí al desarrollo de un pipeline de procesamiento de **datos GPS de autobuses urbanos** utilizado en **2 publicaciones académicas** del grupo