

NICOLAS MARTINS SILVA

Data Engineer

-  Aveiro, Portugal
 -  Open to Remote / Hybrid opportunities across Europe
 -  nicolasn892@gmail.com
 -  +351 924 914 744
 -  [LinkedIn](#) |  [GitHub](#) |  [Portfolio](#)
-

PROFESSIONAL SUMMARY

Computer Engineer specialized in **Data Engineering**, with strong experience in building **scalable ETL/ELT pipelines**, **Data Warehousing**, and **process automation**. Proficient in **Python**, **SQL**, and **Apache Spark (PySpark)**, with hands-on experience in **dimensional modeling (Star Schema/Kimball)**, **API integrations**, and **large-scale data processing using Databricks**. Currently working as a **Data Engineering Intern at Grupo Fleury**, contributing to real-world, production-grade data solutions. Based in Portugal and available for opportunities across Europe.

TECHNICAL SKILLS

Languages & Programming

- Python (Pandas, NumPy, Requests, OOP)
- SQL (Advanced)
- Bash / Shell

Big Data & Processing

- Apache Spark (PySpark)
- ETL / ELT Design Patterns
- Databricks

Databases & Warehousing

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

Tools & DevOps

- Git & GitHub
- Docker
- Linux Environment

- VS Code

Analytics & BI

- Power BI
 - Exploratory Data Analysis (EDA)
 - Data Storytelling
-

PROFESSIONAL EXPERIENCE

Data Engineering Intern | Grupo Fleury

May 2023 – Present

- Contribute to the development and maintenance of **data pipelines** using Python, SQL, and Databricks.
 - Work with **large-scale judicial and healthcare-related datasets**, ensuring data quality, performance, and reliability.
 - Perform **API integrations (SOAP and REST)** with pagination, retries, and fault tolerance mechanisms.
 - Support **data modeling initiatives** and analytics-ready datasets for business consumption.
 - Collaborate with cross-functional teams in an agile environment to deliver scalable data solutions.
-

ENGINEERING PROJECTS (FEATURED)

Full source code and documentation available at: nicolasn892.github.io/portfolio



Weather Data Pipeline – ETL with PySpark

Designed and implemented an end-to-end ETL pipeline to ingest and process meteorological data.

- **Architecture:** Built a robust extraction layer consuming the IPMA REST API, handling semi-structured JSON data.
- **Processing:** Applied PySpark transformations including schema enforcement, array explosion, and data type casting.
- **Storage:** Implemented a Data Lake strategy using partitioned Parquet files by date to optimize analytical queries.

Tech Stack: Python, PySpark, REST APIs, JSON, Parquet



Tourism Data Warehouse – SQL & Dimensional Modeling

Developed a complete Data Warehouse solution to analyze tourism metrics in Portugal using a Star Schema.

- **Modeling:** Designed Fact and Dimension tables (fact_tourism, dim_location, dim_time) following Kimball methodology.
- **Implementation:** Transformed raw Excel datasets into a relational SQLite database, ensuring data quality and referential integrity.
- **Analytics:** Enabled complex SQL queries to analyze revenue trends, seasonality, and regional performance.

Tech Stack: SQL, Python, Pandas, SQLite, Data Modeling

Automated Data Ingestion Bot

Created an object-oriented Python application to automate corporate file ingestion workflows.

- **Automation:** Developed a monitoring bot to detect new files (Excel, CSV, JSON) and standardize naming conventions.
- **Efficiency:** Automated conversion from proprietary formats (.xlsx) to open formats (.csv), reducing manual effort and processing time.

Tech Stack: Python (OOP), Watchdog, Pathlib, Pandas

EDUCATION

Bachelor's Degree in Computer Engineering

Universidade do Estado de Minas Gerais (UEMG) – Brazil

Focus on Software Engineering, Databases, Algorithms, and Data Processing

LANGUAGES

- **Portuguese:** Native
- **English:** Intermediate (B1/B2)

Entao vamos la , adicionei o curriculo em ingles , porem os links estao quebrados no index e somente o linkedin esta azul , queria tambem um botao a mais o qual daria a opçao da pessoa fazer o dowload do meu curriculo em portugues