

NICOLAS MARTINS SILVA

Data Engineer

 Aveiro, Portugal

 Disponível para oportunidades Remotas / Híbridas na Europa

 nicolasn892@gmail.com

 +351 924 914 744

 LinkedIn |  GitHub |  Portfólio

RESUMO PROFISSIONAL

Engenheiro da Computação especializado em **Engenharia de Dados**, com sólida experiência na construção de **pipelines ETL/ELT escaláveis, Data Warehousing e automação de processos**. Possui forte domínio de **Python, SQL e Apache Spark (PySpark)**, com experiência prática em **modelagem dimensional (Star Schema/Kimball), integrações com APIs e processamento de dados em larga escala utilizando Databricks**. Atualmente atua como **Estagiário em Engenharia de Dados no Grupo Fleury**, contribuindo para soluções de dados em ambiente produtivo. Residente em Portugal e disponível para oportunidades em toda a Europa.

COMPETÊNCIAS TÉCNICAS

Linguagens & Programação

- Python (Pandas, NumPy, Requests, Programação Orientada a Objetos)
- SQL (Avançado)
- Bash / Shell

Big Data & Processamento

- Apache Spark (PySpark)
- Padrões de Arquitetura ETL / ELT
- Databricks

Bancos de Dados & Data Warehousing

- PostgreSQL, SQLite
- Modelagem Dimensional (Star Schema – Kimball)
- Data Lakes (Parquet)

Ferramentas & DevOps

- Git & GitHub
- Docker
- Ambiente Linux

- VS Code

Analytics & BI

- Power BI
 - Análise Exploratória de Dados (EDA)
 - Data Storytelling
-

EXPERIÊNCIA PROFISSIONAL

Estagiário em Engenharia de Dados | Grupo Fleury

Maio 2023 – Atual

- Atuação no desenvolvimento e manutenção de **pipelines de dados** utilizando Python, SQL e Databricks.
 - Trabalho com **grandes volumes de dados** nas áreas jurídica e de saúde, garantindo **qualidade, desempenho e confiabilidade** das informações.
 - Realização de **integrações com APIs SOAP e REST**, incluindo paginação, mecanismos de retry e tolerância a falhas.
 - Apoio em iniciativas de **modelagem de dados** e construção de datasets prontos para consumo analítico.
 - Colaboração com times multidisciplinares em ambiente ágil, entregando soluções de dados escaláveis.
-

PROJETOS DE ENGENHARIA (DESTAQUE)

Código-fonte completo e documentação disponíveis em: nicolasn892.github.io/portfolio

Pipeline de Dados Meteorológicos – ETL com PySpark

Desenvolvimento de um pipeline ETL de ponta a ponta para ingestão e processamento de dados meteorológicos.

- **Arquitetura:** Implementação de uma camada de extração robusta consumindo a API REST do IPMA, com tratamento de dados JSON semiestruturados.
- **Processamento:** Aplicação de transformações com PySpark, incluindo definição de schema, explosão de arrays e conversão de tipos de dados.
- **Armazenamento:** Estratégia de Data Lake com arquivos Parquet particionados por data, otimizando o desempenho de consultas analíticas.

Tecnologias: Python, PySpark, REST APIs, JSON, Parquet



Data Warehouse de Turismo – SQL & Modelagem Dimensional

Desenvolvimento de um Data Warehouse completo para análise de métricas de turismo em Portugal, utilizando Star Schema.

- **Modelagem:** Criação de tabelas Fato e Dimensão (fact_tourism, dim_location, dim_time) seguindo a metodologia Kimball.
- **Implementação:** Transformação de dados brutos em Excel para um banco relacional SQLite, assegurando integridade referencial e qualidade dos dados.
- **Análises:** Possibilitou consultas SQL complexas para análise de receitas, sazonalidade e desempenho regional.

Tecnologias: SQL, Python, Pandas, SQLite, Modelagem de Dados



Bot Automatizado de Ingestão de Dados

Criação de uma aplicação Python orientada a objetos para automatizar fluxos de ingestão de arquivos corporativos.

- **Automação:** Desenvolvimento de um bot para monitoramento de diretórios, detecção de arquivos (Excel, CSV, JSON) e padronização de nomenclaturas.
- **Eficiência:** Automação da conversão de formatos proprietários (.xlsx) para formatos abertos (.csv), reduzindo esforço manual e tempo de preparação dos dados.

Tecnologias: Python (OOP), Watchdog, Pathlib, Pandas

FORMAÇÃO ACADÊMICA

Bacharelado em Engenharia da Computação

Universidade do Estado de Minas Gerais (UEMG) – Brasil

Ênfase em Engenharia de Software, Bancos de Dados, Algoritmos e Processamento de Dados

IDIOMAS

- **Português:** Nativo
- **Inglês:** Intermediário (B1/B2)