

Franklin Ferreira

Staff Data Engineer

Brasil · Remoto/Global

Arquitectura de Datos · Big Data · Gobernanza · FinOps · ML e IA Generativa

franklinfs390@gmail.com | linkedin.com/in/franklinfs390 | github.com/franklin390 | franklin390.github.io

RESUMEN

Staff Data Engineer y arquitecto de datos. Diseño la arquitectura de datos en la que se apoya toda una empresa, del warehouse al lakehouse, del batch al streaming, del ML clásico a la IA generativa: más de 50% menos costo y ~3x menos incidentes. Sigo hands-on, con un equipo técnico de 15+ personas entre Brasil y Rumania.

>50%	~3x	~50	10+	100+	15+
menos costo de la plataforma de datos	menos incidentes y retrabajo	fuentes unificadas en una sola plataforma	dominios de Unity Catalog	pipelines de Airflow en Clicksign	personas lideradas

TRAYECTORIA

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

COMPETENCIAS

Arquitectura y plataformas: Modelado dimensional, modelado de datos, data warehouse, data lake, lakehouse, Databricks, Delta Lake, Apache Spark, Databricks Workflows, Amazon Redshift, streaming

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

Gobernanza y confiabilidad: Gobernanza de datos, catálogos de datos (Unity Catalog, AWS Glue Data Catalog, Hive Metastore), calidad de datos, LGPD, observabilidad, FinOps

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

Liderazgo: Liderazgo técnico de equipos, arquitectura y formas de trabajo, definición de prioridades

EXPERIENCIA

Flutter Brazil	ago 2024 – Actualidad
Staff Data Engineer	jul 2026 – Actualidad
Plataforma de datos Databricks usada por miles de personas en toda la empresa, con un equipo de 15+ personas (ingeniería, analytics y ciencia de datos) entre Brasil y Rumania.	
- Defino la arquitectura y las formas de trabajo de un equipo de 15+ personas (10 como Tech Lead). - Sigo hands-on en arquitectura, modelado y decisiones de plataforma. Cuando el equipo lo necesita, defino prioridades y destrabo decisiones de negocio.	
Tecnologías: Databricks, Unity Catalog, Delta Lake, Apache Spark, dbt, SQL, Python, AWS, Terraform, GitHub Actions	
Data Engineer – Tech Lead	jul 2025 – jun 2026
Lideré un equipo de 10 personas (ingeniería, analytics y ciencia de datos) en una plataforma Databricks de decenas de TB.	
- Reduje en más de 50% el costo de la plataforma de datos con una iniciativa de FinOps (clústeres, jobs, almacenamiento, gobernanza) y modelos de previsión de costos. Gasto anual: seis cifras (USD). - Reduje los incidentes y el retrabajo ~3x al implementar CI/CD con trunk-based development. - Entregué dashboards de apoyo a la decisión, observabilidad de toda la plataforma y monitoreo de métricas de negocio casi en tiempo real (latencia de pocos minutos), y reestructuré Unity Catalog por dominio (10+). - Definé las arquitecturas y las formas de trabajo de casi todo el equipo de datos. Ayudé a crear los modelos de LTV y payback de la empresa.	
Tecnologías: Databricks, Unity Catalog, Delta Lake, Apache Spark, dbt, SQL, Python, AWS, Terraform, Streamlit, Metabase	
Data Engineer	ago 2024 – jul 2025
Lideré la migración de toda la infraestructura de datos de la empresa a Databricks: ~50 fuentes (Google Analytics 4, sistemas internos y herramientas de terceros) en una sola plataforma, con pipelines que pasaron de horas a minutos.	

- Construí desde cero el data lake y el modelo dimensional de la empresa, con Python, dbt y SQL. Hasta hoy todas las áreas usan esa arquitectura. También contribuí al proceso regulatorio que permitió a la empresa operar legalmente en Brasil.

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

● Clicksign

oct 2020 – ago 2024

○ Data Team Lead

oct 2021 – ago 2024

- Uno de los primeros ingenieros del equipo de datos de Clicksign. Hice crecer el grupo de 4 a ~20 personas, entre ingeniería de datos, análisis de datos y aseguramiento de la calidad de datos, y construí los procesos que sostuvieron ese crecimiento.
- Lideré los equipos de ingeniería de datos, análisis de datos y aseguramiento de la calidad de datos y fui directamente responsable de la confiabilidad de los datos usados en las decisiones de la empresa.

Tecnologías: Amazon Redshift, dbt, Apache Airflow, modelado dimensional

○ Data Architect

jun 2021 – ago 2024

- Diseñé la arquitectura de prácticamente todas las estructuras de datos que la empresa tuvo en ese período, incluidos 100+ pipelines de Airflow (modelado e ingesta).
- Lideré la implementación de dbt en Redshift, que trajo control de calidad, gobernanza, modelado más fácil, observabilidad y una orquestación más simple.
- Construí 10+ APIs para enviar datos a partners y terceros, y diseñé los procesos de protección de datos (en cumplimiento de la LGPD y alineados con el GDPR) y de continuidad de la entrega de datos.

Tecnologías: Apache Airflow, dbt, Amazon Redshift, modelado dimensional, gobernanza de datos

○ Data Scientist (Python Developer)

oct 2020 – jun 2021

- Equipo de Business Intelligence: scripts e integraciones entre bases de datos, APIs y web services REST con Python, PostgreSQL y Apache Airflow, además de análisis predictivos y modelado.

Tecnologías: Python, PostgreSQL, Apache Airflow, REST APIs, modelado predictivo

● Let's Code from Ada *en paralelo a Clicksign*

jun 2022 – jul 2024

○ Profesor docente

jun 2022 – jul 2024

- Enseñé cursos de Big Data e Ingeniería de Datos a cientos de estudiantes: modelado de bases de datos, data warehouses, data lakes y lakehouses, arquitectura de software y desarrollo en cloud.

— FORMACIÓN

XP Educação — MBA en Ingeniería de Datos

jul 2022 – jul 2023

PUCRS — Especialización (Lato Sensu) en Ciencia de Datos e Inteligencia Artificial

jun 2021 – jul 2022

UEZO — Licenciatura en Ciencias de la Computación

2015 – 2019

— CERTIFICACIONES

Formación Analista de Datos — Data Science Academy

ene 2023

SQL avanzado, análisis de datos con Python y Power BI sobre data warehouses y data lakes.

Certificación Astronomer: DAG Authoring for Apache Airflow — Astronomer

dic 2022

Formación Arquitecto de Datos — Data Science Academy

dic 2022

Gobernanza de datos y LGPD, modelado relacional, NoSQL y data stores, arquitectura de pipelines ETL/ELT y automatización con Spark, Airflow, Kafka y AWS Glue.

Certificación Astronomer: Apache Airflow Fundamentals — Astronomer

nov 2022

Formación Científico de Datos — Data Science Academy

oct 2021

Machine learning aplicado con R, Python y Spark en el ecosistema Hadoop, con foco en visualización de datos y storytelling.

Cursos seleccionados: Diseño y administración de AWS Cloud Computing; Diseño y automatización de pipelines de datos; Bootcamp de ingeniería de datos.

— CHARLAS E INVESTIGACIÓN

Charla en TDC: Airflow — ¿quién dijo que no se pueden procesar datos en él?

dic 2021

Investigación de grado en UEZO: técnicas de reconocimiento y detección; optimización de detección de bordes en células

ago 2018 – ago 2019

— IDIOMAS

Portugués — Nativo o bilingüe

Inglés — Competencia profesional completa

Español — Competencia profesional