

Targon, Juan Cruz

Data Engineer | Information Systems Engineer

Barcelona, España - Cel: +34 600551568 - [linkedin.com/in/juan-cruz-targon](https://www.linkedin.com/in/juan-cruz-targon)

Mail: juantargon@gmail.com

Data Engineer with experience in data ingestion, modeling, and processing within big data environments. Former member of the core data ingestion and modeling team at Santander Tecnología, collaborating with engineering and business teams to ensure data availability, consistency, and quality. Previously worked at Disbyte (E-commerce) as a Data Engineer, developing scalable data solutions in AWS. Passionate about continuous improvement, process optimization, and solving complex analytical challenges. Currently pursuing a Master's in Economics, Finance, and Computation at the Universidad Internacional de Andalucía (UNIA), focused on Big Data, Economics, and Artificial Intelligence. Graduated in Information Systems Engineering from UTN-FRC.

Experience

- **Data Engineer-Datalytics (client Santander Tecnología) [05/2023] - [05/2025]** : member of the main data ingestion team towards datalake. Responsible for the incorporation of data from multiple sources, modeling and design of tables according to business requirements. Optimization of loading processes and quality assurance and availability of information. Active participation in the redefinition of data domains according to current standards. Applied technologies: PySpark, Python, SQL, PL/SQL, NoSQL, API, AWS, Airflow, Hive, Git, Jira.

Achievements

- Redesign the process of ingesting data from multiple sources into Santander's datalake, reducing load times by 30% through optimization with PySpark and Airflow.
- Model data for consumption by business users, based on good modeling practices and domain standardization, facilitating traceability and data analysis.
- I collaborated in continuous improvement initiatives in ETL processes and data quality validations, proposing automations that increased the efficiency of the team.

- **Data Engineer – Disbyte (E-commerce) [05/2025] – [10/2025]:** Member of the data engineering team responsible for the design, development, and implementation of data ingestion and processing pipelines from multiple sources. Actively participated in defining the data architecture and modeling structures for both analytical and operational use. Implemented scalable solutions in AWS and optimized ETL/ELT processes with a strong focus on efficiency, reliability, and data quality.

Technologies: AWS (*Lambda, Redshift, S3, RDS, Aurora, API Gateway, SNS, SQS, IAM*), *Airflow, Python, PL/SQL, Spark, Pandas, Polars, Git, Jira*.

Achievements

- Designed and implemented data ingestion pipelines from various APIs, including major marketplaces (*Amazon, Mercado Libre*) and external logistics services, integrating data into transactional systems, the Data Lake, and *Redshift*.
- Defined data models for business consumption, ensuring consistency and traceability throughout the entire data lifecycle, from capture to analytics.
- Proposed and implemented architectural improvements using *Spark, Pandas, and Polars*, reducing processing time and infrastructure costs.
- Collaborated closely with business teams to gather requirements and validate deliverables, ensuring alignment between analytical needs and technical solutions.

Education & Certification

- Information Systems Engineering - UTN - FRC [2019-2024] - Average Note: 8.38 - GPA: 3.7
Master's in Economics, Finance and Computation – Universidad Internacional de Andalucía (UNIA) [11/2025 – Present].
Focus on *Big Data, Economics, and Artificial Intelligence*, combining economic and financial foundations with advanced data analysis and computational tools.
- AWS Cloud Practitioner [2024].
- AWS Data Engineer [2025-current].
- Apache Airflow Fundamentals - Astronomer [2024].
- Git and Github Professional Course [2023].
- Tableau and data visualization course [2022].

Languages

- English – Advanced (B2, TOEIC).
- Spanish - Native.

Technologies & tools

- **Programming Languages:** Python (scripting, automation, and data analysis), SQL and PL/SQL (complex queries, stored procedures, and optimization in relational environments).
- **Big Data & Distributed Processing:** PySpark and Apache Airflow for developing and orchestrating data pipelines; Hive for large-scale querying and Data Lake modeling.
- **Cloud Computing & APIs:** Advanced experience with AWS (*Lambda, Redshift, S3, RDS, Aurora, API Gateway, SNS, SQS, IAM*); design and implementation of scalable and secure architectures; **Postman** for API development and testing.
- **Databases:** Work with relational databases (*PostgreSQL, Oracle*) and NoSQL (*MongoDB, DynamoDB*); data modeling, query optimization, and integration across heterogeneous sources.
- **Version Control & Project Management:** Git, GitLab, Jira; applying agile methodologies in data engineering and analytics projects.