

NICOLÁS ROJAS DÍAZ

Computer Science Engineer | Data Engineering | Software Development

+56 9 4934 3967 | nicord2002@gmail.com

<https://www.linkedin.com/in/nicolas-rojas> | <https://github.com/Niicolas-Rojas> | <https://niicolas-rojas.github.io/>

PROFESSIONAL SUMMARY

Computer Science Engineer and Software Developer focused on Data Engineering. Professional experience in backend development, REST APIs, PostgreSQL, CI/CD with GitHub Actions, and cloud infrastructure, complemented by experience in Data Governance, SQL Server, relational data modeling, and data quality.

I build end-to-end Data Engineering projects using Azure Data Factory, Azure Data Lake Storage Gen2, Azure Databricks, Apache Spark, PySpark, Delta Lake, and Unity Catalog, working with historical and incremental ingestion, Medallion Architecture, idempotency, Data Quality, orchestration, security, automation, and production deployment practices.

TECHNICAL SKILLS

Data Engineering: SQL, Python, PySpark, Apache Spark, Spark SQL, ETL/ELT, Medallion Architecture, Data Lake, Lakehouse, Data Modeling, Data Quality, and Data Governance.

Azure: Azure Data Factory, Azure Data Lake Storage Gen2, Azure Key Vault, Managed Identities, and Azure RBAC.

Databricks: Azure Databricks, Delta Lake, Unity Catalog, and Databricks Jobs.

Databases: PostgreSQL and Microsoft SQL Server.

Software Development & CI/CD: Go, REST APIs, Git, GitHub, GitHub Actions, CI/CD, Docker, Linux, AWS Lightsail, Cloudflare Workers, and Scrum.

Analytics & AI: Power BI, Tableau, GPT/LLMs, and Prompt Design.

PROFESSIONAL EXPERIENCE

Software Developer

12/2025 - Present

Nissi Advisor S.P.A.

- Develop and maintain a production platform for gym operations management using Go, PostgreSQL, and REST APIs.
- Design and maintain PostgreSQL data models and automate processes used in the platform's daily operations.
- Design and maintain CI/CD pipelines with GitHub Actions for a Go backend and Next.js frontend, including build validation on pushes and Pull Requests and automated production deployments.
- Implement CI/CD workflows with build validation, artifacts, post-deployment health checks, and rollback mechanisms to reduce deployment risk.
- Automate backend and frontend deployments to cloud infrastructure using GitHub Actions, AWS, and Cloudflare, securely managing credentials and configuration variables.

Professional Internship – Data Engineering

11/2024 - 01/2025

Universidad Católica de Temuco – Data Governance Office

- Conducted a feasibility study to assess the potential adoption of a Data Lake, comparing architecture alternatives, costs, and technical considerations across AWS, Azure, Google Cloud, and open-source technologies.
- Built a local proof of concept using Docker, MinIO, Trino, Hive Metastore, PostgreSQL, and Metabase to validate storage, metadata management, and SQL querying.
- Designed a 13-table relational database in Microsoft SQL Server, including the data model, SQL queries, data dictionary, and data-loading documentation.
- Worked under Scrum methodology and used AI-assisted text extraction followed by manual validation, correction, and classification; subsequently developed dashboards in Tableau.

Professional Internship – Data Analyst

01/2024 - 06/2024

Universidad Católica de Temuco – Data Governance Office

- Analyzed and validated institutional data used for reporting and operational monitoring.
- Developed dashboards and reports using Power BI and Tableau to support analysis and KPI monitoring.
- Used SQL to query and prepare data for analysis and visualization.

DATA ENGINEERING PROJECTS

ChileCompra Data Platform

2026

<https://github.com/Niicolás-Rojas/chilecompra-data-platform>

Azure Data Factory | ADLS Gen2 | Azure Databricks | PySpark | Spark SQL | Delta Lake | Unity Catalog

- Designed and built an end-to-end batch data platform on Azure to process more than **1.1 million purchase orders** and **2.8 million line items** from Mercado Público / ChileCompra, combining historical CSV ingestion with daily incremental ingestion from REST APIs.
- Implemented a Medallion Architecture **Landing** → **Bronze** → **Silver** → **Gold** and automated the daily pipeline using Azure Data Factory and Databricks Jobs; the purchase order and tender branches converge before five independent Gold transformations executed in parallel.
- Implemented idempotent and recoverable processing using `process_date`, `run_id`, `replaceWhere`, Delta MERGE, and freshness rules, supporting reruns and backfills without generating duplicates or overwriting newer versions.
- Added critical pre-write and post-write Data Quality controls for keys, duplicates, statuses, and referential integrity; critical inconsistencies stop the task and propagate failures from Databricks to Azure Data Factory.
- Configured service-to-service security using Azure Key Vault, Managed Identity, Azure RBAC, Databricks Access Connector, and Unity Catalog, and automated the refresh of a Databricks AI/BI dashboard only after all Gold tasks complete successfully.

Olist E-Commerce Lakehouse

2026

<https://github.com/Niicolás-Rojas/Brazilian-E-Commerce>

Databricks | Apache Spark | PySpark | Spark SQL | Delta Lake | Unity Catalog

- Built a Lakehouse pipeline **Raw** → **Bronze** → **Silver** → **Gold**, processing CSV files into analytical tables stored in Delta Lake.
- Implemented data cleaning, type casting, standardization, deduplication, quality controls, referential integrity, and temporal consistency checks.
- Worked with MERGE, Time Travel, RESTORE, OPTIMIZE, ZORDER, and VACUUM to validate Delta Lake capabilities.
- Analyzed 99,441 orders and identified an association between delivery delays and lower customer satisfaction: an average rating of 2.27 for delayed orders versus 4.29 for non-delayed orders.

EDUCATION

Civil Engineering in Computer Science – Information Technologies Concentration

03/2020 - 08/2025

Universidad Católica de Temuco

Capstone Project: Generative AI applied to preventive education.

Designed and validated a library of seven custom GPTs using ChatGPT-4, structured prompts, requirements analysis, and quality criteria. Evaluated results with 11 participants through surveys conducted before and after using the tool.

ADDITIONAL TRAINING

Mining Industry

- Strategic Planning for the Mining Business.
- Mineral Economics.
- Underground Mining Planning.

Data & Artificial Intelligence

- Google Data Analytics Professional Certificate – Google, 07/2026.
- Google AI Essentials – Google, 06/2026.