



ANDERSON BISPO



Engenheiro e Analista de Dados Sênior | Python · SQL · Databricks

Sobre Mim

Profissional de dados sênior com atuação em **engenharia, análise e ciência de dados**. Construo pipelines ETL/ELT em produção — batch e streaming — e transformo dados em **modelos, indicadores e recomendações** para o negócio.

Experiência com **Python, SQL avançado, Airflow, dbt, Kafka e Databricks** (PySpark, Delta Lake), em cloud (GCP em produção, AWS), com data lakes, data warehouses e modelagem dimensional.

Trabalho com **dados financeiros e transacionais de alto volume**, aplicando governança, qualidade, observabilidade e CI/CD, e entrego dashboards em Power BI e modelos preditivos.

Como **docente universitário e CTO de edtech**, traduzo necessidades de negócio em soluções de dados, comunico resultados para a liderança e compartilho conhecimento com o time.



+55 81 99524-1524



anderson.bispox@gmail.com



Caruaru, Pernambuco, Brasil



@anderson-bishopx



@bishoppython

Idiomas

- Português (Nativo)
- Inglês (Profissional)

Formação Acadêmica

Mestrado em Engenharia da Computação

UPE · PPGEC — Universidade de Pernambuco

📅 2024 – 2026

Especialização em Data Science

UNOPAR — Universidade do Norte do Paraná

📅 2021 – 2022

Bacharelado em Sistemas de Informação

UFRPE — Universidade Federal Rural de Pernambuco

📅 2017 – 2022

Técnico em Redes de Computadores

IFPE — Instituto Federal de Pernambuco

📅 2015 – 2016

Experiência Profissional

Histórico completo: linkedin.com/in/anderson-bishopx

Engenheiro de Dados Sênior — Grupo Aposta Ganha | RecebaBet

📍 Caruaru, PE | 📅 2026

Plataforma de dados para uma operação de **alto volume de transações financeiras** (depósitos, saques, saldos e apostas), da ingestão em streaming ao consumo analítico.

- ♦ Arquitetura de **ingestão near real-time** de 29 tópicos Kafka: consumidor Python em Cloud Run → Parquet no **data lake** (GCS) → **data warehouse BigQuery**, integrando também fontes SQL Server.
- ♦ **Modelagem em camadas com dbt e SQL avançado**: tabelas de contas, saldos e transações para indicadores financeiros, com deduplicação idempotente, *unnesting* de JSON e *event sourcing*.
- ♦ **Governança e qualidade**: testes automatizados, documentação e linhagem dos modelos antes da entrega aos times de Analytics, Produto e Negócio.
- ♦ Orquestração com **Airflow** (Cloud Composer) e troubleshooting em produção: falhas de memória, esquema, partição e permissões.
- ♦ **CI/CD** com GitHub Actions e Docker; segurança com Secret Manager, IAM, mTLS e Workload Identity Federation.

Python SQL Kafka Airflow dbt BigQuery Data Lake GCP CI/CD

Competências Técnicas

- **Python**
pandas · numpy · scikit-learn · APIs
- **SQL Avançado**
CTEs · janelas · BigQuery · PostgreSQL · SQL Server
- **Databricks e PySpark**
Delta Lake · Databricks SQL
- **Pipelines ETL / ELT**
Airflow · dbt · Kafka · batch e streaming
- **Cloud**
GCP (produção) · AWS · Azure Data Factory (estudos)
- **Data Lake e Data Warehouse**
Parquet · BigQuery · Snowflake · Medallion
- **Modelagem de Dados**
dimensional · estrela e snowflake
- **BI e Visualização**
Power BI · Streamlit
- **Machine Learning e IA**
modelos preditivos · NLP · LLMs · RAG
- **Governança e DevOps**
qualidade · linhagem · Git · Docker · CI/CD

Competências Comportamentais

- Tradução de negócio em dados
- Comunicação com a liderança
- Pensamento analítico
- Liderança técnica e mentoria
- Colaboração multidisciplinar
- Resolução de problemas
- Aprendizado contínuo
- Autonomia em trabalho remoto

Diferenciais

- **Dados financeiros em alto volume**
depósitos, saques e saldos
- **IA no dia a dia do desenvolvimento**
Claude Code · agentes · RAG
- **Docência e mestrado**
Engenharia de Dados e IA

Analista de Dados Sênior | Engenheiro de Dados Jr. — Azul Linhas Aéreas

📍 Barueri, SP · Híbrido | 📅 2023 – 2025

- ♦ **ETLs em Python e SQL** integrando dados operacionais de voos, com foco em integridade e reprodutibilidade.
- ♦ Análise de **eficiência operacional** e do processo **TAT (Turn-Around-Time)**, com insights estratégicos e ferramentas de automação **apresentados à diretoria** de Operações.
- ♦ **Modelos preditivos** de Machine Learning (scikit-learn) para manutenção de aeronaves, validados antes da implantação.
- ♦ **NLP**, análise de sentimentos e GenAI/LLMs para o Comitê de Recrutamento de Pilotos, em colaboração com a Engenharia de Dados.

Fundador & CTO — NeuroKids (Edtech)

📍 Caruaru, PE | 📅 2026 – atual

- ♦ Defino a **estratégia técnica e a arquitetura da plataforma** (aplicação, dados e IA) de uma edtech SaaS B2B2C.
- ♦ **Arquitetura de indicadores**: métricas de desempenho calculadas de forma determinística, separadas da geração de relatórios por IA; PostgreSQL, FastAPI e infraestrutura própria.
- ♦ **IA generativa com LGPD**: RAG com pgvector, anonimização com NER local e validação com especialistas (metodologia Kappa).

Docente — Centro Universitário UniFavip

📍 Caruaru, PE | 📅 2024 – 2026

- ♦ Leciono **Python, SQL, Big Data, Ciência e Engenharia de Dados, Machine Learning**, Cloud (AWS, Docker) e GenAI.
- ♦ **Mentoria e comunicação didática**: traduzo conceitos complexos em projetos práticos — habilidade que aplico ao apresentar análises e orientar times.

Projetos Relevantes

Pipeline ELT com Modelo Estrela — Airflow · dbt · Power BI

🔗 github.com/bishoppython/pipeline-elt-airflow-dbt · Case técnico

- ♦ Pipeline **ELT ponta a ponta** e containerizado (Docker Compose), de uma base operacional para um data warehouse PostgreSQL.
- ♦ **Modelo estrela em dbt** (raw → staging → analytics) com DAG diária idempotente no Airflow e **20 testes de qualidade**.
- ♦ KPIs de receita, ticket médio e clientes em **Power BI** e dashboard Streamlit.

Resumo do Perfil

Engenharia de dados

Pipelines batch e streaming em produção, data lake e DW.

Análise e BI

KPIs, dashboards em Power BI e insights para a diretoria.

Ciência de dados

Modelos preditivos, NLP e IA generativa.

Cloud e Big Data

GCP em produção, AWS, Databricks e PySpark.

Governança

Qualidade, linhagem, observabilidade e CI/CD.

Comunicação

Docência, mentoria e ponte com o negócio.