

Franklin Ferreira

Staff Data Engineer

Brasil · Remoto/Global

Arquitetura de Dados · Big Data · Governança · FinOps · ML e IA Generativa

franklins390@gmail.com | linkedin.com/in/franklins390 | github.com/franklin390 | franklin390.github.io

RESUMO

Staff Data Engineer e arquiteto de dados. Desenho a arquitetura de dados em que uma empresa inteira se apoia, do warehouse ao lakehouse, do batch ao streaming, do ML clássico à IA generativa: mais de 50% menos custo e ~3x menos incidentes. Continuo mão na massa, com um time técnico de 15+ pessoas entre o Brasil e a Romênia.

>50%	~3x	~50	10+	100+	15+
menos custo da plataforma de dados	menos incidentes e retrabalho	fontes unificadas em uma só plataforma	domínios do Unity Catalog	pipelines Airflow na Clicksign	pessoas lideradas

TRAJETÓRIA

out/2020	jun/2021	out/2021	ago/2024	jul/2025	jul/2026
Data Scientist	Data Architect	Data Team Lead	Data Engineer	Data Engineer – Tech Lead	Staff Data Engineer
Clicksign	Clicksign	Clicksign	Flutter Brazil	Flutter Brazil	Flutter Brazil

COMPETÊNCIAS

Arquitetura e plataformas: Modelagem dimensional, modelagem de dados, data warehouse, data lake, lakehouse, Databricks, Delta Lake, Apache Spark, Databricks Workflows, Amazon Redshift, streaming

Engenharia: Python, PySpark, SQL, dbt, Apache Airflow, pipelines de dados, AWS, Terraform, Docker, Kubernetes, GitHub Actions, CI/CD

Governança e confiabilidade: Governança de dados, catálogos de dados (Unity Catalog, AWS Glue Data Catalog, Hive Metastore), qualidade de dados, LGPD, observabilidade, FinOps

Apps de dados e IA: Streamlit, Metabase, Google Analytics 4, machine learning, IA generativa (Lovable)

Liderança: Liderança técnica de times, arquitetura e formas de trabalho, definição de prioridades

EXPERIÊNCIA

Flutter Brazil	ago/2024 – Atual
Staff Data Engineer	jul/2026 – Atual
Plataforma de dados Databricks usada por milhares de pessoas em toda a empresa, com um time de 15+ pessoas (engenharia, analytics e ciência de dados) entre o Brasil e a Romênia.	
- Defino a arquitetura e as formas de trabalho de um time de 15+ pessoas (10 como Tech Lead). - Continuo mão na massa em arquitetura, modelagem e decisões de plataforma. Quando o time precisa, defino prioridades e destravo decisões de negócio.	
Tecnologias: Databricks, Unity Catalog, Delta Lake, Apache Spark, dbt, SQL, Python, AWS, Terraform, GitHub Actions	
Data Engineer – Tech Lead	jul/2025 – jun/2026
Liderei um time de 10 pessoas (engenharia, analytics e ciência de dados) em uma plataforma Databricks de dezenas de TB.	
- Reduzi em mais de 50% o custo da plataforma de dados com uma iniciativa de FinOps (clusters, jobs, armazenamento, governança) e modelos de previsão de custo. Gasto anual: seis dígitos (USD). - Reduzi incidentes e retrabalho em ~3x ao implantar CI/CD com trunk-based development. - Entreguei dashboards de apoio à decisão, observabilidade de toda a plataforma e monitoramento de métricas de negócio quase em tempo real (latência de poucos minutos), e reestruturei o Unity Catalog por domínio (10+). - Defini as arquiteturas e as formas de trabalho de quase todo o time de dados. Ajudei a criar os modelos de LTV e payback da empresa.	
Tecnologias: Databricks, Unity Catalog, Delta Lake, Apache Spark, dbt, SQL, Python, AWS, Terraform, Streamlit, Metabase	
Data Engineer	ago/2024 – jul/2025
Liderei a migração de toda a infraestrutura de dados da empresa para o Databricks: ~50 fontes (Google Analytics 4, sistemas internos e ferramentas de terceiros) em uma única plataforma, com pipelines que passaram de horas para minutos.	

