

Gabriel Franco Josino da Silva

Data Engineer | PySpark | SQL | Databricks | Delta Lake | Airflow | Financial Data Pipelines

Sao Paulo, Brazil | Open to full remote data engineering roles
+55 11 95881-4398 | ds.gabrielfranco@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Data Engineer with 2+ years of experience in banking data environments, working with SQL, Spark SQL, PySpark, Databricks, Delta Lake, Python automation, batch orchestration concepts, schema validation, and business-facing data delivery. Experienced in translating financial business rules for fixed-income products and trade finance into maintainable data transformations, supporting on-premise to cloud modernization, and using config-driven/metadata-driven patterns for financial parameterization and legacy-rule modernization. Practical exposure to Git, branches, pull requests, GitHub Actions CI validation, and controlled deployments across environments. Currently expanding Control-M batch operations experience into Apache Airflow through certification study and a public orchestration project, while also strengthening GitHub Actions, CI/CD, and cloud data engineering for full remote and international Data Engineering roles.

CORE SKILLS

Languages	SQL, Spark SQL, Python, PySpark
Data platforms	Azure Databricks, Delta Lake, Oracle/Exadata, Power BI
Operations	Airflow fundamentals, Control-M orchestration/monitoring, Git, GitHub, GitHub Actions, CI/CD fundamentals, Jira, Confluence
Data engineering	Batch pipelines, ETL/ELT, data validation, schema validation, financial parameterization, config-driven/metadata-driven processing, documentation, automation
Learning focus	Apache Airflow Fundamentals certification, GitHub Actions, Azure Databricks certification path, Spark performance

PROFESSIONAL EXPERIENCE

Data Engineer - BRQ Digital Solutions	Oct 2024 - Present Sao Paulo, Brazil Hybrid
Allocated to F1rst/Santander Financial MIS	
- Develop and maintain PySpark data pipelines in Databricks for financial data processing and internal business consumption. - Create and maintain SQL business rules for fixed-income products and trade finance, including CRA, CRI, debentures, data treatment, transformation, validation, and parameterization. - Build Python automations for SQL parameterization and validation using config-driven/metadata-driven patterns, reducing repetitive manual work and operational risk.	

- Support modernization of legacy financial rules by mapping sources, fields, SQL dependencies, and validation/parity strategies.
- Validate schemas and troubleshoot Spark/Databricks pipeline issues, including data types, write behavior, and consistency between DataFrames and target tables.
- Support modernization of Spark data processes from on-premise environments to Azure Databricks and cloud lakehouse platforms using Delta Lake.
- Work with Oracle/Exadata integrations and maintain data processing routines connected to financial calculation drivers.
- Maintain, monitor, and troubleshoot batch jobs with Control-M, including job status tracking, log analysis, dependency checks, FileWatcher routines, and operational support.
- Map batch orchestration concepts from Control-M to Apache Airflow patterns such as DAGs, sensors, retries, dependencies, logs, and success markers for portfolio development.
- Work with corporate versioning and delivery flows using Git, branches, pull requests, GitHub Actions CI validation, and controlled deployments across environments.
- Collaborate with POs, backoffice teams, and business stakeholders to define, validate, and document data rules.

Data Analyst Intern - Santander Brasil

Oct 2023 - Oct 2024 | Sao Paulo, Brazil | Hybrid

Insurance and Assistance Platform

- Queried, treated, and joined datasets in Databricks using SQL and PySpark across silver and gold data layers.
- Built Power BI dashboards connected to Databricks to support insurance business indicators and management reporting.
- Developed Python automations for text and Excel file processing, reducing manual preparation before dashboard consumption.
- Cleaned, standardized, and normalized datasets involving products, regions, payers, debtors, and credit analysis.
- Supported internal analytics requests using Databricks, on-premise query environments, Jira, and Pipefy.

SELECTED WORK**Financial Data Pipelines and Business Rules**

Development and maintenance of SQL and PySpark routines for financial data treatment, parameterization, validation, and delivery to internal teams in a banking environment, including fixed-income products, trade finance, config-driven rule processing, schema validation, and traceability improvements.

Technologies: SQL, Spark SQL, PySpark, Databricks, Delta Lake, Python, Oracle/Exadata, Control-M

Airflow Orchestration Transition - Portfolio Track

Study and project track to translate production batch concepts from Control-M into Apache Airflow, including DAG authoring, sensors, retries, task dependencies, monitoring, and failure handling.

Technologies: Apache Airflow, Docker, Python, GitHub Actions, CI/CD, data validation

Certification Simulator App - Portfolio Project

Active prototype for certification practice, built with Next.js, TypeScript, Tailwind, and Supabase. The next planned step is to add analytics around attempts, weak topics, learning progress, and data modeling.

Technologies: TypeScript, Next.js, Supabase/PostgreSQL, product analytics roadmap

EDUCATION

FIAP | Technologist Degree in Data Science | Feb 2023 - Dec 2024

Practical education focused on Python, SQL, statistics, machine learning, data analysis, modeling, and pipelines.

CERTIFICATIONS

- Databricks Academy Accreditation - Get Started with Databricks for Data Engineering
- Databricks Academy Accreditation - Databricks Fundamentals
- PCP Federado - Low Platform - F1rst
- In progress: Databricks Certified Data Engineer Associate study path
- Planned: AZ-900, Microsoft DP-750, GitHub Actions/CI-CD, Apache Airflow Fundamentals

LANGUAGES

- Portuguese: Native
- English: Intermediate, actively improving for remote and international data engineering work