

Gabriel Franco Josino da Silva

São Paulo, SP, Brasil · +55 11 95881-4398 · ds.gabrielfranco@gmail.com · [LinkedIn](#) · [GitHub](#) · [Portfolio](#)

Engenheiro de Dados com 2+ anos de experiência no setor financeiro, atuando com SQL, Spark SQL, Python, PySpark, Databricks, Delta Lake, automações, conceitos de orquestração batch, validação de schemas e documentação técnica. Experiência em regras de negócio para produtos de renda fixa e trade finance, parametrizações financeiras config-driven/metadata-driven, migração de rotinas on-premise para cloud e modernização de regras legadas para camadas mais rastreáveis, testáveis e reutilizáveis. Vivência com Git, branches, pull requests, validações de CI com GitHub Actions e deploy controlado entre ambientes. Atualmente expandindo a base prática em Control-M para Apache Airflow por meio de certificação e projeto público de orquestração.

COMPETENCIAS

Linguagens	Python, SQL, Spark SQL, PySpark
Stack de dados	Databricks, Delta Lake, Oracle/Exadata, Power BI
Operacao	Apache Airflow em estudo/projeto, Control-M, Git, GitHub, GitHub Actions, fundamentos de CI/CD, Jira, Confluence
Focos	Pipelines batch, validação de dados, validação de schemas, parametrização financeira, arquitetura Lakehouse, documentação, automação

EXPERIENCIA PROFISSIONAL

Data Engineer · BRQ Digital Solutions

Out 2024 - Atual · São Paulo, Brasil · Híbrido

Alocado na F1rst/Santander · MIS Financeiro

- Criação e manutenção de regras de negócio em SQL para produtos de renda fixa (CRA, CRI, Debêntures) e trade finance.
- Desenvolvimento de automações em Python para montagem e validação de parametrizações SQL com abordagem config-driven/metadata-driven.
- Atuação em modernização de regras financeiras legadas, mapeando fontes, campos, dependências SQL e estratégias de validação/paridade.
- Validação de schemas e diagnóstico de problemas em pipelines Spark/Databricks, incluindo análise de tipos, escrita e consistência entre DataFrame e tabela destino.
- Refatoramento e migração de rotinas Spark de ambiente on-premise para Azure Databricks, utilizando Python/PySpark e Delta Lake.

- Manutenção e monitoramento de pipelines batch no Control-M, incluindo análise de logs, troubleshooting, dependências e FileWatchers.
- Mapeamento de conceitos de orquestração batch do Control-M para padrões de Apache Airflow, incluindo DAGs, sensores, retries, dependências, logs e sinalização de sucesso.
- Integração com Oracle/Exadata e rotinas relacionadas a drivers de cálculo financeiro.
- Atuação com fluxo corporativo de versionamento e entrega utilizando Git, branches, pull requests, validações de CI com GitHub Actions e deploy controlado entre ambientes.
- Documentação técnica e interação com POs, backoffice e áreas de negócio para definição e validação de regras.

Data Analyst Intern · Santander Brasil

Out 2023 - Out 2024 · São Paulo, Brasil · Híbrido

Plataforma de Seguros e Assistências

- Consulta, tratamento e cruzamento de dados em Databricks utilizando SQL e PySpark em camadas silver e gold.
- Desenvolvimento de automações em Python para tratamento e preparação de arquivos texto e Excel.
- Construção de dashboards em Power BI para acompanhamento de indicadores de negócio.
- Padronização, limpeza e normalização de bases envolvendo produtos, regiões, pagadores, devedores e análise de crédito.
- Apoio a áreas internas com análises de dados, relatórios e estudos gerenciais.

PROJETOS RELEVANTES

Pipelines Financeiros e Regras de Negócio

Desenvolvimento e manutenção de rotinas SQL e PySpark para tratamento, parametrização, validação e disponibilização de dados financeiros para consumo por áreas internas, com foco em rastreabilidade, paridade e redução de complexidade em regras legadas.

Tecnologias: SQL, Spark SQL, PySpark, Databricks, Delta Lake, Python, Control-M

Transição para Orquestração com Airflow

Trilha de estudo e projeto para traduzir conceitos de orquestração batch usados em ambiente corporativo para Apache Airflow, com DAGs, sensores, retries, dependências, logs e tratamento de falhas.

Tecnologias: Apache Airflow, Docker, Python, GitHub Actions, validação de dados

Modernização e Manutenção de Processos de Dados

Atuação em ajustes, refatorações, versionamento, documentação técnica e melhoria de processos para aumentar confiabilidade, rastreabilidade e facilitar sustentação.

Tecnologias: Python, SQL, PySpark, Databricks, Git, GitHub, GitHub Actions, CI/CD, Confluence

FORMACAO

FIAP · Tecnólogo em Data Science · Fev 2023 - Dez 2024

Formação com foco em Python, SQL, estatística, machine learning, análise de dados, modelagem e pipelines.

CERTIFICACOES

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

IDIOMAS

- Português: nativo
- Inglês: intermediário, em evolução para entrevistas e trabalho remoto internacional