- Construí o data lake e o modelo dimensional da empresa do zero, com Python, dbt e SQL. Até hoje todas as áreas usam essa arquitetura. Também contribuí para o processo regulatório que permitiu à empresa operar legalmente no Brasil.

Tecnologias: Python, dbt, SQL, Databricks, Google Analytics 4

● Clicksign out/2020 – ago/2024

○ Data Team Lead out/2021 – ago/2024

- Um dos primeiros engenheiros do time de dados da Clicksign. Ampliei o grupo de 4 para ~20 pessoas, entre engenharia de dados, análise de dados e garantia de qualidade de dados, e construí os processos que sustentaram esse crescimento.
- Liderei os times de engenharia de dados, análise de dados e garantia de qualidade de dados e fui diretamente responsável pela confiabilidade dos dados usados nas decisões da empresa.

Tecnologias: Amazon Redshift, dbt, Apache Airflow, modelagem dimensional

○ Data Architect jun/2021 – ago/2024

- Arquiteteei praticamente todas as estruturas de dados que a empresa teve naquele período, incluindo 100+ pipelines Airflow (modelagem e ingestão).
- Liderei a implementação do dbt no Redshift, que trouxe controle de qualidade, governança, modelagem mais fácil, observabilidade e orquestração mais simples.
- Construí 10+ APIs para envio de dados a parceiros e terceiros, e arquiteteei os processos de proteção de dados (em conformidade com a LGPD e alinhados ao GDPR) e de continuidade da entrega de dados.

Tecnologias: Apache Airflow, dbt, Amazon Redshift, modelagem dimensional, governança de dados

○ Data Scientist (Python Developer) out/2020 – jun/2021

- Time de Business Intelligence: scripts e integrações entre bancos de dados, APIs e web services REST com Python, PostgreSQL e Apache Airflow, além de análises preditivas e modelagem.

Tecnologias: Python, PostgreSQL, Apache Airflow, REST APIs, modelagem preditiva

● Let's Code from Ada *em paralelo à Clicksign* jun/2022 – jul/2024

○ Professor Docente jun/2022 – jul/2024

- Lecionei cursos de Big Data e Engenharia de Dados para centenas de alunos: modelagem de bancos de dados, data warehouses, data lakes e lakehouses, arquitetura de software e desenvolvimento em cloud.

— FORMAÇÃO

XP Educação — MBA em Engenharia de Dados jul/2022 – jul/2023

PUCRS — Especialização (Lato Sensu) em Ciência de Dados e Inteligência Artificial jun/2021 – jul/2022

UEZO — Bacharelado em Ciência da Computação 2015 – 2019

— CERTIFICAÇÕES

Formação Analista de Dados — Data Science Academy jan/2023

SQL avançado, análise de dados com Python e Power BI sobre data warehouses e data lakes.

Certificação Astronomer: DAG Authoring for Apache Airflow — Astronomer dez/2022

Formação Arquiteto de Dados — Data Science Academy dez/2022

Governança de dados e LGPD, modelagem relacional, NoSQL e data stores, arquitetura de pipelines ETL/ELT e automação com Spark, Airflow, Kafka e AWS Glue.

Certificação Astronomer: Apache Airflow Fundamentals — Astronomer nov/2022

Formação Cientista de Dados — Data Science Academy out/2021

Machine learning aplicado com R, Python e Spark no ecossistema Hadoop, com foco em visualização de dados e storytelling.

Cursos selecionados: Design e Administração de Cloud Computing AWS; Design e Automação de Pipelines de Dados; Bootcamp de Engenharia de Dados.

— PALESTRAS E PESQUISA

Palestra no TDC: Airflow — quem disse que não se pode fazer processamento de dados nele? dez/2021

Iniciação científica na UEZO: técnicas de reconhecimento e detecção; otimização de detecção de bordas em células ago/2018 – ago/2019

— IDIOMAS

Português — Nativo ou bilíngue **Inglês** — Proficiência profissional plena

Espanhol — Proficiência profissional de trabalho