

Lucas Cruz

Engenheiro de Dados, com **3 anos de experiência** em soluções de **Big Data** em ambientes de **nuvem (AWS)** e **on-premises**. Possui conhecimento em **Python, SQL, Airflow, Trino** e **Docker**, além de estar em aprendizado contínuo com **Databricks, PySpark, DBT, Kafka & MLflow**. Tem experiência na construção de arquiteturas de **Data Lakes, Data Warehouses** e **Data Lakehouse**, integrando dados de diversas fontes e atuando na camada final de dados com ferramentas de **BI** como **Power BI, Qlik Sense** e **Superset**, entregando análises de alta qualidade e gerando insights para o negócio.

EXPERIÊNCIAS

Spacecom, Brasil — Engenheiro de dados;

04/2023 - 07/2024 (1y3m)

Reuniões com stakeholders, arquitetura e desenvolvimento de **pipelines de dados batch end-to-end**, integrando fontes como **PostgreSQL, SQL Server, APIs** e arquivos **Excel/CSV** com **Python** e **SQL**. Construí e mantive uma plataforma de **Big Data Lakehouse** em **AWS S3** com o modelo **Medalhão (Bronze, Silver, Gold)** e **Delta parquet**, processando dados de mais de 90 mil dispositivos IoT e centenas de milhões de registros diários. Implementei transformações com **PySpark** e **Pandas**, virtualizei tabelas com **TrinoSQL** e orquestrei fluxos de ETL com **Apache Airflow**. O deploy foi feito com **Docker** em **AWS EC2**, e desenvolvi dashboards em **QlikSense** e **Apache Superset** para análise de dados.

wDiscovery, Brasil — Engenheiro de dados (Consultor);

06/2022 - 05/2024 (1y11m)

Reuniões com Stakeholders, desenvolvimento de **projetos de dados end-to-end** para ingestão de dados de máquinas de cartão (PagSeguro, Stone e Cielo) e máquinas agrícolas (John Deere), integrando múltiplas fontes como **APIs, SQL Server, PostgreSQL**, arquivos **Excel/CSV** e **web scraping**, utilizando **Python** e **SQL**. Criação de **Data Warehouses** em **SQL Server** e **Oracle Autonomous Database**, com processos de limpeza, transformação e carga de dados utilizando **Pandas**. Orquestrei os fluxos de ETL com **Apache Airflow** e realizei o deploy dos projetos em **VMs da Oracle Cloud** utilizando **Docker**. Também desenvolvi dashboards em **Power BI** para suporte às análises de negócio.

Grupo Itaeté, Brasil — Analista de dados;

06/2020 - 02/2023 (2y8m)

Desenvolvimento de **projeto de dados** para coleta de dados de dispositivos IoT, **PostgreSQL, SQL Server, APIs** e arquivos **Excel/CSV**, utilizando **Python** e **SQL**. Processamento e tratamento dos dados com **Pandas** e **construção de Data Warehouse em SQL Server**. Consumo dos dados por meio de **Power BI** e **Excel**. Desenvolvimento de **sistema interno para controle de estoque** e **automação de processos financeiros**, incluindo emissão de notas fiscais e envio automático de dashboards semanais para clientes. Atuação em **reuniões e suporte a clientes**, garantindo a qualidade da entrega dos dados e o gerenciamento da plataforma **IoT**.

EDUCAÇÃO

Uni Dom Bosco, Brasil — Análise e Desenvolvimento de Sistemas

01/2021 - 08/2023 (2y6m)

Curitiba, PR - BR

+44 07512 225720 (Whatsapp)

dev.lucasacruz@gmail.com

<https://github.com/devzurc>

<https://linkedin.com/in/lucas-cruz>

SKILLS

[0~1 y exp]:

- Databricks;
- DBT;
- PySpark;
- Kafka;
- MLFlow;

[1~2y exp]:

- Power B.I;
- Qlik Sense;
- Apache Superset;
- Apache Airflow;
- Trino/Presto;

[2~3y exp]:

- Git;
- Docker;
- Scrum & Kanban;
- Cloud; (AWS & Oracle);

[3~4y exp]:

- Python;
- SQL & NoSQL;
- Bash scripts;

[4y+ exp]:

- Linux/Windows
- Advanced MS Office;

IDIOMAS

- Portugues (Nat);
- Inglês (C1);
- Espanhol (B1);

