SAMUEL DE ANDRADE MAIA

Engenheiro de Dados Pleno na Avanade | Analytics Engineering | Governança e Qualidade de Dados

Fortaleza/CE | +55 85 98409-6353 | samuelmaia.carreira@gmail.com | Pessoa com Deficiência (PcD)
LinkedIn | GitHub | Portfólio

RESUMO PROFISSIONAL

Engenheiro de Dados Pleno na **Avanade**, com trajetória em operações, BI e Analytics Engineering. Estruturo pipelines e produtos analíticos com Python, SQL, dbt e práticas de qualidade, governança e automação. No portfólio pessoal, desenvolvi o **Governed Analytics Platform**, que integra contratos de dados, publicação controlada, rastreabilidade e uma interface executiva em Streamlit. A validação local documentada em 13/11/2026 registrou **1.353 testes aprovados e 92% de cobertura (src + app)**. Experiência complementar em Lakehouse com PySpark/Delta Lake, soluções analíticas e integração de APIs.

COMPETÊNCIAS TÉCNICAS

Engenharia de dados: Python, SQL, PySpark, Apache Spark, Delta Lake, ETL/ELT, arquitetura Lakehouse; conhecimentos em Azure Databricks.

Analytics Engineering e BI: dbt, DuckDB, PostgreSQL, SQL Server, modelagem dimensional, Power BI, DAX, Power Query.

Governança e qualidade: Data Quality, Data Contracts, Data Lineage, Publication Gate, reconciliação, testes e controles de privacidade.

IA e automação: RAG com TF-IDF local e citações, OpenAI API opcional, JSON estruturado, pandas, NumPy, Streamlit, FastAPI, REST, GraphQL e n8n.

Cloud e desenvolvimento: AWS, Git, GitHub Actions, CI/CD, pytest, Ruff e mypy.

EXPERIÊNCIA PROFISSIONAL

Avanade | Analyst, Data Engineering — Engenheiro de Dados Pleno

Remoto | out/2026 – atual

- Atuação em soluções de dados, da ingestão e transformação à modelagem e disponibilização analítica, com foco em confiabilidade e manutenção.
- Aplicação de conhecimentos em SQL, Python, PySpark, arquitetura Lakehouse, Azure Databricks, qualidade, versionamento e documentação técnica.

Samuel Maia Analytics | Analytics Engineer / Data & BI Consultant

Fortaleza/CE · remoto | ago/2022 – set/2026

- Desenvolvi pipelines, modelos dimensionais e aplicações analíticas com Python, SQL, dbt, DuckDB, PostgreSQL, Power BI e Streamlit para organizar informações e apoiar decisões.
- Estruturei validações, contratos e testes de dados; desenvolvi integração de IA para extração de atributos de produtos em JSON estruturado.

Piauí Consult — Grupo Boticário (Regional do Piauí) | Assistente de TI

Picos/PI | mar/2019 – mar/2021

- Implantei OCS Inventory, reduzindo em cerca de 80% o tempo de localização e conferência do inventário; integrei informações de PDV, ERP e inventário, reduzindo em cerca de 30% o esforço de consolidação manual.
- Desenvolvi automações em Python e SQL e dashboards em Power BI, Excel e Power Query para acompanhamento comercial e operacional.

Via Bebidas LTDA — Operação Ambev | Assistente Financeiro

Picos/PI | jul/2018 – dez/2018

- Consolidei informações financeiras e comerciais em Excel e sistemas corporativos, reduzindo em cerca de 20% o tempo de consolidação; elaborei relatórios de metas, variações e conferência de registros.

i9life | Analista de Dados

Picos e região/PI | jun/2015 – jun/2018

- Estruturei indicadores comerciais e tratei dados com SQL, Excel, Python e pandas; automatizei rotinas de reporting, reduzindo em cerca de 30% o esforço manual, e conduzi reuniões periódicas de acompanhamento de metas.

Banco Santander | Assistente de Atendimento PF/PJ

Manaus/AM | set/2009 – mar/2014

- Atuei em atendimento, operações e backoffice, acompanhando solicitações, documentação e indicadores. Recebi dois certificados de reconhecimento pela qualidade do atendimento.

SAMUEL DE ANDRADE MAIA

Projetos pessoais, formação e qualificações

PROJETOS EM DESTAQUE

Governed Analytics Platform | 2026

Para enfrentar inconsistências na publicação de dados analíticos, construí uma plataforma com **Python, SQL, dbt e DuckDB**, camadas Bronze/Silver/Gold/Quarantine, Data Contracts, Data Quality, Data Lineage e **Publication Gate** para aprovar, revisar ou bloquear publicações. A nova **Executive Overview em Streamlit** reúne indicadores e navegação por áreas; a solução também inclui FastAPI com autenticação Bearer e assistente RAG com TF-IDF local, contratos aprovados e citações. **Validação local de 13/11/2026: 1.353 testes aprovados, 92% de cobertura em src + app e zero falhas.** Interface modernizada e documentada em 14/11/2026, com checks de CI e deploy aprovados. Projeto demonstrativo; arquitetura Azure é proposta, não implantação corporativa.

Código e documentação | Demonstração Streamlit

Azure Databricks Governed Lakehouse | 2026

Implementei pipeline Lakehouse em **Python, PySpark e Delta Lake** com camadas Raw/Bronze/Silver/Gold/Quarantine, schemas explícitos, regras de qualidade, reconciliação e Publication Gate. A execução local documentada produziu **seis datasets Silver e quatro tabelas Gold, com 186 testes automatizados**. Azure Databricks é a plataforma-alvo; implantação em nuvem não comprovada.

github.com/samuelmaia-analytics/azure-databricks-governed-lakehouse

Central de Automação e Operações | 2026

Desenvolvi uma central em **Python, SQL, DuckDB e Streamlit**, integrada ao Pipefy por API/GraphQL, para consolidar SLA, backlog, alertas e criticidade, apoiando a identificação de demandas vencidas ou sem responsável. Permite demonstração com dados de exemplo quando não há credenciais.

github.com/samuelmaia-analytics/central-automacao-operacoes

Contador de Acessos Serverless na AWS | 2026

Projeto em grupo do AWS re/Start: frontend estático e função Lambda em Python com incremento atômico em DynamoDB. Arquitetura documentada para API, segurança, observabilidade e custos; endpoint real e infraestrutura como código ainda pendentes no código publicado.

github.com/samuelmaia-analytics/aws-serverless-access-counter

FORMAÇÃO ACADÊMICA

Pós-graduação em Inovação e Transformação Digital — Gran Faculdade | out/2026

Pós-graduação em Ciência de Dados — Gran Faculdade | abr/2026

Bacharelado em Administração — FAEL | jul/2024

FORMAÇÃO COMPLEMENTAR

Databricks Data Engineer Associate: curso preparatório — Udemy, 13,5 h | out/2026 (não é certificação profissional)

AWS re/Start, Inteligência Artificial e No Code — Escola da Nuvem, 270 h | jul/2026

Python do Básico ao Avançado — Udemy / Geek University, 64 h | abr/2026

dbt: Crie e Execute Pipelines de Engenharia de Dados — Udemy, 4,5 h | abr/2026

Power BI Completo — Udemy, 9 h | set/2025

IDIOMAS

Inglês intermediário: boa leitura técnica; conversação em desenvolvimento.