

Bruno Oliveira

Data Engineer | Databricks • Azure • PySpark

📍 Americana – SP (or remote) • 📞 +55 (19) 9.9850-8547 • 📩 brunoolivei@protonmail.com

🔗 LinkedIn: <https://www.linkedin.com/in/brunoolivei> | GitHub: <https://github.com/BrunoOlivei>

Summary

Senior Data Engineer specialized in designing and operating ingestion pipelines (batch/daily) on Databricks + Azure Data Factory, building Data Lakehouse architectures (Bronze/Silver/Gold). Hands-on experience in end-to-end ETL/ELT: connectors for APIs/files/databases, data quality (validations, schema drift, deduplication), governance (tagging, anonymization, versioning), documentation/KT, and monitoring/alerts.

Key achievement: Led reconciliations and monitoring that identified a business rule flaw impacting ~2.3k records over 7 months. Corrected ISS/IR calculations, normalized historical data, and implemented preventive checks to avoid recurrence.

Professional Experience

Senior Data Engineer

NAVA — Full-time | Jan/25 – Present

- Built data reconciliations and daily monitoring (Databricks/PySpark), identifying rule inconsistencies affecting ~2.3k records over 7 months; corrected financial calculations, normalized history, and implemented preventive validations.
- Designed and orchestrated ingestion pipelines (Databricks + ADF), integrating CSV/XLSX/APIs; ETL/ELT across Bronze/Silver/Gold layers with governance (tagging, anonymization, versioning) and quality checks.
- Automated remuneration processes by migrating Excel spreadsheets to PySpark notebooks, applying business rules (pro-labore, brokerage, co-brokerage), with calculated columns, unit tests, email alerts, and Power BI exposure.
- Developed cancellation (refund) logic with multiple joins, rule hierarchies, and historical/audit tables; reconciliations uncovered errors in transactional systems.
- Supported deployments/PRs, troubleshooting ingestion workflows, and created validation notebooks for business teams.

Back-end Developer

DMX Design & Development — Full-time | Oct/24 – Dec/24

- Led MVP development with LLMs for salon scheduling via WhatsApp, personalized offers, and microservices architecture.

- Integrated Twilio and Asaas, improving operational efficiency while optimizing infrastructure costs.
- Implemented user preference learning for tailored services, enhancing customer engagement.
- Adopted unit testing with Jest, ensuring code reliability and maintainability.

Data Engineer

Qyon — Full-time | Mar/22 – Oct/24

- Developed an automated scheduling system (FastAPI, SQLAlchemy, MySQL, MongoDB, RabbitMQ, WhatsApp API, Microsoft Teams), reducing client onboarding time from 122 to 22 days.
- Built daily data pipelines in AWS EKS, collecting customer usage data across PostgreSQL, MySQL, SQL Server, DynamoDB, and APIs with Python/asyncio/pandas/numpy, exposing results via APIs and Power BI.
- Migrated workloads to Databricks + Airflow, creating a medallion architecture for lead enrichment (public Receita Federal data, RD Station), feeding CRM with enriched leads aligned to business strategy.

Education

B.Sc. in Data Science (in progress)

Faculdade Descomplica — Expected graduation: Jul/2028

Additional Information

Technical Skills

Databricks, Azure Data Factory, PySpark, Python, SQL, Power BI, Data Lakehouse (Bronze/Silver/Gold), PostgreSQL, MySQL, SQL Server, DynamoDB, MongoDB, Git, Docker, Agile Development, CI/CD, Unit Testing

Languages

English – Intermediate

Certifications & Courses

- Workshop Databricks – ETL from Scratch (Jornada de Dados, 2024)
- Python Sudeste 2024 (UFMG, 2024)
- Python for Data Science: Intro to Numpy/Pandas (Alura, 2023)
- Design Patterns in Python I (Alura, 2023)
- Power BI: DAX & Reports (Alura, 2023)

Soft Skills

Adaptability in dynamic environments • Critical thinking & business acumen for process optimization • Excellent communication & teamwork across multidisciplinary teams